Mathematical Methods for IoT-based Signal and Image Processing Applications

Lead Guest Editor: Savita Gupta
Guest Editors: Neelam Goel, Yanhui Guo, and Yu-Chen Hu
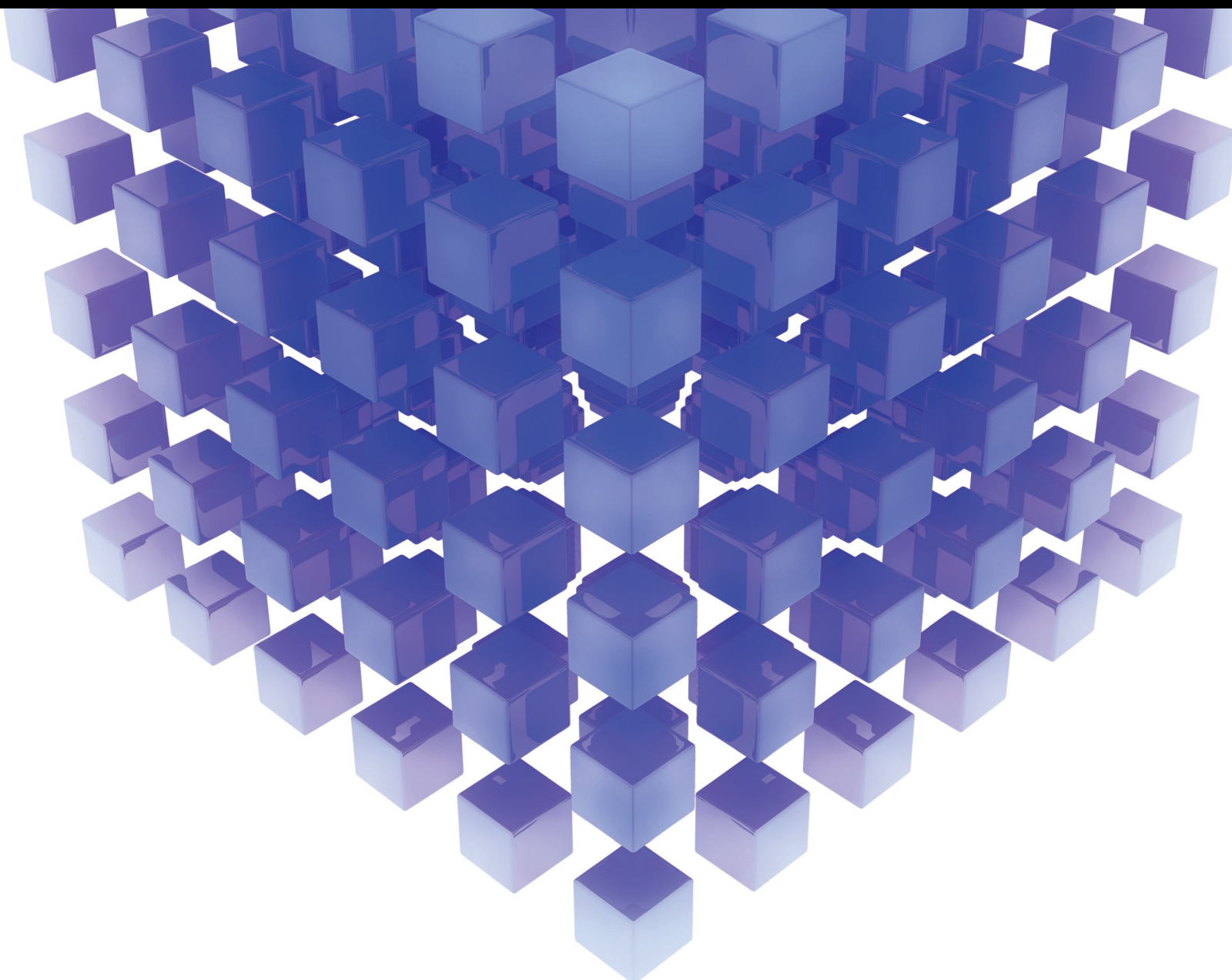
Mathematical Methods for IoT-based Signal and Image Processing Applications

Mathematical Methods for IoT-based Signal and Image Processing Applications

Lead Guest Editor: Savita Gupta
Guest Editors: Neelam Goel, Yanhui Guo, and Yu-Chen Hu

Paolo Boscariol, Italy
Daniela Boso, Italy
Guillermo Botella-Juan, Spain
Abdesselem Boulkroune, Algeria
Boulaïd Boulkroune, Belgium
Fabio Bovenga, Italy
Francesco Braghin, Italy
Ricardo Branco, Portugal
Julien Bruchon, France
Matteo Bruggi, Italy
Michele Brun, Italy
Maria Elena Bruni, Italy
Maria Angela Butturi, Italy
Bartłomiej Błachowski, Poland
Dhanamjayulu C, India
Raquel Caballero-Águila, Spain
Filippo Cacace, Italy
Salvatore Caddemi, Italy
Zuowei Cai, China
Roberto Caldelli, Italy
Francesco Cannizzaro, Italy
Maosen Cao, China
Ana Carpio, Spain
Rodrigo Carvajal, Chile
Caterina Casavola, Italy
Sara Casciati, Italy
Federica Caselli, Italy
Carmen Castillo, Spain
Inmaculada T. Castro, Spain
Miguel Castro, Portugal
Giuseppe Catalanotti, United Kingdom
Alberto Cavallo, Italy
Gabriele Cazzulani, Italy
Fatih Vehbi Celebi, Turkey
Miguel Cerrolaza, Venezuela
Gregory Chagnon, France
Ching-Ter Chang, Taiwan
Kuei-Lun Chang, Taiwan
Qing Chang, USA
Xiaoheng Chang, China
Prasenjit Chatterjee, Lithuania
Kacem Chehdi, France
Peter N. Cheimets, USA
Chih-Chiang Chen, Taiwan
He Chen, China

Kebing Chen, China
Mengxin Chen, China
Shyi-Ming Chen, Taiwan
Xizhong Chen, Ireland
Xue-Bo Chen, China
Zhiwen Chen, China
Qiang Cheng, USA
Zeyang Cheng, China
Luca Chiapponi, Italy
Francisco Chicano, Spain
Tirivanhu Chinyoka, South Africa
Adrian Chmielewski, Poland
Seongim Choi, USA
Gautam Choubey, India
Hung-Yuan Chung, Taiwan
Yusheng Ci, China
Simone Cinquemani, Italy
Roberto G. Citarella, Italy
Joaquim Ciurana, Spain
John D. Clayton, USA
Piero Colajanni, Italy
Giuseppina Colicchio, Italy
Vassilios Constantoudis, Greece
Enrico Conte, Italy
Alessandro Contento, USA
Mario Cools, Belgium
Gino Cortellessa, Italy
Carlo Cosentino, Italy
Paolo Crippa, Italy
Erik Cuevas, Mexico
Guozeng Cui, China
Mehmet Cunkas, Turkey
Giuseppe D'Aniello, Italy
Peter Dabnichki, Australia
Weizhong Dai, USA
Zhifeng Dai, China
Purushothaman Damodaran, USA
Sergey Dashkovskiy, Germany
Adiel T. De Almeida-Filho, Brazil
Fabio De Angelis, Italy
Samuele De Bartolo, Italy
Stefano De Miranda, Italy
Filippo De Monte, Italy

Yong (Aaron) Tan, United Kingdom
Marco Antonio Taneco-Hernández (iD), Mexico
Lu Tang (iD), China
Tianyou Tao, China
Hafez Tari (iD), USA
Alessandro Tasora (iD), Italy
Sergio Teggi (iD), Italy
Adriana del Carmen Téllez-Anguiano (iD), Mexico
Ana C. Teodoro (iD), Portugal
Efstathios E. Theotokoglou (iD), Greece
Jing-Feng Tian, China
Alexander Timokha (iD), Norway
Stefania Tomasiello (iD), Italy
Gisella Tomasini (iD), Italy
Isabella Torcicollo (iD), Italy
Francesco Tornabene (iD), Italy
Mariano Torrisi (iD), Italy
Thang nguyen Trung, Vietnam
George Tsiatas (iD), Greece
Le Anh Tuan (iD), Vietnam
Nerio Tullini (iD), Italy
Emilio Turco (iD), Italy
Ilhan Tuzcu (iD), USA
Efstratios Tzirtzilakis (iD), Greece
FRANCISCO UREÑA (iD), Spain
Filippo Ubertini (iD), Italy
Mohammad Uddin (iD), Australia
Mohammad Safi Ullah (iD), Bangladesh
Serdar Ulubeyli (iD), Turkey
Mati Ur Rahman (iD), Pakistan
Panayiotis Vafeas (iD), Greece
Giuseppe Vairo (iD), Italy
Jesus Valdez-Resendiz (iD), Mexico
Eusebio Valero, Spain
Stefano Valvano (iD), Italy
Carlos-Renato Vázquez (iD), Mexico
Martin Velasco Villa (iD), Mexico
Franck J. Vernerey, USA
Georgios Veronis (iD), USA
Vincenzo Vespri (iD), Italy
Renato Vidoni (iD), Italy
Venkatesh Vijayaraghavan, Australia

Anna Vila, Spain
Francisco R. Villatoro (iD), Spain
Francesca Vipiana (iD), Italy
Stanislav Vítek (iD), Czech Republic
Jan Vorel (iD), Czech Republic
Michael Vynnycky (iD), Sweden
Mohammad W. Alomari, Jordan
Roman Wan-Wendner (iD), Austria
Bingchang Wang, China
C. H. Wang (iD), Taiwan
Dagang Wang, China
Guoqiang Wang (iD), China
Huaiyu Wang, China
Hui Wang (iD), China
J.G. Wang, China
Ji Wang (iD), China
Kang-Jia Wang (iD), China
Lei Wang (iD), China
Qiang Wang, China
Qingling Wang (iD), China
Weiwei Wang (iD), China
Xinyu Wang (iD), China
Yong Wang (iD), China
Yung-Chung Wang (iD), Taiwan
Zhenbo Wang (iD), USA
Zhibo Wang, China
Waldemar T. Wójcik, Poland
Chi Wu (iD), Australia
Qiuhong Wu, China
Yuqiang Wu, China
Zhibin Wu (iD), China
Zhizheng Wu (iD), China
Michalis Xenos (iD), Greece
Hao Xiao (iD), China
Xiao Ping Xie (iD), China
Qingzheng Xu (iD), China
Binghan Xue (iD), China
Yi Xue (iD), China
Joseph J. Yame (iD), France
Chuanliang Yan (iD), China
Xinggang Yan (iD), United Kingdom
Hongtai Yang (iD), China
Jixiang Yang (iD), China
Mijia Yang, USA
Ray-Yeng Yang, Taiwan

Zaoli Yang (iD), China
Jun Ye (iD), China
Min Ye (iD), China
Luis J. Yebra (iD), Spain
Peng-Yeng Yin (iD), Taiwan
Muhammad Haroon Yousaf (iD), Pakistan
Yuan Yuan, United Kingdom
Qin Yuming, China
Elena Zaitseva (iD), Slovakia
Arkadiusz Zak (iD), Poland
Mohammad Zakwan (iD), India
Ernesto Zambrano-Serrano (iD), Mexico
Francesco Zammori (iD), Italy
Jessica Zangari (iD), Italy
Rafal Zdunek (iD), Poland
Ibrahim Zeid, USA
Nianyin Zeng (iD), China
Junyong Zhai (iD), China
Hao Zhang (iD), China
Haopeng Zhang (iD), USA
Jian Zhang (iD), China
Kai Zhang, China
Lingfan Zhang (iD), China
Mingjie Zhang (iD), Norway
Qian Zhang (iD), China
Tianwei Zhang (iD), China
Tongqian Zhang (iD), China
Wenyu Zhang (iD), China
Xianming Zhang (iD), Australia
Xuping Zhang (iD), Denmark
Yinyan Zhang, China
Yifan Zhao (iD), United Kingdom
Debao Zhou, USA
Heng Zhou (iD), China
Jian G. Zhou (iD), United Kingdom
Junyong Zhou (iD), China
Xueqian Zhou (iD), United Kingdom
Zhe Zhou (iD), China
Wu-Le Zhu, China
Gaetano Zizzo (iD), Italy
Mingcheng Zuo, China

Contents

Retraction

Retracted: The Use of Artificial Intelligence for Smart Decision-Making in Smart Cities: A Moderated Mediated Model of Technology Anxiety and Internal Threats of IoT

Mathematical Problems in Engineering

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] A. Alloulbi, T. Öz, and A. Alzubi, "The Use of Artificial Intelligence for Smart Decision-Making in Smart Cities: A Moderated Mediated Model of Technology Anxiety and Internal Threats of IoT," *Mathematical Problems in Engineering*, vol. 2022, Article ID 6707431, 12 pages, 2022.

Research Article

Maximize the Security for Image Processing Using an Improved Watermarking Approach

Manoj Kumar Tyagi,[1] Cuddapah Anitha,[2] R. Ramyadevi,[3] Sunita Pachar,[4] Ravi Kumar Tata,[5] Preetam Suman,[6] and Fardin Ahmadi [7]

[1]*Computer Science and Information Technology, KIET Group of Institutions, Delhi-NCR, Muradnagar, Ghaziabad, Uttar Pradesh, India*
[2]*School of Computing, Computer Science and Engineering, Mohan Babu University (Erstwhile Sree Vidyaniketan Engineering College), Tirupati 517102, Andhra Pradesh, India*
[3]*Department of Computer Science and Engineering, Saveetha Engineering College, Chennai, India*
[4]*IBM Department, GLA University, Mathura, Uttar Pradesh, India*
[5]*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India*
[6]*School of Computing Science and Engineering, VIT Bhopal University, Bhopal-Indore Highway, Kothrikalan, Sehore, Madhya Pradesh 466114, India*
[7]*Rana University, Kabul, Afghanistan*

Correspondence should be addressed to Fardin Ahmadi; fardin.ahmadi@bcs.ru.edu.af

Creating a dependable and effective digital image watermarking (WM) method necessitates balancing imperceptibility, resilience, capacity, and security. Several studies have combined spatial and transform domains to meet these needs. We investigated the current state of hybrid digital image WM. When developing a hybrid WM strategy, it is critical to consider software compatibility. Following a brief review of the literature, we used a table to analyse and evaluate current hybrid approaches. Furthermore, the limitations and possibilities of these approaches are discussed. We investigated the limitations of current research methodologies and proposed new research directions. The internet's role in distributing digital material has resulted in more digitalisation and more complex copyright concerns. Copyright breaches are being reduced through the deployment of innovative digital WM techniques. As a result of this research, a WM system capable of dealing with a wide range of threats is being developed. The goal of this study is to maximise security while maintaining visibility and resilience. Discrete wavelets transform (DWT) and singular value decomposition (SVD) were used to investigate covert operations; DWT was used to isolate each level of the host image, which was then processed using SVD. At the time, normalized correlation was the best method for measuring watermarked images. In research, DWT–SVD was more accurate in detecting assaults; it shows structural similarity index and normalized correlation are 98.4 and 98.35, respectively. Watermarked images can withstand a variety of assaults due to their invisibility and resilience; this system repels assaults that alter pixel values better than traditional methods.

1. Introduction

The following sections summarise the main points, every day, multimedia advances without sacrificing quality, digital images can be instantly modified, copied, reproduced, and transmitted over local networks and the Internet. The multi-media revolution has an impact on image security and privacy. Digital image watermarking (WM) has become an important way to protect multimedia content and determine who owns its use in recent years. The watermark information is implanted in a media item (text, image, sound, and video), then recovered or recognised to approve the item's genuineness [1]. An eavesdropper cannot change or replace the watermark data in the host data, thus protecting it. By using this strategy, you can be confident that the content you are using is genuine, that its integrity has been verified, and that your images have

been safeguarded. A watermark safeguards the image quality and aesthetic appeal of the cover. Several spatial or transform-based approaches have been developed in recent years. The study made the following contributions: we exposed flaws in existing hybrid digital image WM methods. The protection of sensitive electronic information has been and will continue to be one of the most important problems facing research in the scientific community. Because of the proliferation of Internet-connected gadgets, it is now simpler to make unauthorised copies of digital information, authenticate it, and distribute it to third parties. These applications include broadcasting and tracking. Since electronic information has been widely published and distributed via the Internet, there has been an increase in the frequency of various violations of copyright, including unauthorised use and copying, as well as theft of online content. The Internet is the distinguishing technology of the information age, just like the electrical engine was the push for technical growth during the industrial age. Users may engage in a wide range of communications in several ways from any location on Earth due to the linked structure of these networks and the fact that the bulk of them are backed by technology that allows wireless transmission. This immediately affects a user's ability to connect to the internet. The Internet is the distinguishing technology of the information age, just like the electrical engine was the push for technical growth during the industrial age. Users may engage in a wide range of communications in several ways from any location on Earth due to the linked structure of these networks and the fact that the bulk of them are backed by technology that allows wireless transmission. This immediately affects a user's ability to connect to the internet. If digital images are ever going to be worth a significant amount of money in the future, it is imperative that their intellectual property rights be safeguarded. It is possible to safeguard these images by utilising the cutting-edge method of digital WM [2, 3], which is not only efficient but also feasible using the WM approach; any data belonging to the owner that is going to be preserved or shared over the internet may be encrypted. The demonstration of ownership may be accomplished by the acquisition of encoded watermark data at the appropriate time using a variety of technologies, application domains, and many other online platform-based systems. There have been very few investigations exploring the use of digital WM as a technique of protecting images that were captured with a digital camera. Over the course of the past several years, several various alternative WM techniques have been created and implemented. The WM of a digital image may be extracted with the use of algorithms, which can then be modified [3–6]. In the context of WM in general, attacks that seek to erase or otherwise damage encoded watermark data could be a real possibility. These kinds of attacks might be a concern. There is a possibility that the WM process might be affected inadvertently by actions that are essential to the maintenance or distribution of the content. The word "unintentional" could be used to describe attacks of this kind. This may be achieved in one of two ways: either computationally or physically. Both approaches are viable options. However, to extract WM, it is necessary to fulfil several additional prerequisite requirements. Both intentional and

accidental assaults have the potential to corrupt the data kept on the host as well as the data that is encoded in WM. Therefore, it's possible that taking a statistical approach will produce better results than trying to extract WM in an organised and planned method.

To create a form of semisequence that may be used to integrate many watermark images into one, a cryptographic approach that uses pseudorandom data may be utilised to build the sequence. As a watermark, you have the option of using either a monochromatic or two-color image [7–11]. Approaches for the protection of digital material can be implemented in WM methods. Two examples of these techniques are scatter-spectral (SS) and quantisation index manipulation (QIM). The abbreviations SS and QIM refer to the procedures of SS manipulation and QIM, respectively. Both alternative and cumulative methods are referred to in a manner that is virtually equivalent when they are referred to in this manner. Spatial domain watermarking (STDM) is a procedure that takes both the dependability of SS and the efficiency of QIM and combines them. Systems for WM anything need to have a payload in addition to being reliable and valid. Images that were processed in the spatial domain utilising filtering techniques such as brightness, softness, and noise reduction are substantially more trustworthy than those that were processed using discrete cosine transform (DCT). The ability of discrete wavelets transform (DWT) to edit digital images that have been watermarked has contributed to the broad use of the tool. An initial set of processes will work their way down, beginning with the DWT domain [12–15], going from the domain with the highest resolution all the way down to the domain with the lowest resolution. When creating digital images, it is important to make sure that the intensity of the watermark masking is raised throughout the process. This ensures that the product will last for a longer period. Currently, the use of singular value decomposition (SVD) as a method for adding a watermark to digital images is already regarded to be standard practise. Even if the user is attacked in a variety of ways while SVD is running, the concealed image will not be corrupted since it will still be protected. Companies are already utilising testing platforms to develop algorithms that can interact with their surroundings in a manner that is more helpful and to glean insights that have never been discovered before.

The removal of noise and an enhancement in the image's overall clarity are two examples of how digital image processing may enhance the overall quality of an image. The use of neural networks to construct mappings between clear and distorted images to enhance deblurring has been increasingly popular in recent years [16, 17]. To improve the effectiveness of the deblurring process, this step has been taken. The images that were sharp and clear were contrasted with those that were fuzzy or otherwise misleading in some other way. Convolutional neural network (CNN) has the potential to offer excellent results in image denoising because of its vast modelling capacity as well as major network and architectural advancements. CNN makes use of a deep structure model, which can make greater use of visual information in both its training and its noise reduction processes. To

do this, the use of a CNN is necessary; batch normalisation and the rectified linear unit, both of which have witnessed significant advancements in recent years, are currently considered to be two of the most essential learning approaches to produce CNNs, as well as the teaching of those methodologies. It is dangerous to employ this method of reducing noise on an image that has a watermark since the watermark itself is nothing more than a string of noise. Digitized watermark images, particularly those whose level of noise has been increased or lowered, offer a lot of room for attack [18–20]. This is especially true for images in which the noise level has been altered. Because of this, the starting value of each pixel in an image is the same for both procedures. An image is composed of a very large number of individual dots called pixels. Researchers conducted tests using Fully Convolutional Neural Networks (FCNNs) to see whether the FCNNs could recognise watermarks in images that already included watermarks. In the course of this study, a noteworthy discovery emerged: the normalization of denoising could potentially serve as an independent avenue for launching attacks when implementing computerized digital WM on images.

Images that have been digitally watermarked are susceptible to a broad variety of different attacks. These assaults make use of a range of tactics, some of which include backdrop removal, geometric alteration, loss compression, and additive noise. Any one of these methods will do the trick in removing watermarks from digital images. The properties of the noise include both salt-and-pepper noise as well as multiplicative noise. The most common type of attack used against digital WM images is one that involves the introduction of Gaussian noise. The use of salt-and-pepper sounds may take a greyscale image with 8 bits that appear to be white and white. The build-up of Gaussian noise results in a decline in both the image's quality and its look. Image filtering methods such as averaging, Wiener, median, and Gaussian filtration can be utilised to remove a watermark from a digital image. Other techniques of image filtration, such as median and Gaussian filtration, are also available. The median filter is an adaptive nonlinear filtering method that may be utilised to retain the borders of the image while simultaneously minimising the amount of noise that is present in the image. The Wiener filter is widely utilised as a means of reducing the overall impression of blur in the images they capture. When an averaging filter is used, the value of each pixel is replaced with a weighted average of the values of its near neighbours as well as itself. This new value is then used to represent the pixel. As a direct consequence of this, there is less variation in the brightness of individual pixels. It is usual practise to blur images, as doing so lessens both the intensity of the image and the amount of distortion it includes when using a Gaussian filter to process the image. Image modification techniques such as scaling, rotating, clipping, and translating may be used to change images that were created using geometric assaults [21, 22]. Other image modification techniques that can be used include skewing and blurring. It is possible to speak about "local" geometric assaults while referring to both "local" and

"global" attacks. Attackers armed with chainsaws are only able to completely wipe off a small area of the final image.

To produce a wavelet tree quantisation-based blind image that is more resistant to geometric distortions like rotation, scaling, or cutting, stenographic techniques were utilised in the creation process. This method was developed to address the challenges using this strategy; it is simple to avoid being caught by cropping and rotating assaults. A method of image WM that is immune to several forms of cryptographic assaults, such as translation and rotation [23]. His method is described in more detail in the reference. Together with his colleagues, he came up with a method for mathematically sound image WM [24] that was founded on the change of histograms. This method was mathematically sound. This method considers the image's translations as part of the total transformation, in addition to rotation, cropping, and scaling, which are the other components of the transformation. In order to protect their images from being stolen using geometrical approaches [25], a method of WM for persistent images. It is not necessary to use DCT to embed a host image; in fact, you can choose from a variety of different methods. There are two other methods for modifying watermark information that may be used in addition to the JPEG compression strategy and the continuous gain attack (CGA). CGA attack settings were utilised to adjust the levels of brightness and darkness in digital images that had been watermarked. By utilising a model that considers the user's point of view [26] were able to make the STDM WM approach more resistant to JPEG compression. Other researchers [27] have demonstrated how they improved a JPEG compression approach [28] that is utilised in image WM systems. The findings of this research were presented in a report that was published in Computer Graphics Forum. In the previous few years, the utilisation of CNNs has become an increasingly crucial component in the process of the generation of an image prior to [29] designed the method known as deep network CNN (DnCNN) with the intention of removing noise from images. In this framework, you should make use of the convolutional kernels that are utilised in the process of identifying attacks on image watermarks. To construct a network with this level of complexity, it is required to use a broad array of architecture and many convolutional layers. Since they are all components of the hidden layer, Convolution, BatchNorm, and Relu are all utilised in the last phase of the downsampling process. The network was taught to perform blinded denoising in conjunction with cumulative Gaussian denoising. With the use of these methods, it is conceivable that the identification of attacks that make use of images with watermarks might be significantly enhanced [30] FFDNet. This apparatus can function normally in a broad variety of loud environments, including those with exceptionally high noise levels. FFDNet utilises a CNN model that is superior in both speed and accuracy compared to that of DnCNN, which uses a CNN model that is equal to FFDNet's. When it comes to precisely detecting specific sounds, this approach falls short, even though it is often believed to be accurate. These images can be utilised in CNN layers. To be able to see images that have been watermarked,

one must first come to terms with the fact that watermarks are both imperceptible and permanent over the course of time [31]. Researchers have investigated the concept of a trade-off or compromise to develop a solution that satisfies both the need for invisibility and the requirement for resistance. Among the numerous important results and discoveries that came out of this research are the following: by employing a hybrid model that can recognise a broad variety of attacks on watermarked images, the WM approach was given an increased level of protection. Through the utilisation of a scaling factor, we were successful in locating a medium between the invisibility and robustness linkages. If this capability is developed, controlling images that have been watermarked may become much simpler. This paper discusses the challenges and potential solutions for future researchers. Part 2 looks at the literature review part. Section 3 shows the proposed hybrid method for the digital WM of images. Section 4 goes into detail about the results, the conclusion is shared in Section 5.

2. Literature Review

The study proposes a copyright-protecting digital image WM approach based on DWTs and DCTs. Watermark images are encrypted using the Arnold transform; the image is more aesthetically pleasing after the system computes DWT LL sub-band block-based DCT. In this situation, a watermark created from the DCT coefficient midfrequency is used. After the host image has been processed with a 2L DWT subband, the watermark is added to the unique values recorded in the HL/LH format. The DWT–SVD digital WM approach can be built using DWT and SVD wavelet fusion techniques. It enables them to do more with less resources while limiting their exposure. The watermark can be understood even if the original image is lost. WM techniques, such as DWT–SVD hybrid WM [32–34], can be utilised to successfully integrate data into an image. Images with watermarks have noticeable changes in their diagonal singular value coefficients. When DWT–DCT–SVD was made, both the privacy of medical imaging and the protection of intellectual property rights were considered. To improve the approach's performance, the proposed WM method employs discrete wavelet transforms, DCTs, and SVDs. The DWT improves the accuracy of potentially uncovered crucial data by basing its analysis on aspects of the human visual system. This is because the DWT based its analysis on these characteristics. As a result, it is easier to select solutions that are beneficial to you. The DCT will remain unknown for an extended period. Finally, the major advantage of employing this technique is that it assures the SVD's single values are accurate. When a small amount of data, known as perturbations, is introduced to an image, it indicates that the values of the individual pixels do not fluctuate much. Several independent analyses have proved that the approach we describe is effective. It is not only easy to add a substantial amount of information without sacrificing image quality (using a 512×512 watermark), but it is also resistant to a wide range of various forms of attacks. As a result, employing this method to safeguard sensitive information is a wise decision.
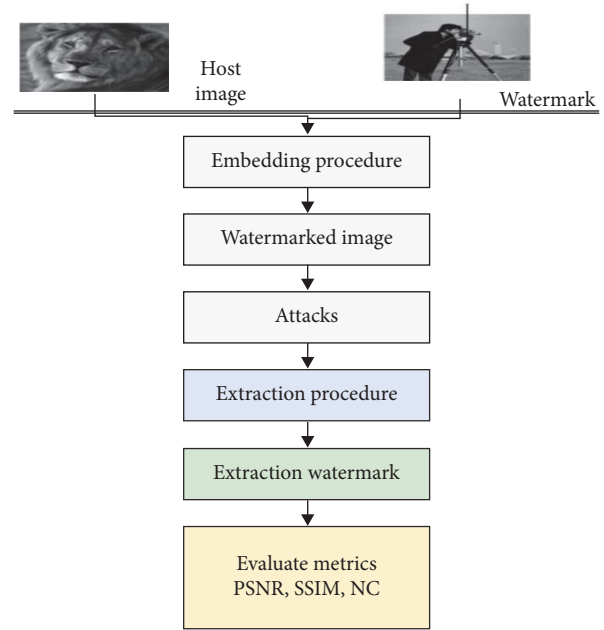
Figure 1: Watermarking procedure.

WM can be used to increase the signal-to-noise ratio (PSNR) of an image; steps for WM are shown in Figure 1. There are no scaling, filtering, cropping, or rotating issues with the approach provided. Another study suggests a method for preserving telemedicine medical data. To safely transmit medical data over a network, this study employs both DWT and DCT. In this medical host scenario, the region of interest (ROI) and non-region of interest (NROI) are distinct concepts (the ROI). In ROI and NROI, the image has both image and text watermarks. To add security before embedding a text watermark, the Rivest–Shamir–Adleman method is used. When adding a watermark, this hybrid method has no effect on the image's visual quality. A hybrid DCT–DWT technique with auto thresholding is another option. Before DWT, the host image is DCT-transformed. DWT and DCT can be combined to create a robust and unnoticed method [35]. The watermark is embedded in this study using DWT–DCT coefficients. This method is illustrative. According to studies, this method is resistant to histogram equalisation, compression, cropping, and noise addition. DWT combines DWT, DCT, and SVD to improve the safety of medical images. This medical image contains a hidden watermark. The medical host image has been 3L DWT. Finally, the host image is DCT and SVD transformed. The system reduces the visibility and resilience of Gaussian, salt-and-pepper, and Wiener noise addition, as well as Wiener, average, and median filtering. Edge detection replaces older methods for determining the best location to include a watermark [36]. To balance imperceptibility and resilience, particle swarm optimisation is used. If you require more storage, consider [37], which describes a frequency domain DCT–SVD hybrid technique. Arnold is used in this method to texture the watermark logo. Both watermark logo symbol vector decomposition and host image
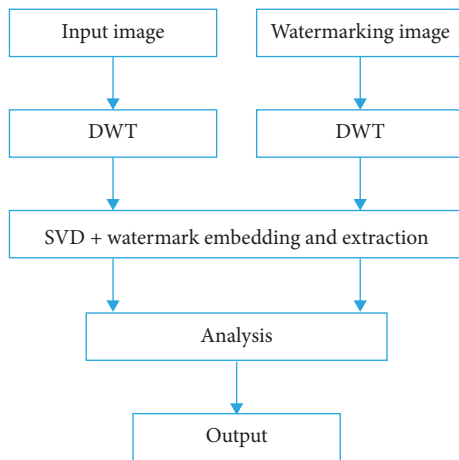
FIGURE 2: The watermark embedding procedure.

DCT processing are completed at this stage. Weights are chosen based on the host image's lowered singular values for the best results and the least distortion. To protect copyrighted material, this method [38] employs DFT and DCT. The demonstration demonstrates that the watermarked and host images are identical. The system can withstand more severe attacks. The new method improved imperceptibility and resilience. Another study employs DCT and DWT to secure watermark colour images. A three-component RGB host image must be split into red, blue, and green before DCT and DWT can be applied. Arnold secures the watermarked image. The encrypted watermark's DCT coefficient is calculated after it has been divided into equal-sized subsets. The system is more imperceptible than other methods. WM technologies that are blind and nonblind are recommended, with the former used on the inside and the latter on the outside. In 2020, researchers published a watermark reinforcement learning approach. Q learning and matrix factorisation were used. To determine which host blocks to integrate, Q learning employs trial-and-error learning. It is more resilient and undetectable than random embedding. It lacks the ability to withstand repeated attacks. Another paper [39] employs a genetic algorithm and SVM to select relevant regions and generate low-frequency regions; fuzzy entropy is used to improve the performance. A genetic algorithm is used to optimise the watermark strength in this case [40]. To find the best embedding function or block, use a genetic algorithm, ant colony optimisation, or the firefly algorithm. Machine learning and neural networks are more popular than ever [41, 42], and we present a less time-consuming DWT-based Spiking Neural Network (SNN). Extraction is viewed as an SNN-solvable optimisation problem by this network. Contourlet transforms, Kurtosis coefficients, and YCbCr spaces have all been implemented using neural networks.

Figure 2 depicts embedded watermarking process, to apply a watermark to an image, you must first do the following steps: a personalised watermark for your work might be created using this approach. After that, the watermark will be applied.

3. Hybrid Method for the Digital Watermarking of Images

One way for securing multimedia material is to use digital watermarks, which can obscure the data's original source [28]. Everything works out in the end, using an embedding strategy and a secret key, a watermark has been imprinted on the host's image. Following that, folks will swap watermarked images to remove the watermark image from the system; a certain technique and key must be used, Figure 3 depicts the hybrid method.

Procedures in space are vulnerable to attacks and manipulation [29]; as a result, transform domain solutions for multimedia security have gained traction [30]. Because of the payload limitations of transform domain methods [31], hybrid domain approaches have emerged. Two or more photo changes are required for hybrid domain WM, and transformation domain methods benefit from hybrid domain strategies [32]. These approaches are most used to protect copyrights and multimedia content, but there are other applications available. An image will first go through the DCT, DFT, DWT, and SVD editing stages before commencing the hybrid digital image WM process. It is most likely the consequence of some form of hybridisation. The n-blocks are then split down into a lower number of bits once the watermark image has been encrypted [43, 44]. Following this, a watermark image is encrypted using a block-based watermark that can be found dispersed across the image in entirely random areas. Each chunk has N bytes at this point, and the previously specified watermark will be applied to the image. Reverse the process to remove the watermark, and the hybrid watermark embedding scheme is depicted in Figure 4.

When working with watermarked images, having a mechanism for extracting features is critical. It is critical to extract characteristics from the watermarked image to determine which elements are the most important. When discussing watermarked images, it is customary to refer to both the image-maker and the result. Figure 5 depicts the steps that must be taken to remove watermarks from previously watermarked images.

WM images makes it more difficult to sell or share them since they cannot be modified or removed after capture. This method can help to avoid data theft and unauthorised use, both transforms and steganographic approaches must be used for hybrid WM to operate. Standards for digital image WM in hybrid systems because of the confluence of multimedia and the Internet, digital images may now be printed, transferred, and published over a private network or the Internet. WM digital images are accomplished by adding extra data to the host medium. This allows you to limit who can see a piece of information, and it is critical that the system works properly. The hybrid digital image WM technique can only be carried out effectively if certain parameters are met. It is critical that the system stays undiscovered while being robust, scalable, and secure. We have included a graphic illustration of their components in Figure 6 for your consideration.
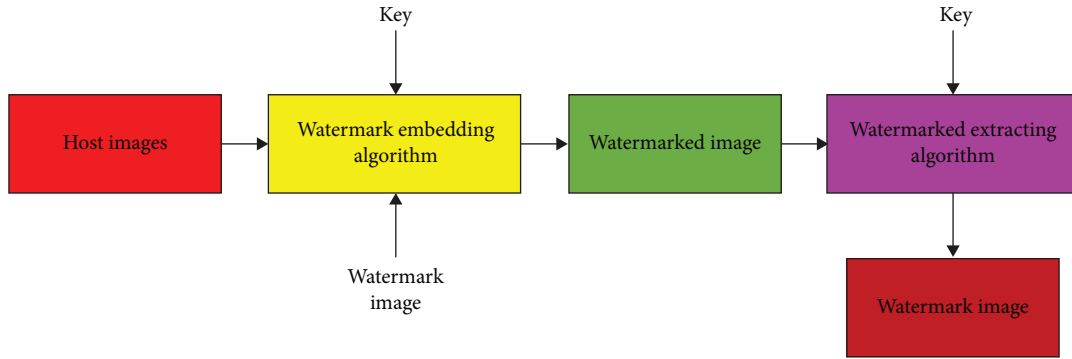
FIGURE 3: Watermark embedding and extraction is a two-step technique.
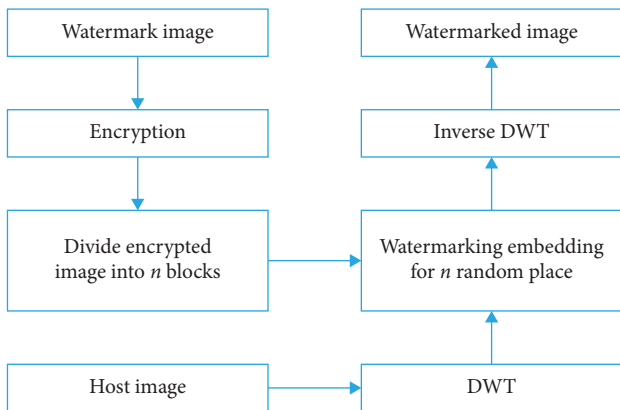
FIGURE 4: Digital image watermark embedding framework for hybrid techniques.
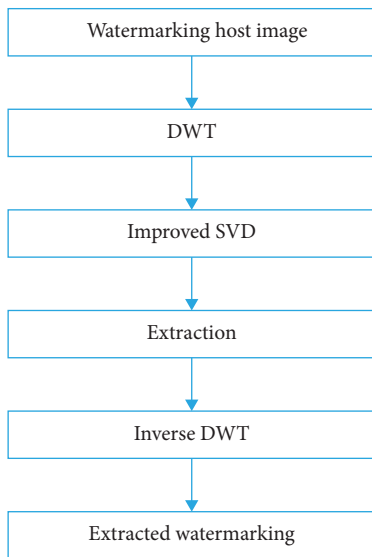
FIGURE 5: Feature extracting procedure.

FIGURE 6: Specifications necessity for hybrid watermarking approach.

The performance of hybrid digital image WM solutions is evaluated using imperceptibility. The watermarked and host images look identical. Human eyes cannot distinguish between dimmed brightness and contrast [28–33]. Because of the durability of the watermark, the watermarked image may still be traceable to its owner after standard image processing. Using this method aids in the retention of the watermark. In addition to "robustness," the word "fragility" and "semi-fragility" can be used to denote it. The amount of data that may be stored inside an image is known as its payload capacity. A host image can only store so much data. It may be challenging to incorporate more watermark bits into the image being stored. Each one of these conditions must be met before the fingerprinting procedure can begin. As a preventive step, the image containing the watermark has been encrypted. Images containing watermarks can be hidden in a variety of ways.

When it comes to digital photo WM, you may choose between a spatial or transformational watermark technique that can be used simultaneously. These approaches are classified as spatial domain techniques and include patchwork, LSB, and ISB. DCT, DFT, DWT, and SVD are a few transformation-domain techniques that may be employed in this sector. The hybrid approach to domains integrates a variety of approaches within its overarching framework. Algorithms of this type include DCT/DFT/DWT/SVD. As seen in Figure 7, there are several techniques for WM images.

The efficacy of assaults like image manipulation and sharpening are greatly decreased when hybrid digital photo WM technologies are utilised. These various techniques are getting increasingly computationally intensive (time and space). It is physically impossible for them to have all these attributes at the same time. Concerns about users' capacity to retain their privacy have directly resulted from the introduction of this technology. Computers, on the other hand, rely on more complex
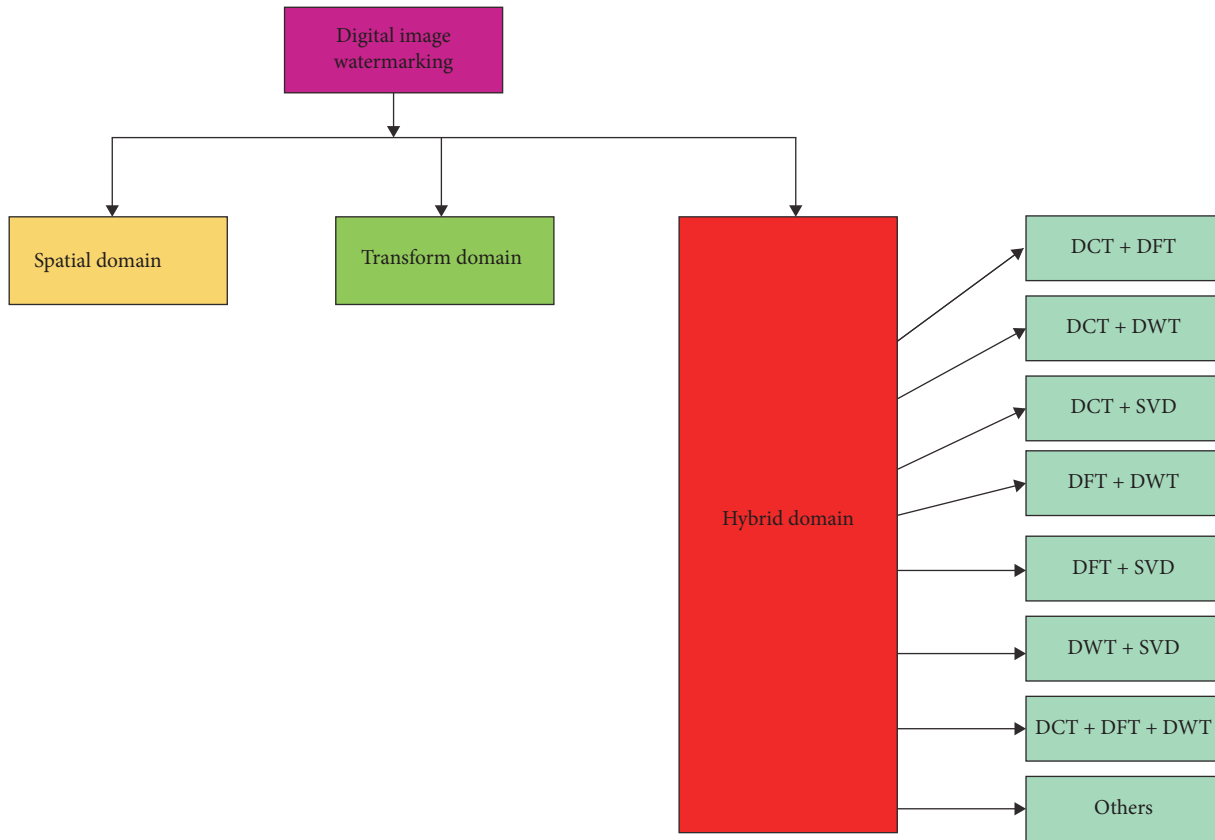
FIGURE 7: Watermarking hybrid domain possibilities.

algorithms for security, whereas internet-connected devices use simpler approaches. The Internet of Things (IoT) necessitates less storage space and processing power [36]. Blockchain technology allows for decentralised authentication. A blockchain is a decentralised digital ledger made up of encrypted, linked data chunks that are protected by encryption. The digital signature of the host photo is saved using blockchain technology. An attacker cannot alter the host's appearance. The updated area and host image can be validated using the blockchain signature. Near-field communication can be used to authenticate pharmaceuticals on the IoT [38]. ROR is used to secure session keys (real-or-random). This model saves money on computers and transmission. Content-based image retrieval does not require information about cloud servers. Photos in this demo are feature vectors, not bitmaps. The feature vector is protected by k-neighbors. LBRAPS, a lightweight authentication technique for RFID data based on blockchain, will benefit 5G mobile supply chains. The method defends against multiple attackers [40]. Outsiders can examine cloud server data from here. Cloud server and smart metre authentication allow for the secure exchange of cloud server data. Bitwise XOR and cryptographic hash algorithms have been shown in experiments to make the procedure more secure.

4. Results

It is required to use a range of sizes when computing the scaling factor, and this information must subsequently be provided to optimal edge detection. Using watermarked images within the suggested framework allows for the management of image invisibility and the strengthening of images against various forms of assaults. Lions, just like filmmakers, have certain requirements when it comes to the size of their images and the values of their important qualities; we are aware of both requirements due to our familiarity with each of these professions. By keeping an eye on this connection, one can gain control over image invisibility and durability. The normalized correlation measure is used to show the relative importance of a variety of unique elements in connection to assaults (Figure 8).

For example, it was revealed that motion blur and average filters had exceptionally low PSNR values. The PSNR values that have been measured are shown on an axis. All the attack scale variables are presented along the labelled axis, which reads. Most of the watermark images that we looked at displayed symptoms associated with a wide range of diseases. These techniques were used to address difficulties such as motion blur, spherical noise compression, and histogram equivalency. These criteria were used to assess the watermark image's resilience to a variety of attacks. Using this approach, you may assess how resistant the watermark image is to various assaults. An assault was carried out on the watermark images to establish how well they would survive the normalized correlation measure.

Figure 9 summarises the experiment's findings because of technological developments; the screening procedure for

	No attack		JPEG compression
	Gaussian low-pass filter		JPEG2000 compression
	Median		Sharpening attack
	Gaussian noise		Histogram equalization
	Salt and pepper noise		Average filter
	Speckle noise		Motion blur
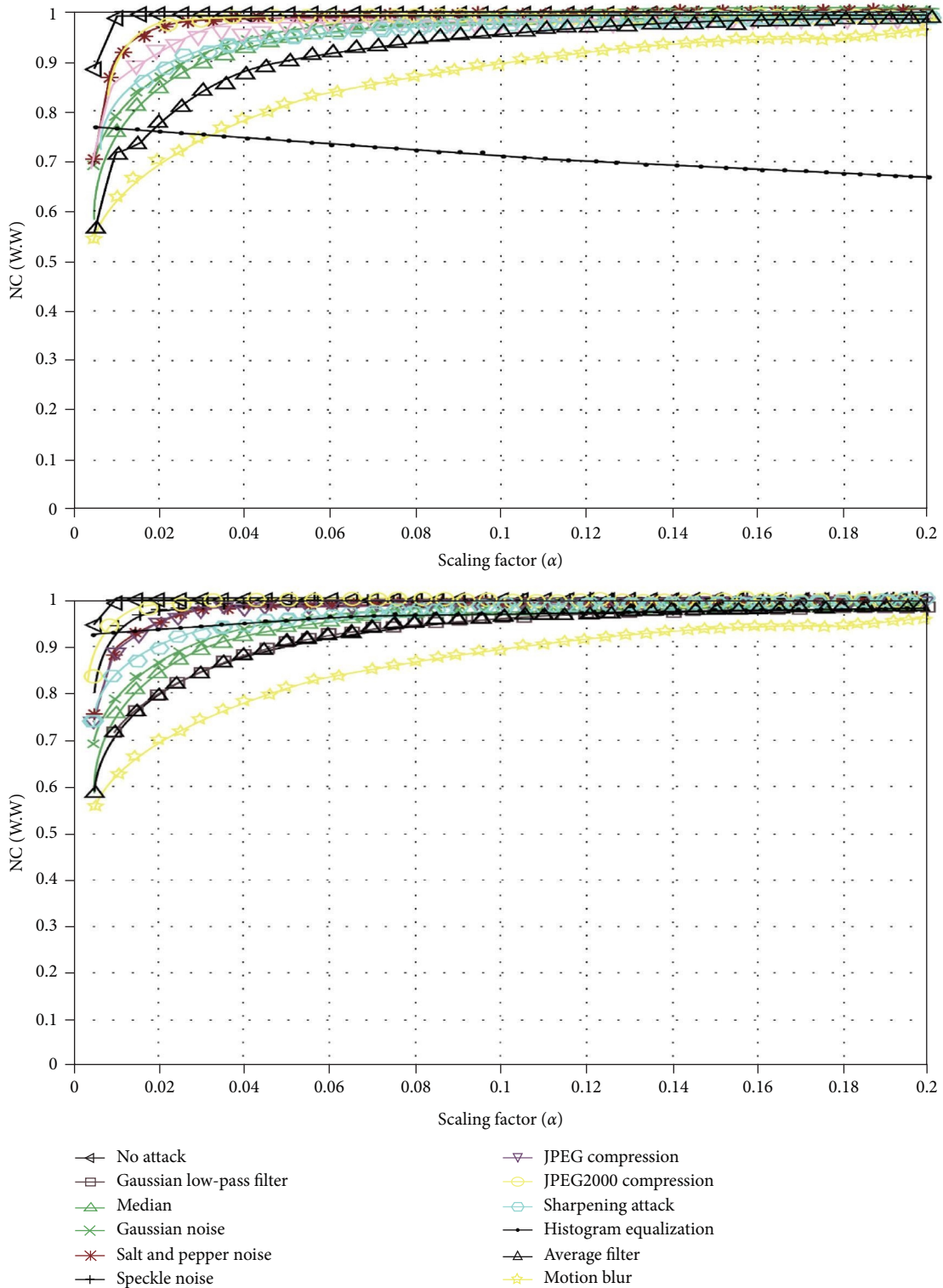
FIGURE 8: The NC approach was utilised to scale both the lion image and the image of Lina.

averting various forms of assaults proceeded quite successfully. Following the assaults, watermarked images were gathered and analysed using DWT–SVD. Table 1 depicts the study's findings; the values offered by NC are superior to the bulk of attacks.

The researchers used a total of six distinct procedures, some of which were rather difficult and included the use of techniques such as Gaussian filtering and median filtering. As a direct result of the investigation, the quality of previously wrecked images has greatly improved. Multiple attacks revealed that the proposed strategy was significantly more successful than the present one. According to the study's findings, this strategy had a considerable influence on the
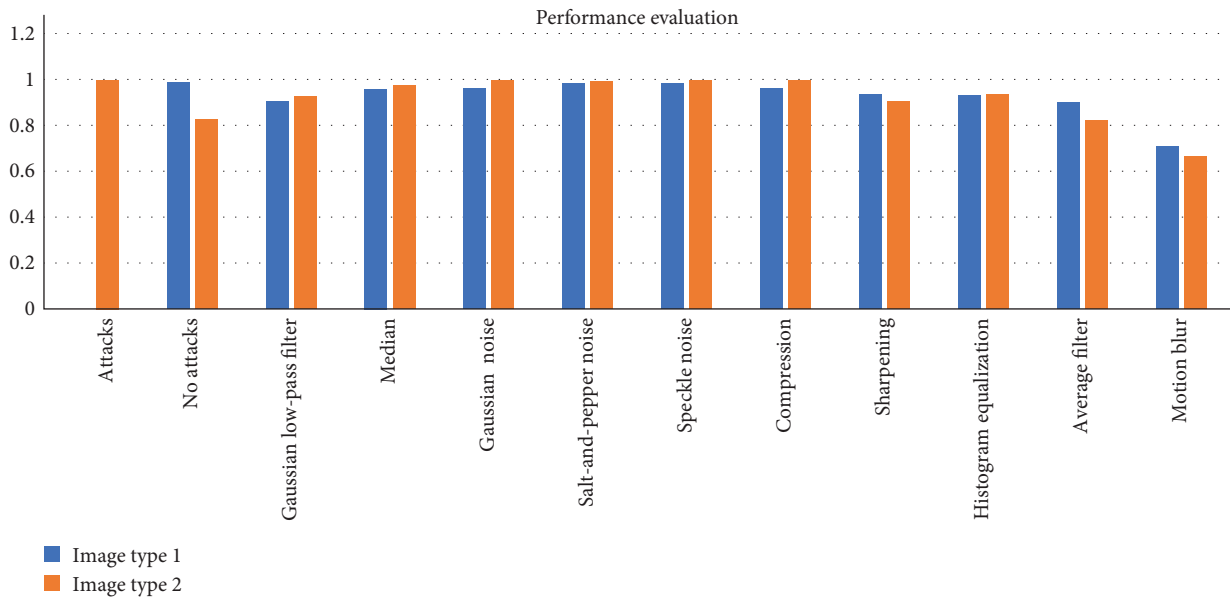
FIGURE 9: Performance evaluation of the proposed method.

TABLE 1: Performance evaluation of the proposed method with traditional methods.

Attacks or noises	Proposed system		Existing system	
	Normalized correlation	Structural similarity index	Normalized correlation	Structural similarity index
Motion blur	96.90	98.2	94.4	86.7
Gaussian noise	98.35	98.4	88.9	62.3
Speckle noise	91.71	97.7	83.1	42.9

desired objectives. When analysing watermarked photos, Structural Similarity Index (SSIM) and PSNR were used to determine their invisibility, whereas normalized correlation was utilised to determine their resistance to the assaults that were employed. The study's empirical data revealed that normalized correlation performed admirably in analysing each of the assaults used. When the suggested system was compared to those currently in existence, it was discovered to have greater performance. This approach may help digital image WM since it allows unfettered mobility of the WM information as well as WM on the host image without causing any damage to either of them.

5. Conclusion

Images are included in multimedia data sets. Image authentication becomes difficult due to Internet traffic. IoT technologies and interactive multimedia data transfer make it easier to copy and distribute information. Making the system invisible, resilient, and data-embedding capable is just as important as ensuring the privacy of image data. Keep these difficulties in mind when WM digital images. To address these challenges, we concluded that existing hybrid techniques must be improved. Because of the dynamic nature of digital multimedia data transmission, it is now possible to easily regenerate information. The DWT–SVD approach will

be used in this study to investigate a range of watermark assaults. We used scaling variables to study the relationship between being invisible and having a resilient character. Watermarks were expected to be implanted and removed in a manner that was both undetectable and resistant to numerous attacks for the purposes of the study. The following are some of the study's most notable findings: the DWT–SVD technique was utilised to extract a wide range of relevant data to identify all watermark attacks. To put the proposed approach to the test, hybrid DWT–SVD was applied to watermarked images. The SSIM, PSNR, and normalized correlation data were used to evaluate the proposed system's efficacy. When analysing watermarked photos, SSIM and PSNR were used to determine their invisibility, whereas normalized correlation was utilised to determine their resistance to the assaults that were employed. The study's empirical data revealed that normalized correlation performed admirably in analysing each of the assaults used. When the suggested system was compared to those currently in existence, it was discovered to have greater performance. This approach may help digital image WM since it allows unfettered mobility of the WM information as well as WM on the host image without causing any damage to either of them. Using the suggested approach, invisibility and resilience may be increased and regulated, allowing the system to respond to a wide range of threats at its best. Our aim is to

perfect this technology to the point where there is scarcely any quality loss in the watermarked image, and the restored watermark is almost immaculate. Watermark extraction accuracy should improve dramatically with the addition of this additional information. If the massive over dictionary was the primary focus of the investigation, other feature learning methodologies or training models might be used in possible future research paths.

Data Availability

ImageNet data is publicly available at the following URL: https://www.image-net.org/.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] P. Morasso, "Spatial control arm movements," *Experimental Brain Research*, vol. 42, pp. 223–227, 1981.

[2] Y. Uno, M. Kawato, and R. Suzuki, "Formation and control of optimal trajectory in human multijoint arm movement," *Biological Cybernetics*, vol. 61, pp. 89–101, 1989.

[3] L. P. Feng, L. B. Zheng, and P. Cao, "ADWT-DCT based blind watermarking algorithm for copyright protection," in *Proceedings of 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT)*, pp. 455–458, Chengdu, China, July 2010.

[4] X. Zhou, J. Ma, and W. Du, "SoW: a hybrid DWT-SVD based secured image watermarking," in *Proceedings of the International Conference on Sensor Network Security Technology and Privacy Communication System (SNS & PCS)*, pp. 197–200, Nangang, China, May 2013.

[5] E. E.-D. Hemdan, N. El-Fishaw, G. Attiya, and F. Abd El-Samii, "C11. Hybrid digital image watermarking technique for data hiding," in *Proceedings of the 30th National Radio Science Conference*, pp. 220–227, Cairo, Egypt, April 2013.

[6] R. Nair and A. Bhagat, "An application of big data analytics in road transportation," in *Advances in Systems Analysis, Software Engineering, and High Performance Computing*, pp. 39–54, IGI Global, Hershey, PA, USA, 2018.

[7] R. Nair and A. Bhagat, "An introduction to clustering algorithms in big data," in *Encyclopedia of Information Science and Technology*, pp. 559–576, IGI Global, Hershey, PA, USA, 5th edition, 2021.

[8] R. Nair, S. N. Zafrullah, P. Vinayasree et al., "Blockchain-based decentralized cloud solutions for data transfer," in *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–12, Hindawi Limited, 2022.

[9] R. Kashyap, "Applications of wireless sensor networks in healthcare," in *Advances in Wireless Technologies and Telecommunication*, pp. 8–40, IGI Global, Hershey, PA, USA, 2020.

[10] R. Nair, S. Gupta, M. Soni, P. K. Shukla, and G. Dhiman, "An approach to minimize the energy consumption during blockchain transaction," in *Materials Today: Proceedings*, Elsevier, 2020.

[11] N. Waoo, R. Kashyap, and A. Jaiswal, "DNA nano array analysis using hierarchical quality threshold clustering," in *2010 2nd IEEE International Conference on Information Management and Engineering*, pp. 81–85, IEEE, Chengdu, China, 2010.

[12] R. Kashyap and A. Piersson, "Big data challenges and solutions in the medical industries," in *Advances in Systems Analysis, Software Engineering, and High Performance Computing*, pp. 1–24, IGI Global, Hershey, PA, USA, 2018.

[13] R. Kashyap, "Big data analytics challenges and solutions," in *Big Data Analytics for Intelligent Healthcare Management*, pp. 19–41, IGI Global, Hershey, PA, USA, 2019.

[14] I. Assini, A. Badri, K. Safi, A. Sahel, and A. Baghdad, "A robust hybrid watermarking technique for securing medical image," *International Journal of Intelligent Engineering and Systems*, vol. 11, no. 3, pp. 169–176, 2018.

[15] T. T. Takore, P. R. Kumar, and G. L. Devi, "A new robust and imperceptible image watermarking scheme based on hybrid transform and PSO," *International Journal of Intelligent Systems and Applications*, vol. 10, no. 11, pp. 50–63, 2018.

[16] P. Jain and U. Ghanekar, "Robust watermarking technique for textured images," in *Proceedings of the 6th International Conference on Smart Computing and Communications (ICSCC)*, pp. 179–186, Kurukshetra, India, December 2018.

[17] M. Hamidi, M. E. Haziti, H. Cherifi, and M. E. Hassouni, "Hybrid blind robust image watermarking technique based on DFT-DCT and Arnold transform," *Multimedia Tools and Applications*, vol. 77, no. 20, pp. 27181–27214, 2018.

[18] Y. Zhang, Y. Li, and Y. Sun, "Digital watermarking based on Joint DWT-DCT and OMP reconstruction," *Circuits, Systems and Signal Processing*, vol. 38, no. 11, pp. 5135–5148, 2019.

[19] A. K. Abdulrahman and S. Ozturk, "A novel hybrid DCT and DWT based robust watermarking algorithm for color images," *Multimedia Tools and Applications*, vol. 78, pp. 17027–17049, 2019.

[20] D. G. Savakar and A. Ghuli, "Robust invisible digital image watermarking using hybrid scheme," *Arabian Journal for Science and Engineering*, vol. 44, no. 4, pp. 3995–4008, 2019.

[21] M. Alizadeh, H. Sajedi, and B. Babaali, "Image watermarking by Q learning and matrix factorization," in *Proceedings of the International Conference on Machine Vision and Image Processing (MVIP)*, Qom, Iran, February 2020.

[22] R. Mehta, K. Gupta, and A. K. Yadav, "An adaptive framework to image watermarking based on the twin support vector regression and genetic algorithm in lifting wavelet transform domain," *Multimedia Tools and Applications*, vol. 79, no. 25-26, pp. 18657–18678, 2020.

[23] G. Dubey, C. Agarwal, S. Kumar, and H. P. Singh, "Image watermarking scheme using cuckoo search algorithm," in *Advances in Data and Information Sciences*, vol. 94 of *Lecture Notes in Networks and Systems*, Springer, Singapore, 2020.

[24] R. A. Dobre, R. O. Preda, C. C. Oprea, and L. Pirnog, "Authentication of JPEG images on the blockchain," in *Proceedings International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO)*, pp. 211–215, Prague, Czech Republic, May 2018.

[25] Y. Guo, B.-Z. Li, and N. Goel, "Optimised blind image watermarking method based on firefly algorithm in DWT-QR transform domain," *IET Image Processing*, vol. 11, no. 6, pp. 406–415, 2017.

[26] Y.-H. Chen, H.-C. Huang, C.-H. Lai, and T.-Y. Chang, "An image watermarking approach based on artificial fish swarm algorithm," in *Proceedings of Analysis of Watermarking Framework for Color Image Through*, pp. 46–50, Kyoto, Japan, March 2020.

[27] M. F. Kazemi, M. A. Pourmina, and A. H. Mazinan, "Analysis of watermarking framework for color image through a neural

network-based approach," *Complex & Intelligent Systems*, vol. 6, no. 1, pp. 213–220, 2020.

[28] M. F. Kazemi, M. A. Pourmina, and A. H. Mazinan, "Novel neural network based CT-NSCT watermarking framework based upon Kurtosis coefficients," *Sensing and Imaging*, vol. 21, no. 1, Article ID 7, 2020.

[29] A. K. Yadav and R. Mehta, "Local coupled extreme learning machine based image watermarking using DCT in YCbCr space," in *Proceedings of the Amity International Conference on Artificial Intelligence (AICAI)*, pp. 527–532, Dubai, UAE, February 2019.

[30] M. Naseri, M. Abdolmaleky, A. Laref et al., "A new cryptography algorithm for quantum images," *Optik*, vol. 171, pp. 947–959, 2018.

[31] S. Heidari, M. M. Abutalib, M. Alkhambashi, A. Farouk, and M. Naseri, "A new general model for quantum image histogram (QIH)," *Quantum Information Processing*, vol. 18, no. 6, pp. 1–20, 2019.

[32] R. Krishnamoorthi, S. Joshi, H. Z. Almarzouki et al., "A novel diabetes healthcare disease prediction framework using machine learning techniques," *Journal of Healthcare Engineering*, vol. 2022, Article ID 1684017, 10 pages, 2022.

[33] I. Ahmad, S. H. Serbaya, A. Rizwan, and M. S. Mehmood, "Spectroscopic analysis for harnessing the quality and potential of gemstones for small and medium-sized enterprises (SMEs)," *Journal of Spectroscopy*, vol. 2021, Article ID 6629640, 12 pages, 2021.

[34] H. Z. Almarzouki, H. Alsulami, A. Rizwan, M. S. Basingab, H. Bukhari, and M. Shabaz, "An internet of medical things-based model for real-time monitoring and averting stroke sensors," *Journal of Healthcare Engineering*, vol. 2021, Article ID 1233166, 9 pages, 2021.

[35] D. Delannay and B. Macq, "Generalized 2-D cyclic patterns for secret watermark generation," in *Conference Proceedings 2000 International Conference on Image Processing (Cat. no. 00ch37101)*, vol. 2, pp. 77–79, Vancouver, BC, Canada, September 2000.

[36] H. Tao, L. Chongmin, J. M. Zain, and A. N. Abdalla, "Robust image watermarking theories and techniques: a review," *Journal of Applied Research and Technology*, vol. 12, no. 1, pp. 122–138, 2014.

[37] N. A. Loani, N. N. Hurrahi, S. A. Parah, J. W. Lee, J. A. Sheikhi, and G. M. Bhat, "Secure and robust digital image watermarking using coefficient differencing and chaotic encryption," *IEEE Access*, vol. 6, pp. 19876–19897, 2018.

[38] T. Yang, G. H. Zhang, L. Liu et al., "New features of authentication scheme for the IoT: a survey," in *Proceedings of the 2nd International ACM Workshop on Security and Privacy for the Internet-of-Things*, pp. 44–49, London, UK, November 2019.

[39] M. Wazid, A. K. Das, M. K. Khan, A. A.-D. Al-Ghaiheb, N. Kumar, and A. V. Vasilakos, "Secure authentication scheme for medicine anti-counterfeiting system in IoT environment," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1634–1646, 2017.

[40] Z. Xia, N. N. Xiong, A. V. Vasilakos, and X. Sun, "EPCBIR: an efficient and privacy-preserving content-based image retrieval scheme in cloud computing," *Information Sciences*, vol. 387, pp. 195–204, 2017.

[41] C. Lin, D. He, X. Huang, K.-K. R. Choo, and A. V. Vasilakos, "BSeIn: a blockchain-based secure mutual authentication with fine-grained access control system for industry 4.0," *Journal of Network and Computer Applications*, vol. 116, pp. 42–52, 2018.

[42] S. Jangirala, A. K. Das, and A. V. Vasilakos, "Designing secure lightweight blockchain-enabled RFID-based authentication protocol for supply chains in 5G mobile edge computing environment," *IEEE Transactions on Industrial Informatics*, vol. 16, pp. 7081–7093, 2019.

[43] S. Challa, A. K. Das, P. Gope, N. Kumar, F. Wu, and A. V. Vasilakos, "Design and analysis of authenticated key agreement scheme in cloud-assisted cyber-physical systems," *Future Generation Computer Systems*, vol. 108, pp. 1267–1286, 2020.

[44] M. Wazid, A. K. Das, K. V. Bhat, and A. V. Vasilakos, "LAM-CIoT: lightweight authentication mechanism in cloud-based IoT environment," *Journal of Network and Computer Applications*, vol. 150, Article ID 102496, 2020.

Retraction

Retracted: Mathematical Methods for Sensitive Information Mining Method of News Communication Platform Based on Big Data IOT Analysis

Mathematical Problems in Engineering

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] C. Jiang and Z. Yang, "Mathematical Methods for Sensitive Information Mining Method of News Communication Platform Based on Big Data IOT Analysis," *Mathematical Problems in Engineering*, vol. 2022, Article ID 1606183, 7 pages, 2022.

Retraction

Retracted: Retinex Algorithm and Mathematical Methods Based Texture Detail Enhancement Method for Panoramic Images

Mathematical Problems in Engineering

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Kang, "Retinex Algorithm and Mathematical Methods Based Texture Detail Enhancement Method for Panoramic Images," *Mathematical Problems in Engineering*, vol. 2022, Article ID 6490393, 8 pages, 2022.

Research Article

Transfer and Deep Learning-Based Gurmukhi Handwritten Word Classification Model

Tajinder Pal Singh ⓘ,[1] **Sheifali Gupta** ⓘ,[1] **Meenu Garg** ⓘ,[1] **Amit Verma** ⓘ,[2] **V. V. Hung,**[3] **H. H. Thien** ⓘ,[3] **and Md Khairul Islam** ⓘ[4]

[1]*Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India*
[2]*University of Petroleum & Energy Studies, Dehradun, India*
[3]*Faculty of Information Technology, Ho Chi Minh City Open University, Ho Chi Minh City, Vietnam*
[4]*Department of Information & Communication Technology, Islamic University, Kushtia 7003, Bangladesh*

Correspondence should be addressed to Amit Verma; amit.uptu2006@gmail.com and Md Khairul Islam; mdkito51@gmail.com

The world is having a vast collection of text with abandon of knowledge. However, it is a difficult and time-taking process to manually read and recognize the text written in numerous regional scripts. The task becomes more critical with Gurmukhi script due to complex structure of characters motivated from the challenges in designing an error-free and accurate classification model of Gurmukhi characters. In this paper, the author has customized the convolutional neural network model to classify handwritten Gurmukhi words. Furthermore, dataset has been prepared with 24000 handwritten Gurmukhi word images with 12 classes representing the month's names. The dataset has been collected from 500 users of heterogeneous profession and age group. The dataset has been simulated using the proposed CNN model as well as various pretrained models named as ResNet 50, VGG19, and VGG16 at 100 epochs and 40 batch sizes. The proposed CNN model has obtained the best accuracy value of 0.9973, whereas the ResNet50 model has obtained the accuracy of 0.4015, VGG19 has obtained the accuracy of 0.7758, and the VGG16 model has obtained value accuracy of 0.8056. With the current accuracy rate, noncomplex architectural pattern, and prowess gained through learning using different writing styles, the proposed CNN model will be of great benefit to the researchers working in this area to use it in other ImageNet-based classification problems.

1. Introduction

In the application of natural language processing, handwritten text recognition is used on records for data analysis and recognition. Through data analysis and recognition, it provides an interface for the improvement of communication between humans and computers. Nowadays, this process become more popular for regional applications as most of the information has got digitized everywhere. However, as compared to non-Indian scripts, handwritten text recognition for Indian regional scripts has not achieved expected perfection due to the complex structure of scripts such as Gurmukhi. Hence, it remains an active research area till now.

Generally, handwritten text recognition systems are developed based on an analytical word recognition approach or a holistic word recognition approach. Handwritten text recognition based on analytical word recognition performs character level segmentation of a word before text recognition. On the other hand, text recognition based on holistic word recognition is a segmentation-free approach in which the whole word is recognized at once using its contour or shape information. Hence, holistic word recognition technique is very common for error-free results on documents comprising of overlapped or touched characters as there is no need of word segmentation. In the present work, the authors have also employed a holistic approach to handwritten Gurmukhi word recognition.

Furthermore, handwritten text recognition systems that have been developed are based on machine learning algorithms or deep learning algorithms. Mostly, machine learning algorithms based on handwritten text recognition systems approach to manual features' extraction for text recognition, in which low-level or high-level character-level features have been extracted for interpreting the text accurately. For example, both [1, 2] have employed machine learning to recognize the online handwritten Gurmukhi characters using manual extracted low-level and high-level features of Gurmukhi characters and have achieved a 90.08% recognition efficiency. In order to extract the moment-based invariant features of 52 bilingual (Roman and Gurmukhi) characters for recognition, Dhir [3] applied various moments such as Zernike moments, pseudo-Zernike moments, and orthogonal Fourier–Mellin moments. Out of these, pseudo-Zernike moments outperformed the other moments. Kumar et al. [4] employed machine learning to compare the handwriting of various writers in Gurmukhi. For this, the authors have prepared a feature set of manually extracted zoning, directional, and diagonal features. The classification results of this article show that writer 6 has achieved the best accuracy. Kumar et al. [5] employed machine learning-based classification techniques for text recognition on manual extracted features of Gurmukhi characters that are parabola curve fitting-based features and power curve fitting-based features. The author has achieved an accuracy rate of up to 98.10% using this approach in the recognition of offline handwritten Gurmukhi characters. Both [6, 7] have developed online handwritten Gurmukhi character recognition systems based on various machine learning-based classification techniques and 64-points feature set extracted from online handwritten Gurmukhi characters. Similarly, Kumar et al. [8] have also developed an offline handwritten Gurmukhi character recognition system using machine learning classifiers and various transformation techniques such as discrete wavelet transformations (DWT2), discrete cosine transformations (DCT2), fast Fourier transformations, and fan beam transformations that were used for Gurmukhi character's feature extraction.

Singh et al. [9] have used machine learning for online handwritten Gurmukhi script recognition based on points feature, discrete Fourier transformation features, and directional features. The proposed approach has achieved recognition accuracy of 97.07%. Both [10, 11], for developing writer identification systems in Gurmukhi text, have used machine learning-based classification techniques and feature sets comprising of zoning, transitions, peak extent, centroid, parabola curve fitting, and power curve fitting-based features. An accuracy of 89.85% and an average of 81.75% have been achieved using this approach. Kumar et al. [12] have proposed an optical character recognition system that has been developed using machine learning classifiers and manual extracted features such as zoning, discrete cosine transformations, and gradient feature extraction. The proposed optical character recognition system has been trained on the 18th to 20th century's mediaeval handwritten documents and achieved a 95.91% accuracy rate in the recognition of works written by distinct writers. On degraded handwritten Gurmukhi characters, Garg et al. [13] have achieved recognition accuracy of 96.03% using machine learning classifiers and zoning, diagonal, shadow, and peak extent-based features. In order to evaluate the performance of character recognition systems based on machine learning algorithms, Kumar et al. [14] have experimented on a dataset of handwritten Gurmukhi characters and numerals. The system's accuracy of up to 87.9% has been achieved when peak extent diagonal and centroid kinds of features were extracted from characters and numbers.

In contrast to handwritten text recognition systems based on machine learning algorithms, a handwritten text recognition system based on deep learning can be developed through customized convolution neural networks or pretrained transfer learning models. Both customized convolution neural networks and pretrained models work on automatically extracting features from handwritten text. Specifically, pretrained models, in the context of deep learning, are the methods that use the features that have been learned by a network on a given problem to solve a different set of problems in the same domain. Previously, various authors have used pretrained models for handwritten text recognition. For example, Ganji et al. [15] employed a transfer learning model named "VGG16" to recognize the multivariant handwritten Telugu character. The proposed approach by the authors in this article is the alternative to the already existing optical character recognition systems, which were not capable of recognizing the variant handwritten Telugu characters due to the availability of a limited dataset. A maximum accuracy of 92% has been achieved using this method. In order to recognize the handwritten words using a holistic approach, Pramanik and Bag [16] used AlexNet, VGG-16, VGG-19, ResNet50, and GoogleNet on a dataset of Bangla city names. The author has achieved a maximum accuracy of 98.86% with ResNet50 in this experiment. Similarly, Pramanik et al. [17] have performed experimentation for the recognition of Bangla handwritten pin codes on postal letters as a part of practical application using an already available pretrained CNN model on various Indic scripts, including Bangla, Devanagari, Oriya, and Telugu. When tested on 28 postcards and 168 total digits, the pertained models, AlexNet and VGG16, achieved maximum efficiency of 94.26% and 92.21%, respectively, in this experimentation. As the pretrained model has been used for developing a handwritten text recognition system based on deep learning, researchers have also employed customized convolution neural networks to develop the same. For example, [18] customized a convolution neural network with two convolutional layers and two max-pooling layers to recognize 3500 Gurmukhi characters. The results show that the customized convolution neural network in this article has achieved 98.32% accuracy on the training set and 74.66% on the texting set data. To effectively recognize the offline handwritten text, a sequence-to-sequence method based on CNN-RNN models has been developed by [19]. The suggested model demonstrated competitive word recognition accuracy when tested on the IAM and RIMES handwritten

datasets. Furthermore, in order to recognize online Gurmukhi words, [20] customized a convolution neural network with seven convolutional layers and three maximum pooling layers that gives 97% accuracy. Similarly, [21] have a customized convolution neural network with 1 convolution and three capsule layers, [22] customized the binary neural network, [23] customized the deep neural network based on ResNet 50 called CORNet, and [24] customized the convolutional neural network with 19 layers including 6 convolutional, 3 max-pooling, 4 dropout layers, 3 batch normalization layers, 1 flatten layer, and 2 dense layers.

Pretrained models somehow give numerous benefits of using it, such as saving computational time and being very useful on small datasets, but it may face some critical issues for solving different problems in the same domain where it is trained, such as negative transfer and inaccuracy in identifying decision boundaries among multiple classes in the dataset of the target domain. This will result in making it unsuitable for real-time applications such as automated month's name recognition. Hence, it is recommended to build a custom convolution neural network whose learning will be initialized from scratch for the purpose of effective performance in such applications of automated text recognition systems.

With the objective of proposing a design for an automated month's name recognition system in the Gurmukhi language, the authors have customized a convolutional neural network with 5 convolutional and 3 pooling layers from scratch. For the present classification problem, the weights of a customized convolution neural network are initialized from the beginning and updated through learning the problem, irrespective of utilizing the frozen weights of pretrained models. This leads to better identification of decision boundaries among the 24 classes on a given dataset in the target domain using a customized convolution neural network. As a result of this, the proposed CNN model has shown better performance than the pretrained model on a custom dataset in terms of training results as well as results for confusion matrix parameters.

The major contribution in terms of novelty of the present research work in the target domain has been well explained in the following points:

(1) A convolutional neural network has been customized from scratch with 5 convolutional and 3 pooling layers for an automated month's name recognition system in the Gurmukhi language for regional applications. The proposed CNN design in term of its architectural pattern is less complex and less prone to over fitting when compared to various transfer learning models used for image classification task in the given context. Hence, the proposed CNN model is suitable for portable automated text recognition systems.

(2) The performance of customized convolutional neural network for handwritten text recognition system has been tested on custom dataset of handwritten words that has been prepared from 500 distinct writers from various professions and age groups. Hence, the proposed CNN model is a prowess model through the learning of different writing styles.

(3) Furthermore, it has been observed from the result analysis that the customized convolutional neural network on handwritten Gurmukhi month's name dataset possessed an accuracy rate of up to 99.73%, which is more than using various transfer learning models named as ResNet50, VGG 19, and VGG16, respectively, on the same dataset. As a result, the proposed CNN model's learning weights can be exploited for the use in other ImageNet-based classification problems in the given context.

The present article has been divided into various sections. Section 1 focuses on the introduction and literature review of the article. Section 2 describes the materials and methods used, while conducting the experiment, and also details the complete description of the proposed CNN model's architecture. In Section 3 of this article, results and analysis of the proposed experimentation work have been presented in terms of results obtained at different training parameters and a comparative analysis performed among the various models that have been chosen in this work. Finally, Section 4 has concluded the overall results of various experiments performed for the performance assessment of the proposed CNN model in the classification of Gurmukhi month's name images and its comparison to other three pretrained models.

2. Materials and Methods

This section provides a detailed description of the step-by-step approach of dataset preparation and the methodology used for the classification of handwritten word datasets.

2.1. Dataset.
The dataset for the present research work has been prepared for 24,000 handwritten word images belonging to 24 different classes of Gurmukhi month's names that represent the handwriting style variation among 500 distinct writers from different professions and age groups.

The various steps for dataset preparation have been shown in Figure 1.

The step-by-step approach of dataset preparation has been described by the following points:

(1) In the first step, an offline handwritten Gurmukhi month's name dataset was created on an A4 sheet of paper that contains 48 blocks of the same size for writing the handwritten words. A single writer wrote the names of the 24 classes of the Gurmukhi months two times on an A4 sheet and gave 48 handwritten words on a single piece of paper. Similarly, the dataset for handwritten words of 24 classes of Gurmukhi month's name has been collected on 500 A4-sized sheets from 500 distinct writers, resulting in the formation of 24,000 words in the Gurmukhi month's name dataset. This step helps in
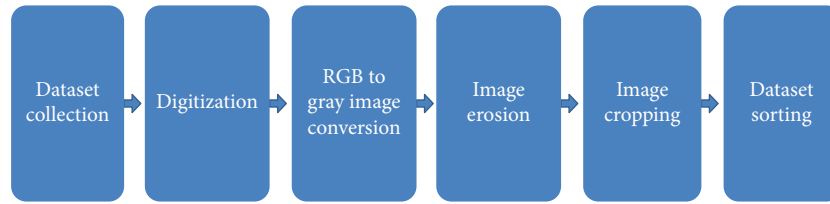
FIGURE 1: Steps to dataset preparation.

incorporating different writing styles into the dataset, which will help to prowess the proposed CNN model through learning it.

(2) All 500 A4-sized sheets collected from 500 distinct writers have been digitized in the second step using an OPPO F1s smart phone camera. Each digitized image of sample sheet is of size 1024×786 pixels. Using this step called digitization, offline handwritten word recognition has become possible and easy.

(3) In the third step of dataset preparation, colour to gray image conversion has been performed on digitized dataset using MATLAB operation "rgb2gray." A conversion of colour to gray image has reduced the data size as well as the computation cost, which results in fast learning and recognition rate.

(4) The dataset preparation in the fourth step has performed eroding on all gray scale images using "imerode" function of MATLAB software. Image eroding helps in retrieving the valuable information from digitized text data.

(5) In the fifth step of dataset preparation, the dataset of eroded images has been cropped into word images and collected 48-word images from each sample erode image. This operation has been performed using "imcrop" function of MATLAB software. This operation helps in sorting the dataset classwise.

(6) In the sixth step of dataset preparation, class namewise dataset sorting has been performed on cropped images. For each class in the 24 classes of Gurmukhi month's name, cropped images have been sorted into the class's respective folders. This results in a collection of 1000 images for each class name in the Gurmukhi month's name dataset. Using this operation, distinction has been incorporated among each image of the different classes of the Gurmukhi month's name.

The results of dataset preparation are shown in Figure 2.

The customized dataset for this experiment was divided into training and testing sets in an $80 : 20$ ratio. That means a total of 19,200 labeled samples from the Gurmukhi handwritten words dataset have been used as a training set and the remaining 4,800 labeled samples have been used as a testing set. The detailed description of dataset as well as the training set and testing set samples selected from each class of Gurmukhi month's name in the present work is shown in Table 1.

2.2. Proposed CNN Model.

This section will illustrate the design and development of a proposed CNN model that is built on alternating convolution layers, normalization layers, max-pooling layers, and dropouts, followed by some fully connected layers. In general, the CNN is a multilayer neural network that stores numerous trainable weights and biases, while training using forward and backpropagation algorithms. A CNN model is constructed through automatic feature extractors and trained classifiers, which help in efficiently learning complex or high-dimensional data to solve image classification problems. For the present research work, the proposed sequential CNN model has been composed of 5 convolutional layers, 3 max-pooling layers, and 1 fully connected layer. The architectural diagram of the proposed CNN model is shown in Figure 3.

2.3. Proposed CNN's Layer Description.

The proposed CNN model's layers have been classified into two types, namely, primary layers and secondary layers. The primary layers of the proposed CNN model include convolution layers, activation layers, pooling layers, flattening layers, and dense layers, which are the main layers used in the CNN. On the other hand, the secondary layers of the proposed CNN model are the optional layers that have been added to improve resistance to over-fitting of CNN networks, which is named as dropout layers and batch normalization layers. The details descriptions of CNN model's layers and learning parameters are presented in Table 2.

2.4. Automatic Feature Extraction.

For automatic extraction of features from an image, three layers in the proposed CNN model play a crucial role. These layers are the convolution layers, the max-pooling layer, and batch normalization. As discussed earlier, the proposed CNN model has been built using five convolutional layers, three max-pooling layers, and one layer of batch normalization after each convolutional operation. All five convolution layers of the proposed CNN model have been designed with the same size of filter (3×3) but with a different filter count in each convolution layer. For example, the first convolutional layer of the proposed CNN model has 32 weight filters, the second and third convolutional layers have 64 weight filters, and the fourth and fifth convolutional layers have 128 weight filters. A batch normalization layer has been added after each and every convolutional layer, as it is required to keep the output away from the saturation region using the mean and variance. Out of three max-pooling layers, the first layer of

FIGURE 2: (a) Blank sample sheet. (b) Sample sheet from writer 1. (c) Image converted from RGB to Gray. (d) Eroded image. (e) Cropped images of Gurmukhi Month's name. (f) Folders containing sorted dataset.

pooling comprises a filter size of 3×3 and the second and third pooling layers comprise a filter size of 3×3. In order to reduce over fitting, a dropout layer has also been used as the next layer, and it has been configured to randomly and temporarily remove 25% of the neurons. These layers all together form an automatic feature extractor for the proposed CNN model.

2.5. Classification. After feature extraction, the next step is data classification using the model. For this, initially flattening layer transforms the two-dimensional matrix of the features extracted through convolutional operation into a single vector and feeds it to the first fully connected layer of the model. The fully connected layer comprises 1024 neurons and a ReLu activation function. After this, a layer of batch normalization and a dropout layer of 0.5 have been added. Finally, a second fully connected layer has been added to the network with softmax as an activation function in order to classify handwritten word data into 24 classes.

2.6. Experimental Setup. For obtaining the results of simulation of Gurmukhi month's name classification using self-created handwritten word dataset, an experiment has been performed on Tesla K80 GPU that is virtually available on Google Colab, under the constrained of its limited use to maximum 12 hours for continuous training. The maximum limit to use the GPU also varies depending upon the traffic on GPU at that time [21, 24]. Along with this, the memory of other hardware RAM and ROM used for this experiment is 12.69 GB and 107.72 GB, respectively.

Furthermore, because of the availability of very useful libraries in Python [21], proposed CNN model's architectures have been built using Keras with a TensorFlow backend.

Other than this, for training the proposed CNN model, selected parameters are presented in Table 3.

For selecting an appropriate deep optimizer for the proposed CNN model's simulation on the Gurmukhi handwritten word dataset with an effective accuracy rate, performance compromise has been made among different

TABLE 1: Detail description of dataset with the number of training and testing samples selected for each class.

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	Class 10	Class 11	Class 12	Class 13	Class 14	Class 15	Class 16	Class 17	Class 18	Class 19	Class 20	Class 21	Class 22	Class 23	Class 24
Class name in English	Vaisakh	Jeth	Harh	Sawan	Bhado	Assu	Katak	Magar	Poh	Magh	Phagun	Chet	January	February	March	April	May	June	July	August	September	October	November	December
Class name in Gurumukhi	ਵੈਸਾਖ	ਜੇਠ	ਹਾੜ੍ਹ	ਸਾਉਣ	ਭਾਦੋਂ	ਅੱਸੂ	ਕੱਤਕ	ਮੱਘਰ	ਪੋਹ	ਮਾਘ	ਫੱਗਣ	ਚੇਤ	ਜਨਵਰੀ	ਫਰਵਰੀ	ਮਾਰਚ	ਅਪ੍ਰੈਲ	ਮਈ	ਜੂਨ	ਜੁਲਾਈ	ਅਗਸਤ	ਸਤੰਬਰ	ਅਕਤੂਬਰ	ਨਵੰਬਰ	ਦਸੰਬਰ
Duration	14 April to 14 May	15 May to 14 June	15 June to 15 July	16 July to 15 August	16 August to 14 September	15 September to 14 October	15 October to 13 November	14 November to 13 December	14 December to 12 January	13 January to 11 February	12 February to 13 March	14 March to 13 April	1 January to 31 January	1 February to 28/29 February	1 March to 31 March	1 April to 30 April	1 May to 31 May	1 June to 30 June	1 July to 31 July	1 August to 31 August	1 September to 30 September	1 October to 31 October	1 November to 30 November	1 December to 31 December
Number of training samples	795	801	799	782	802	809	813	789	798	788	802	819	770	817	804	786	803	822	797	788	805	804	808	799
Number of testing samples	205	199	201	218	198	191	187	211	202	212	198	181	230	183	196	214	197	178	203	212	195	196	192	201

FIGURE 3: Proposed CNN model's architecture.

TABLE 2: Layer's description of the proposed CNN model.

Model: "Sequential" layer (type)	Output shape	Param#
conv2d (Conv2D)	(None, 50, 50, 32)	320
activation (Activation)	(None, 50, 50, 32)	0
batch normalization (Batch Normalization)	(None, 50, 50, 32)	128
max_pooling2d (MaxPooling2D)	(None, 16, 16, 32)	0
dropout (Dropout)	(None, 16, 16, 32)	0
conv2d_1 (Conv2D)	(None, 16, 16, 64)	18496
activation_1 (Activation)	(None, 16, 16, 64)	0
batch_normalization_1 (Batch Normalization)	(None, 16, 16, 64)	256
conv2d_2 (Conv2D)	(None, 16, 16, 64)	36928
activation_2 (Activation)	(None, 16, 16, 64)	0
batch_normalization_2 (Batch Normalization)	(None, 16, 16, 64)	256
max_pooling2d_1 (Max-Pooling 2D)	(None, 8, 8, 64)	0
dropout_1 (Dropout)	(None, 8, 8, 64)	0
conv2d_3 (Conv2D)	(None, 8, 8, 128)	73856
activation_3 (Activation)	(None, 8, 8, 128)	0
batch_normalization_3 (Batch Normalization)	(None, 8, 8, 128)	512
conv2d_4 (Conv2D)	(None, 8, 8, 128)	147584
activation_4 (Activation)	(None, 8, 8, 128)	0
batch_normalization_4 (Batch Normalization)	(None, 8, 8, 128)	512
max_pooling2d_2 (MaxPooling2D)	(None, 4, 4, 128)	0
dropout_2 (Dropout)	(None, 4, 4, 128)	0
flatten (Flatten)	(None, 2048)	0
dense (Dense)	(None, 1024)	2098176
activation_5 (Activation	(None, 1024)	0
batch_normalization_5 (Batch Normalization)	(None, 1024)	4096
dropout_3 (Dropout)	(None, 1024)	0
dense_1 (Dense)	(None, 24)	24600
activation_6 (Activation)	(None, 24)	0

Total params: 2,405,720, trainable params: 2,402,840, and nontrainable params: 2,880.

optimizers such as stochastic gradient descent (SGD), Adagrad, Adadelta, RMSprop, Nadam, and Adam, and it has been observed that the proposed CNN model with the Adam optimizer has given the best accuracy, where the accuracy rate with other optimizers has been observed to be less than that of Adam.

Accuracy serves as the most fundamental and frequently used metric for measuring CNN models. However, F1 score, precision, and recall are all necessary to evaluate the mode's quality. As a result, for the performance assessment of the proposed CNN model, all of these parameters have been incorporated.

While selecting an appropriate learning rate, the proposed CNN model has been initially simulated at the highest learning rate, but the best result has been observed using 0.0001 learning rate in terms of accuracy.

3. Results and Analysis

This section will discuss the training parameters results of the proposed CNN model and the pretrained model in detail. Furthermore, this section will also perform a comparative analysis between the proposed CNN model and the pretrained transfer learning models (ResNet, VGG19, and VGG16) in terms of accuracy as well as various confusion

TABLE 3: Training parameters of the proposed model.

Adam Optimizer's specification	Learning rate (LR)	Loss function	Matrix	Number of epochs	Batch size (BS)
Learning rate = $1E - 03$, beta1 = 0.9, beta2 = 0.999, epsilon = $1e - 07$, and decay = learning rate/epochs	0.0001	Categorical cross entropy	Accuracy	100	20

matrix parameters in order to examine the performance of the proposed CNN model thoroughly.

3.1. Training Parameters' Results of the Proposed CNN Model.
In order to examine the performance of the proposed CNN model, the self-prepared handwritten word dataset has been simulated using the proposed CNN model at 100 and 40 epochs using 20 and 40 batch sizes, respectively. The various training parameters results of the proposed CNN model at different epochs and batch sizes are presented in Table 4.

From Table 4, it has been analyzed that the proposed CNN model at 100 epoch and 40 batch size gives the highest accuracy and minimum loss when simulated using a self-prepared handwritten word dataset. The values of the highest training and validation accuracy for the proposed CNN model on the handwritten word dataset are 99% and 99.73%, respectively.

On the other hand, the poor results have been analyzed for the proposed CNN model at 40 epochs using 20 batch sizes on the same dataset, whose values are 94.86% and 89.95% for training and validation accuracy, respectively.

The accuracy and loss graph for the proposed CNN model at different epochs (100 and 40) and batch sizes (20 and 40) is presented in Figure 4.

3.2. Training Parameters' Results of ResNet 50.
The ResNet50 is designed with 48 convolutional layers that belong to four stages and four convolutional blocks of the model's architecture. The filter sizes that have been used for ResNet's convolutional and max-pooling layers are 7×7 and 3×3. The complete architectural diagram of ResNet50 is presented in Figure 5.

The first stage of ResNet50 has 3 residual blocks and 3 layers in each block. Each layer in the first stage contains 64 or 256 feature maps of sizes of 1×1 or 3×3 for the convolution operation. The second stage of ResNet50 has also 3 residual blocks and each residual block has 4 layers of convolution. The numbers of the filters used to perform the operations are 128 or 512 in each layer. The model's third stage has 3 residual blocks, each with 6 layers. The numbers of filters that have been used in the third stage convolutional layers are either 256 or 1024. Finally, stage 4 of ResNet 50 consists of 3 residual blocks with 3 convolutional layers each. The sizes of the filters used to perform the operation are 512 and 2048. The last layer is the fully connected layer of the architecture that classifies handwritten word dataset into 24 classes.

The pretrained model ResNet 50 has been simulated on self-prepared handwritten dataset of Gurmukhi month's name at 100 and 40 epochs with 20 and 40 batch sizes, respectively. The training parameters' results obtained for this model is presented in Table 5.

From Table 5, it has been analyzed that the pretrained model ResNet 50 has performed the worst on the proposed handwritten words dataset of Gurmukhi month's name. The value of highest validation accuracy that has been obtained using this model is 40.15% at 100 epochs and 40 batch sizes. However, this accuracy result of ResNet 50 at 100 epochs and 40 batch sizes is comparatively higher than the performance of the model at 100 and 40 epochs using 20 batch sizes.

The accuracy loss graph of ResNet 50 at different epochs and batch sizes is presented in Figure 6.

3.3. Training Parameters' Results of VGG 19.
The architectural diagram of VGG 19 is presented in Figure 7. This model has been designed with 19 total layers, of which 16 layers are convolutional layers and 3 layers are fully connected layers. Other than this, the model has also contained 5 max pool layers. The size of filter that has been used for operation in VGG 19 is of the same size 3×3 for convolutional and pooling layers. The number of feature maps that has been used by each layer of VGG 19 is different and varies from 64 to 512. The last three layers are the fully connected layer of the architecture that classifies handwritten word dataset into 24 classes.

The pretrained model VGG19 has also been simulated on self-prepared handwritten dataset of Gurmukhi month's name at 100 and 40 epochs with 20 and 40 batch sizes, respectively. The training parameters' results obtained for this model is presented in Table 6.

It has been analyzed from Table 6 that the highest validation accuracy of pretrained mode VGG19 is 77.71% and 77.58%, respectively, when tested on 100 epochs with 20 and 40 batch sizes. Other than this, the least validation accuracy has been noticed for 40 epochs and 40 batch sizes, whose value is 73.17%. Hence, this model performed well on 20 and 40 batch sizes with 100 epochs in comparison to its performance of 40 epochs.

The VGG19's accuracy loss graph at different epochs and batch sizes has been presented in Figure 8.

3.4. Training parameters' Results of VGG 16.
The VGG 16 model's constructional design is shown in Figure 9. The model comprises 16 layers in total, out of which 13 layers are convolutional layers and 3 layers are fully connected layers. Other than this, VGG16 also consists of five max-pooling layers. The filter size throughout all of the layers of the architecture is of same size, that is, 3×3. However, the count for the feature maps has varied from 64 to 512 from the layer first to the last layer of the model. The last three layers of the model are the fully connected layers that classify the handwritten word dataset into 24 classes.

The training parameters' results obtained for VGG16 when it is simulated on a self-prepared handwritten dataset of Gurmukhi month's name at 100 and 40 epochs with 20 and 40 batch sizes, respectively, have been presented in Table 7.

Table 6 depicted that the model's highest validation accuracy has been obtained at 100 epochs with 20 and 40 batch sizes, whose value is 81.38% and 80.56%, respectively. In contrast to this, the least accuracy for the model has been observed at 40 epochs and 40 batch sizes with an accuracy value of 75.63%. Hence, this model performed well on 20 and 40 batch sizes with 100 epochs in comparison to its performance of 40 epochs.

TABLE 4: Training parameters' results of the proposed CNN model at different epochs and batch sizes.

Results	Epoch = 100		Epoch = 40	
	Batch size 20	Batch size 40	Batch size 20	Batch size 40
Training accuracy	0.9878	0.99	0.9486	0.9568
Validation accuracy	0.9973	0.9973	0.8985	0.9885
Training loss	0.0432	0.0313	0.1580	0.1440
Validation loss	0.0225	0.0188	1.0888	0.0404
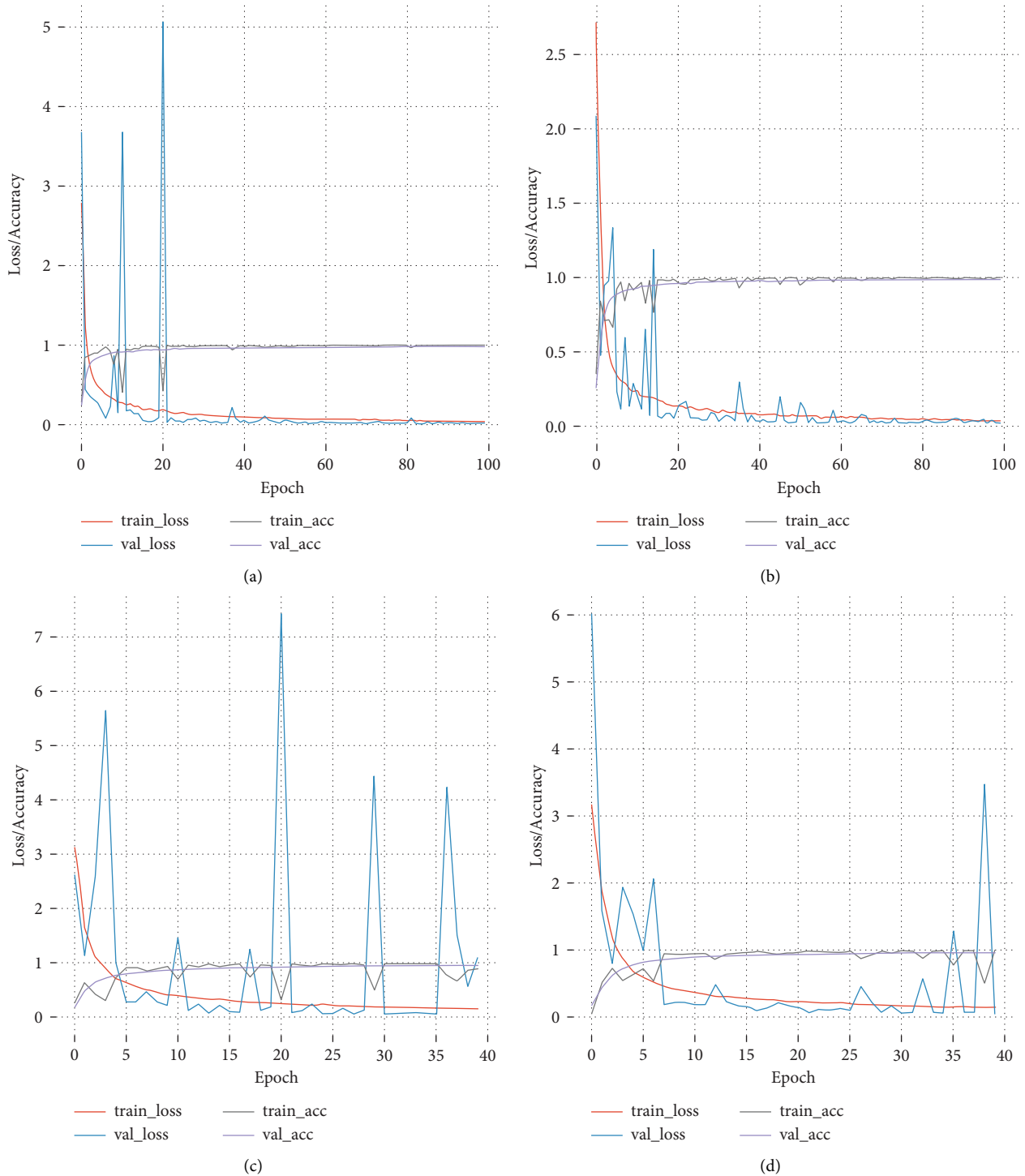
FIGURE 4: Accuracy loss graph of the proposed CNN model at (a) 100 epochs and 20 batch size, (b) 100 epochs and 40 batch size, (c) 40 epochs and 20 batch size, and (d) 40 epochs and 40 batch size.
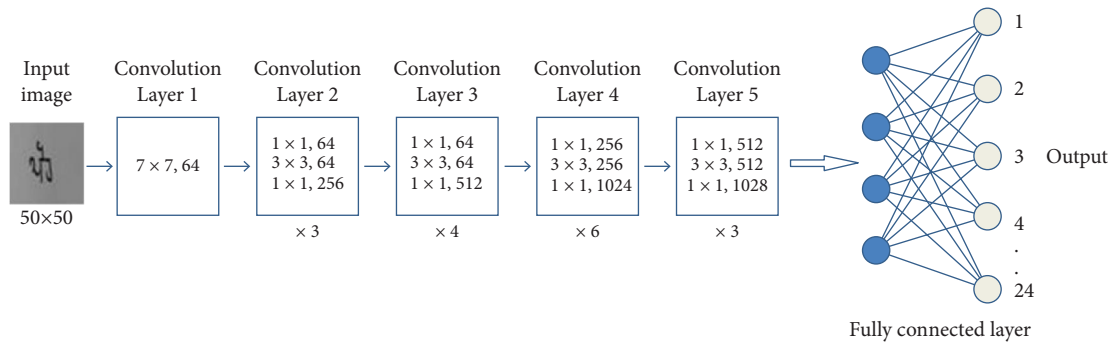
Figure 5: Architectural diagram of ResNet-50.

Table 5: Training parameters' results of ResNet50 at different epochs and batch sizes.

Results	Training parameters			
	Epoch = 100		Epoch = 40	
	Batch size 20	Batch size 40	Batch size 20	Batch size 40
Training accuracy	0.3299	0.3335	0.2503	0.2491
Validation accuracy	0.3929	0.4015	0.3031	0.3310
Training loss	2.1693	2.1658	2.4568	2.4608
Validation loss	1.9268	1.8842	2.2492	2.2037

(a)

(b)
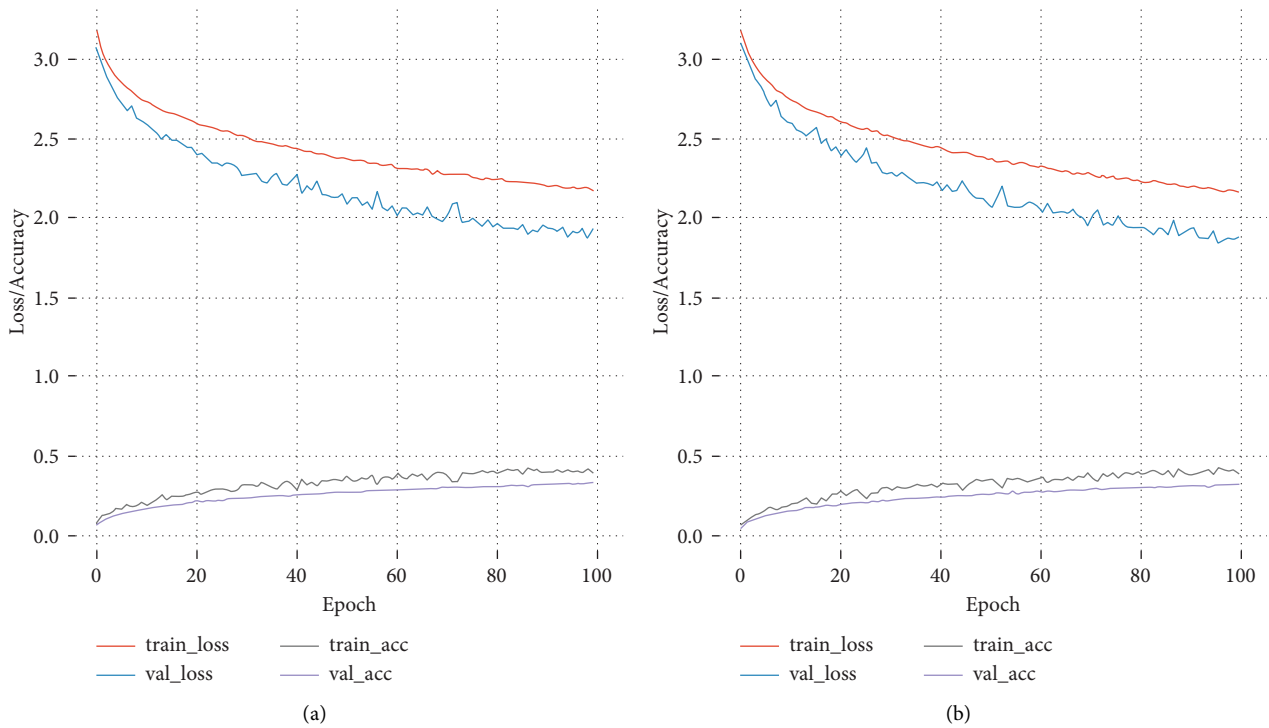
Figure 6: Continued.

(c)

(d)
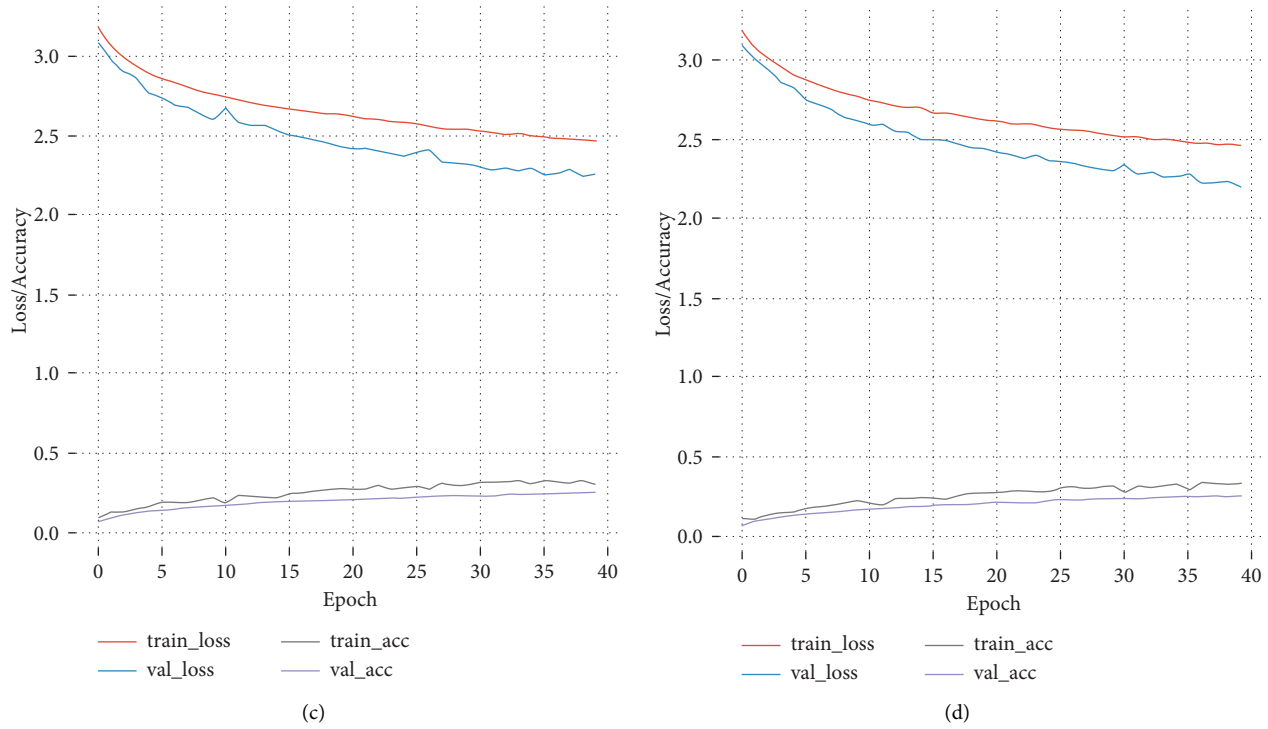
FIGURE 6: Accuracy loss graph of ResNet50 at (a) 100 epochs and 20 batch size, (b) 100 epochs and 40 batch size, (c) 40 epochs and 20 batch size, and (d) 40 epochs and 40 batch size.
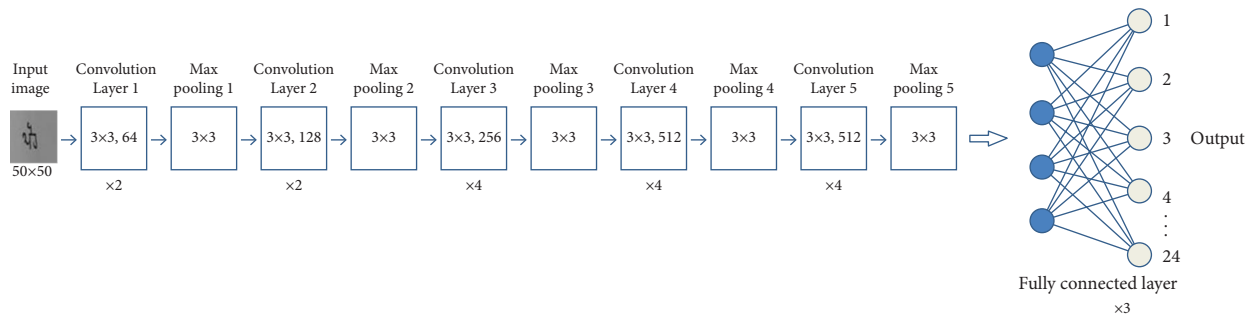
FIGURE 7: Architectural diagram of VGG19.

TABLE 6: Training parameters' results of VGG19 at different epochs and batch sizes.

Results	Training parameters			
	Epoch = 100		Epoch = 40	
	Batch size 20	Batch size 40	Batch size 20	Batch size 40
Training accuracy	0.7530	0.7452	0.6816	0.6793
Validation accuracy	0.7771	0.7758	0.7346	0.7317
Training loss	0.7560	0.7693	0.9775	0.9901
Validation loss	0.6647	0.6744	0.8217	0.8330

The accuracy loss graph for VGG16 at different epochs and batch sizes is presented in Figure 10.

It has been analyzed from the present section that the proposed CNN model as well as pretrained models including ResNet50, VGG19, and VGG16 have performed well on a self-prepared handwritten word dataset of Gurmukhi

month's name when simulated on 100 epochs and 40 batch sizes as compared to its performance in experimental settings on other training parameters such as 20 batch sizes with 100 epochs and 40 epochs with 20 and 40 batch sizes, respectively. Hence, in order to further assess the proposed CNN model's quality, a comparative analysis has been
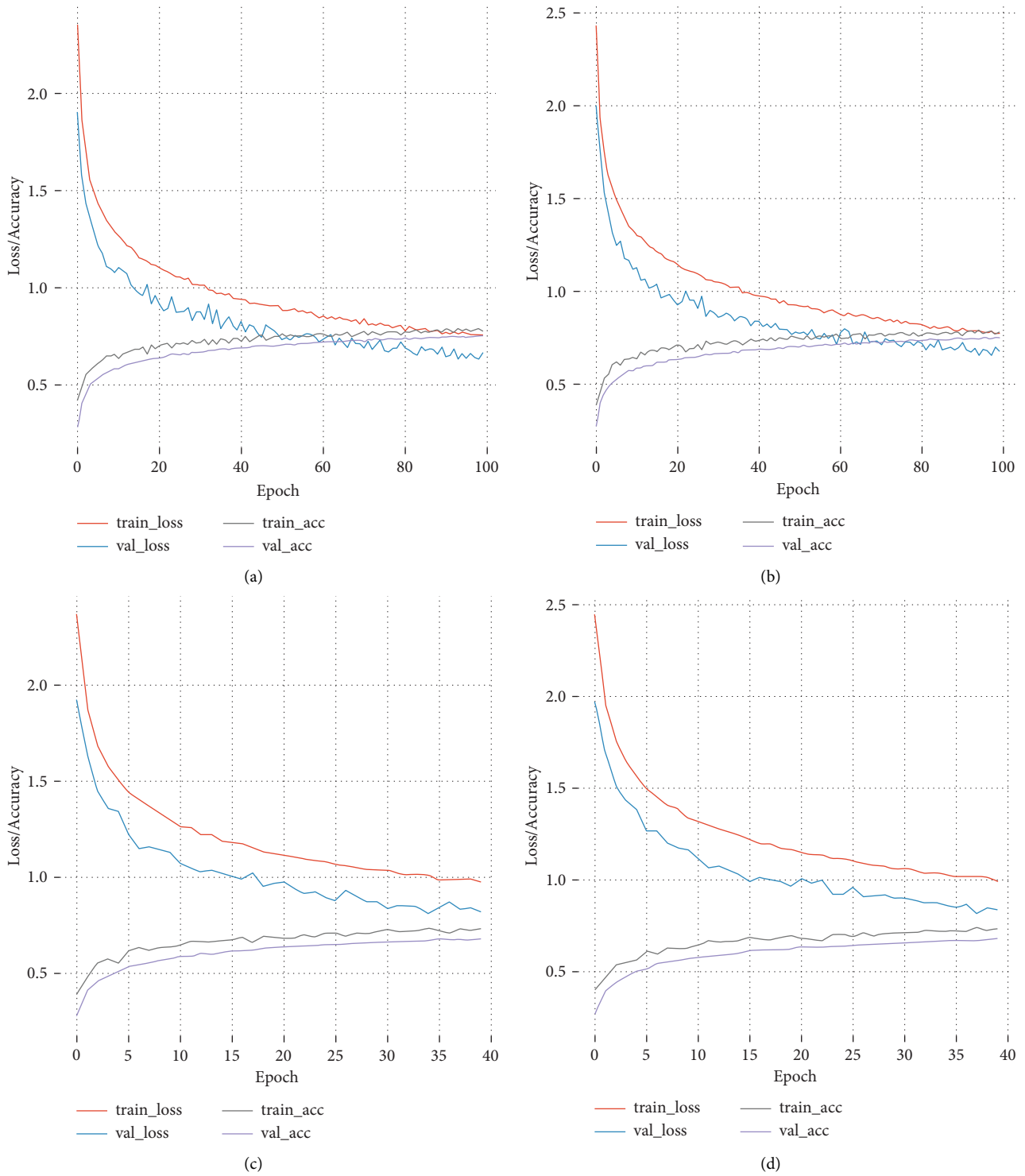
Figure 8: Accuracy loss graph of VGG19 at (a) 100 epochs and 20 batch size, (b) 100 epochs and 40 batch size, (c) 40 epochs and 20 batch size, and (d) 40 epochs and 40 batch size.

performed between the proposed model and a pretrained model at 100 epochs and 40 batch sizes using widely used evaluation metrics such as accuracy, F1, recall, and precision in Section 3.5.

3.5. Comparative Analysis at 100 Epochs and 40 Batch Sizes. It has been already motioned in the previous section that the proposed CNN model as well as pretrained models including ResNet50, VGG19, and VGG16 has been performed well on
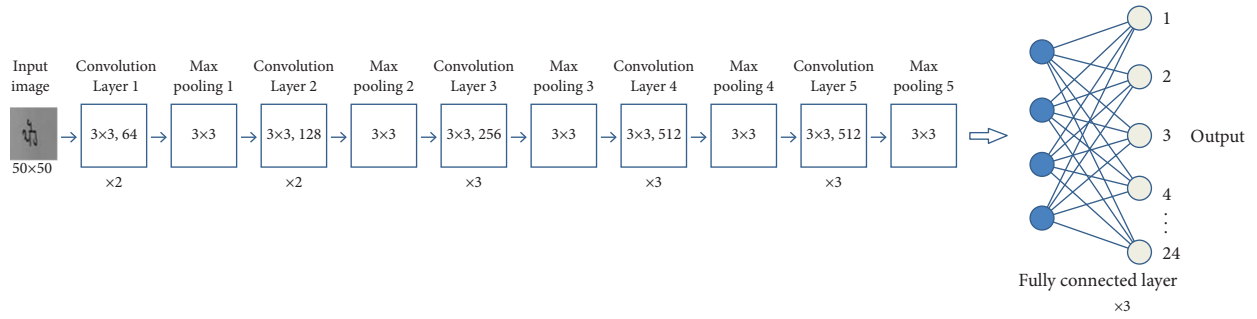
FIGURE 9: Architectural diagram of VGG16.

TABLE 7: Training parameters' results of VGG16 at different epochs and batch sizes.

Results	Training parameters			
	Epoch = 100		Epoch = 40	
	Batch size 20	Batch size 40	Batch size 20	Batch size 40
Training accuracy	0.7925	0.7832	0.7265	0.7212
Validation accuracy	0.8138	0.8056	0.7758	0.7563
Training loss	0.6274	0.6515	0.8382	0.8525
Validation loss	0.5484	0.5720	0.6908	0.7279

(a)

(b)
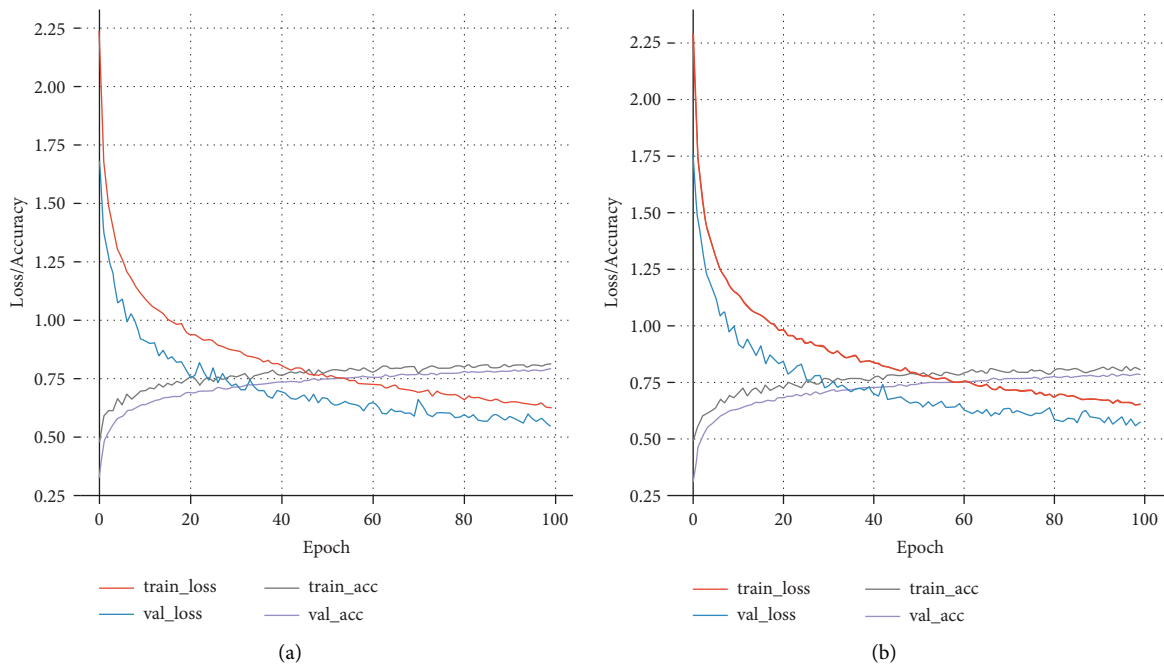
FIGURE 10: Continued.

(c)

(d)
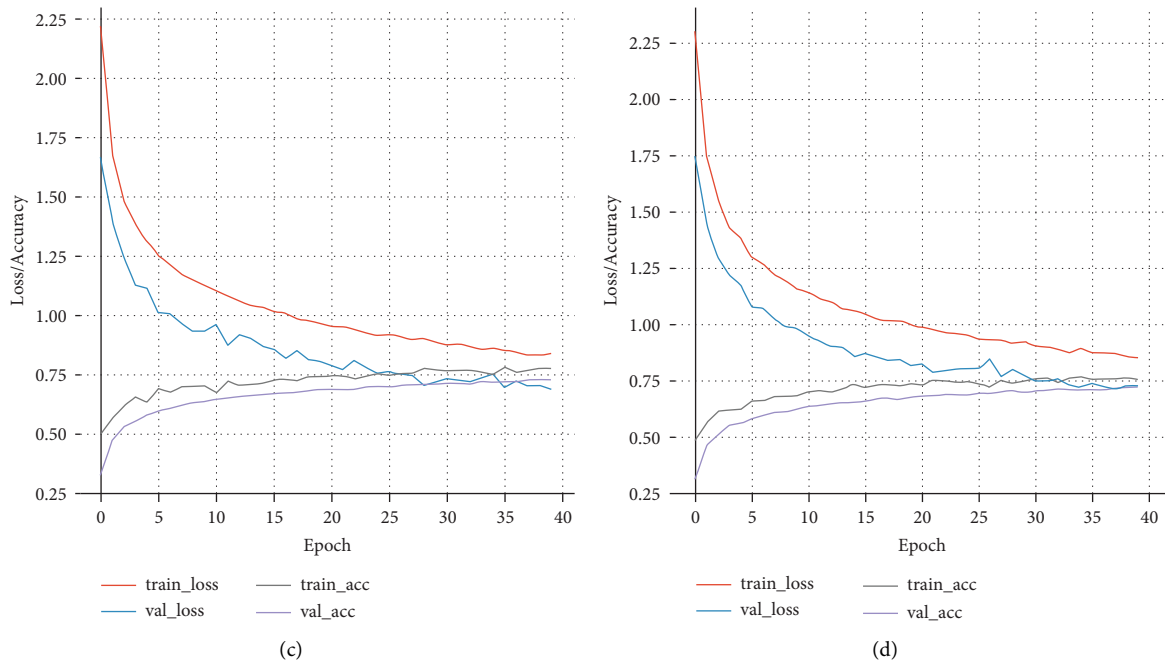
FIGURE 10: Accuracy loss graph of VGG16 at (a) 100 epochs and 20 batch size, (b) 100 epochs and 40 batch size, (c) 40 epochs and 20 batch size, and (d) 40 epochs and 40 batch size.

the self-prepared handwritten word dataset of Gurmukhi month's name on 100 epochs and 40 batch sizes. Hence, for the comparative analysis between the models in this section, the same numbers of epochs and batch sizes have been chosen.

Furthermore, for this comparative analysis, accuracy, F1, recall, and precision have been taken as evaluation parameters.

3.5.1. Comparative Analysis in Term of Accuracy. The 24 class classification results of Gurmukhi month"s name dataset on the proposed model and pretrained models have been obtained using 100 epochs and 40 batch sizes. As per the obtained results, it has been noticed that the proposed CNN model has outperformed the other three pretrained models in terms of classification accuracy. The accuracy of the proposed model is 99.73%, which is almost 59.58%, 22.15%, and 19.17% more than the accuracy obtained using ResNet50, VGG19, and VGG16, respectively, as shown in Figure 11.

Other than this, Figure 12 has shown the classwise accuracy comparison of the proposed model and pretrained models.

From Figure 12, it has been observed that the proposed CNN model has given the 100% classification accuracy result for "April," "Assu," "June," "Magar," "March," "November," "Phagun" and "Poh" classes of Gurmukhi months, where the other three models such as ResNet50, VGG19, and VGG16 give comparatively poor accuracy on the same classes.

The least accuracy of the proposed model is on "May" and "Sawan" classes of Gurmukhi months, which is about 99.92%.
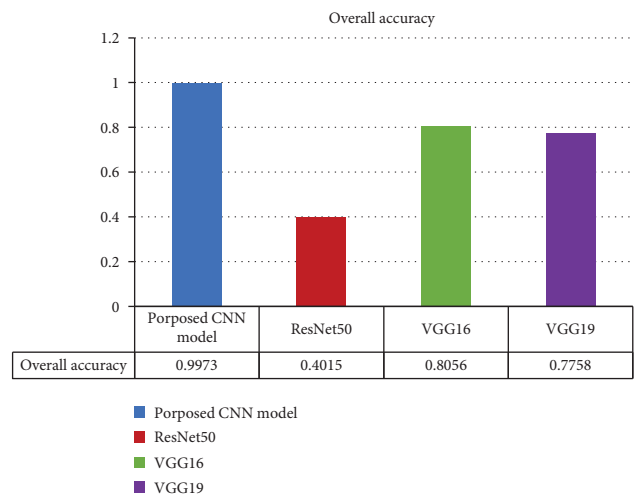
FIGURE 11: Comparison based on overall accuracy.

3.5.2. Comparative Analysis in Term of F1 Score. In order to test the performance of the proposed CNN model, a comparative analysis has been conducted based on the F1 score in this section. Figure 13 shows the comparison between the proposed CNN model and the other three pretrained models based on overall F1 score. It has been observed that the F1 score for the proposed CNN model has been 0.9973, which is 0.6006, 0.222, and 0.1950 higher than the F1 score value of ResNet50, VGG19, and VGG16, respectively.

In addition to this, a classwise F1 score comparison between the proposed CNN model and pretrained models is shown in Figure 14. It has been observed from Figure 14 that the proposed CNN model's F1 score resulting in 1 for

	April	Assu	August	Bhado	Chet	December	February	Harh	Junuray	Jeth	July	June	Katak	Magar	Magh	March	May	November	October	Phagun	Poh	Sawan	September	Vaisakh
■ Porposed CNN model	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
■ ResNet50	0.64	0.70	0.42	0.36	0.43	0.17	0.51	0.51	0.30	0.39	0.45	0.56	0.08	0.39	0.16	0.26	0.47	0.23	0.64	0.15	0.50	0.44	0.25	0.51
■ VGG16	0.87	0.93	0.89	0.86	0.93	0.76	0.75	0.96	0.73	0.66	0.86	0.92	0.75	0.80	0.89	0.81	0.82	0.58	0.87	0.73	0.83	0.86	0.47	0.79
■ VGG19	0.84	0.73	0.89	0.90	0.79	0.73	0.67	0.92	0.78	0.71	0.72	0.89	0.75	0.84	0.89	0.89	0.79	0.42	0.82	0.79	0.84	0.75	0.47	0.77

FIGURE 12: Comparison based on classwise accuracy.

	Porposed model	ResNet50	VGG16	VGG19
Overall f1	0.9973	0.396789097	0.802301975	0.775362205

■ Porposed model
■ ResNet50
■ VGG16
■ VGG19
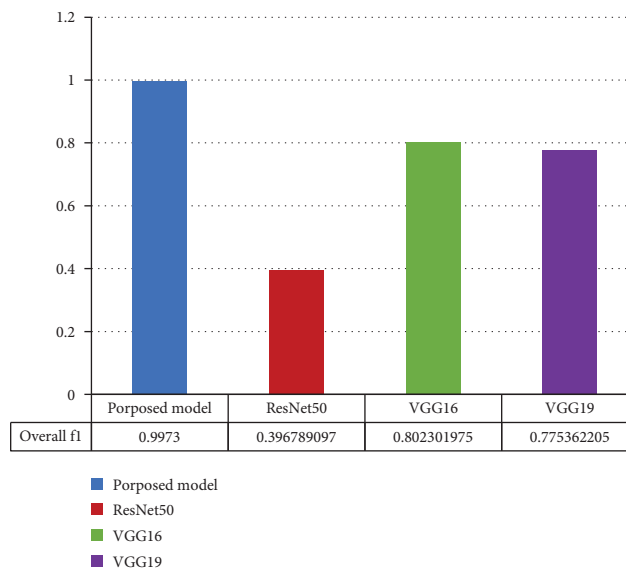
FIGURE 13: Comparison based on overall F1 score.

"April," "Assu," "June," "Magar," "March," "November," "Phagun," and "Poh" month names, where the results for F1 score for the same classes from other three pretrained model have been comparatively low.

On the other hand, the least value of F1 score from the proposed CNN model has been obtained for "Vaisakh," which is around 0.9951.

3.5.3. Comparative Analysis in Term of Recall. A comparison between the proposed CNN model and pretrained networks based on results of overall recall and classwise recall values is shown in Figures 15 and 16, respectively.

As per Figure 15, the proposed CNN model has outperformed the other three pretrained models in terms of the recall value when tested on handwritten word dataset of Gurmukhi months. The overall recall for the proposed CNN model is 0.9973 which is higher than the recall value of ResNet 50, VGG19, and VGG 16 by 0.5952, 0.2216, and 0.1931, respectively.

From Figure 16, it has been clear that the classwise recall score is resulting in 1 for "April," "Assu," "Bhado," "Harh," "January," "June," "Magar," "Magh," "March," "May," "November," "Phagun," "Poh," "September," and "Vaisakh" month names in case of the proposed CNN model, where F1 score's results is comparatively poor for the same classes from other three pretrained models.

	Apr-il	Ass-u	Aug-ust	Bha-do	Che-t	Dec-emb-er	Feb-rua-ry	Har-h	Jun-ura-y	Jet-h	July	June	Kat-ak	Ma-gar	Ma-gh	Ma-rch	Ma-y	Nov-emb-er	Oct-obe-r	Pha-gun	Poh	Saw-an	Sep-tem-ber	Vai-sak-h
■ Porposed CNN model	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	1.00	0.99	1.00	0.99	1.00	0.98	1.00	0.99	1.00	1.00	0.99	0.99	0.99
■ ResNet50	0.39	0.29	0.53	0.36	0.50	0.26	0.47	0.55	0.38	0.39	0.51	0.51	0.13	0.45	0.24	0.34	0.42	0.29	0.53	0.21	0.44	0.40	0.26	0.55
■ VGG16	0.89	0.90	0.87	0.85	0.63	0.49	0.77	0.94	0.76	0.77	0.84	0.92	0.79	0.81	0.89	0.83	0.84	0.60	0.90	0.74	0.86	0.82	0.58	0.86
■ VGG19	0.82	0.81	0.87	0.79	0.69	0.49	0.68	0.93	0.71	0.78	0.74	0.90	0.75	0.78	0.91	0.85	0.81	0.56	0.89	0.72	0.85	0.78	0.52	0.84

FIGURE 14: Comparison based on classwise F1 score.

	Porposed model	ResNet50	VGG16	VGG19
Overall recall	0.9973	0.4021	0.8042	0.7757

■ Porposed model
■ ResNet50
■ VGG16
■ VGG19
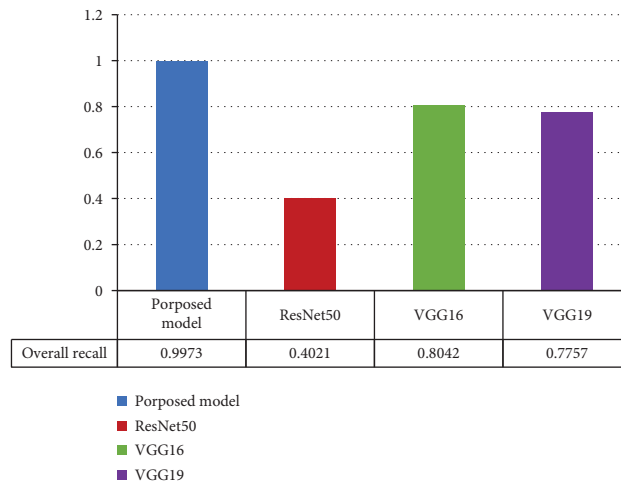
FIGURE 15: Comparison based on overall recall.

It has also been observed that the minimum value of recall is 0.9905, which is for "August" class name using the proposed CNN model.

3.5.4. Comparative Analysis in Term of Precision.
In this section, comparative analysis in terms of the precision value has been performed for 24-class classification results of Gurmukhi month's name dataset using the proposed model and pretrained models at 100 epochs and 40 batch sizes. The comparative analysis in Figure 17 depicted that the proposed CNN model has outperformed the other three pretrained models in terms of precision. As the results of precision using the proposed CNN model is 0.9974, which is almost 0.5446, 0.2026, and 0.1801 more than the precision value obtained using ResNet50, VGG19, and VGG16, respectively, as shown in Figure 17.

A classwise precision comparison between the proposed CNN model and pretrained models, ResNet 50, VGG19, and VGG16, is shown in Figure 18. It has been observed from Figure 18 that, for class name "April," "Assu," "August," "Chet," "December," "February," "Jeth," "July," "June," "Katak," "Magar," "March," "November," "October," "Phagun," and "Poh," the obtained value of precision is one using the proposed CNN model. It has also been observed that, by using other three pretrained models, the precision value has been comparatively low for these classes.

Using the proposed CNN model, the least precision value (0.9903) has been obtained for class name "Vaisakh."

From the comparative analysis performed in this section, it has been found that the proposed CNN model has outperformed the other three pretrained models in terms of four performance assessment matrices such as accuracy, recall, precision, and F1 score when simulated on

	April	Assu	August	Bhado	Chet	December	February	Harh	Junuray	Jeth	July	June	Katak	Magar	Magh	March	May	November	October	Phagun	Poh	Sawan	September	Vaisakh
Porposed model	1	1	0.99	1	0.99	1	0.99	1	1	0.99	1	1	0.99	1	1	1	1	1	0.99	1	1	0.99	1	1
ResNet50	0.64	0.70	0.42	0.36	0.43	0.17	0.51	0.51	0.30	0.39	0.45	0.56	0.08	0.39	0.16	0.26	0.47	0.24	0.64	0.15	0.51	0.44	0.25	0.51
VGG16	0.89	0.86	0.85	0.84	0.47	0.36	0.77	0.91	0.79	0.90	0.82	0.91	0.85	0.81	0.87	0.85	0.85	0.61	0.92	0.74	0.89	0.77	0.73	0.92
VGG19	0.79	0.91	0.86	0.71	0.61	0.36	0.68	0.91	0.66	0.85	0.77	0.91	0.75	0.72	0.92	0.81	0.82	0.79	0.92	0.65	0.84	0.77	0.57	0.92

FIGURE 16: Comparison based on classwise recall.

	Porposed model	ResNet50	VGG16	VGG19
Overall precision	0.9974	0.4528	0.8173	0.7908

■ Porposed model
■ ResNet50
■ VGG16
■ VGG19
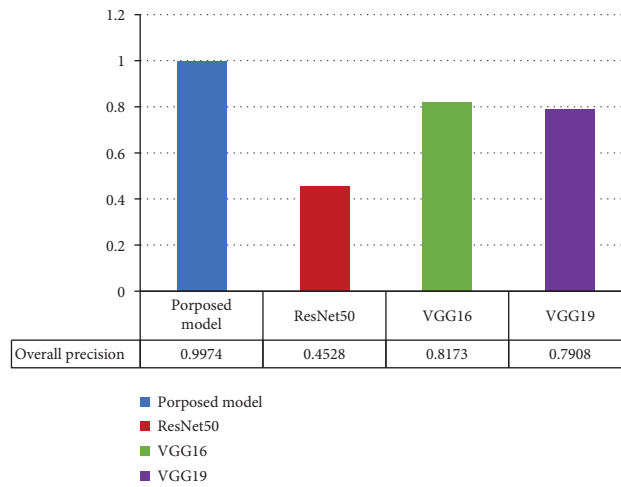
FIGURE 17: Comparison based on overall precision.

	April	Assu	August	Bhado	Chet	December	February	Harh	Junuray	Jeth	July	June	Katak	Magar	Magh	March	May	November	October	Phagun	Poh	Sawan	September	Vaisakh
Porposed model	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ResNet50	0.3	0.2	0.7	0.4	0.6	0.5	0.4	0.6	0.5	0.4	0.6	0.5	0.4	0.5	0.5	0.5	0.4	0.4	0.5	0.3	0.4	0.4	0.3	0.6
VGG16	0.9	0.9	0.9	0.9	0.9	0.8	0.8	1	0.7	0.7	0.9	0.9	0.8	0.8	0.9	0.8	0.8	0.6	0.9	0.7	0.8	0.9	0.5	0.8
VGG19	0.9	0.7	0.9	0.9	0.8	0.7	0.7	1	0.8	0.7	0.7	0.9	0.8	0.9	0.9	0.9	0.8	0.4	0.9	0.8	0.9	0.8	0.5	0.8

FIGURE 18: Comparison based on classwise precision.

a self-prepared handwritten word dataset of Gurmukhi month's name for image classification. Hence, the proposed CNN model is the best fit model for various regional applications of month's name classification.

4. Conclusion

In this research article, the classification of a handwritten word dataset of Gurmukhi month's name has been done into 24 different classes. For this, the CNN model has been prepared from scratch with five convolutional, three polling layers, one flattening layer, and one dense layer. Furthermore, a performance assessment of the proposed CNN model has been conducted at different numbers of epochs and batch sizes, and it has been observed that the proposed CNN model has performed well at epoch 100 and 40 batch sizes. In addition to this, a comparative analysis has also been performed between the proposed CNN model and three pretrained models named as ResnEt50, VGG19, and VGG16 based on different performance assessment matrixes such as accuracy, F1 score, recall, and precision. From this experimentation, it has been concluded that the proposed CNN model at 100 epochs and 40 batch sizes outperformed the other three pretrained models in the classification of handwritten words dataset in 24 classes.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. Sharma, R. Kumar, and R. K. Sharma, "Online handwritten gurmukhi character recognition using elastic matching," *2008 Congress on Image and Signal Processing*, vol. 2, pp. 391–396, 2008.

[2] A. Sharma, R. Kumar, and R. K. Sharma, "Rearrangement of recognized strokes in online handwritten gurmukhi words recognition," in *Proceedings of the 2009 10th International Conference on Document Analysis and Recognition*, Barcelona, Spain, July 2009.

[3] R. Dhir, "Moment based invariant feature extraction techniques for bilingual character recognition," in *Proceedings of the 2010 2nd International Conference on Education Technology and Computer*, Shanghai, China, June 2010.

[4] M. Kumar, M. K. Jindal, and R. K. Sharma, "Classification of characters and grading writers in offline handwritten Gurmukhi script," in *Proceedings of the 2011 International Conference on Image Information Processing*, Shimla, India, September 2011.

[5] M. Kumar, R. K. Sharma, and M. K. Jindal, "Efficient feature extraction techniques for offline handwritten gurmukhi character recognition," *National Academy Science Letters*, vol. 37, no. 4, pp. 381–391, 2014.

[6] R. Kumar, R. K. Sharma, and A. Sharma, "Recognition of multi-stroke based online handwritten gurmukhi aksharas,"

Proceedings of the National Academy of Sciences, India, Section A: Physical Sciences, vol. 85, no. 1, pp. 159–168, 2015.

[7] K. Verma and R. K. Sharma, "Recognition of online handwritten Gurmukhi characters based on zone and stroke identification," *Sadhana*, vol. 42, no. 5, pp. 701–712, 2017.

[8] M. Kumar, M. K. Jindal, and R. K. Sharma, "Offline handwritten gurmukhi character recognition: analytical study of different transformations," *Proceedings of the National Academy of Sciences, India, Section A: Physical Sciences*, vol. 87, no. 1, pp. 137–143, 2017.

[9] H. Singh, R. K. Sharma, and V. P. Singh, "Efficient zone identification approach for the recognition of online handwritten Gurmukhi script," *Neural Computing & Applications*, vol. 31, no. 8, pp. 3957–3968, 2019.

[10] M. Kumar, M. K. Jindal, R. K. Sharma, and S. R. Jindal, "A novel framework for writer identification based on presegmented Gurmukhi characters," *Sadhana*, vol. 43, no. 12, 2018.

[11] Sakshi, N. K. Garg, and M. Kumar, "Writer identification system for handwritten gurmukhi characters: study of different feature-classifier combinations," in *Proceedings of International Conference on Computational Intelligence and Data Engineering*, Springer, Singapore, 2018.

[12] M. Kumar, S. R. Jindal, M. K. Jindal, and G. S. Lehal, "Improved recognition results of medieval handwritten gurmukhi manuscripts using boosting and bagging methodologies," *Neural Processing Letters*, vol. 50, no. 1, pp. 43–56, 2019.

[13] A. Garg, M. K. Jindal, and A. Singh, "Degraded offline handwritten Gurmukhi character recognition: study of various features and classifiers," *International Journal of Information Technology*, vol. 14, no. 1, pp. 145–153, 2019.

[14] M. Kumar, M. K. Jindal, R. K. Sharma, and S. R. Jindal, "Performance evaluation of classifiers for the recognition of offline handwritten Gurmukhi characters and numerals: a study," *Artificial Intelligence Review*, vol. 53, no. 3, pp. 2075–2097, 2020.

[15] T. Ganji, M. S. Velpuru, and R. Dugyala, "Multi variant handwritten Telugu character recognition using transfer learning," *IOP Conference Series: Materials Science and Engineering*, vol. 1042, no. 1, Article ID 012026, 2021.

[16] R. Pramanik and S. Bag, "Handwritten Bangla city name word recognition using CNN-based transfer learning and FCN," *Neural Computing & Applications*, vol. 33, no. 15, pp. 9329–9341, 2021.

[17] R. Pramanik, P. Dansena, and S. Bag, "A study on the effect of CNN-based transfer learning on handwritten indic and mixed numeral recognition," in *Document Analysis and Recognition*, Springer, Singapore, 2019.

[18] U. Jindal, S. Gupta, V. Jain, and M. Paprzycki, "Offline handwritten Gurumukhi character recognition system using deep learning," in *Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals*, Springer, Singapore, 2020.

[19] R. Geetha, T. Thilagam, and T. Padmavathy, "Effective offline handwritten text recognition model based on a sequence-to-sequence approach with CNN–RNN networks," *Neural Computing & Applications*, vol. 33, no. 17, Article ID 10923, 2021.

[20] S. Singh, A. Sharma, and V. K. Chauhan, "Online handwritten Gurmukhi word recognition using fine-tuned Deep Convolutional Neural Network on offline features," *Machine Learning with Applications*, vol. 5, Article ID 100037, 2021.

[21] M. Ragab, A. S. A.-M. Al-Ghamdi, B. Fakieh, H. Choudhry, R. F. Mansour, and D. Koundal, "Prediction of diabetes

through retinal images using deep neural network," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7887908, 6 pages, 2022.

[22] S. Shambhu, D. Koundal, P. Das, and C. Sharma, "Binary classification of COVID-19 CT images using CNN: COVID diagnosis using CT," *International Journal of E-Health and Medical Communications*, vol. 13, no. 2, pp. 1–13, 2021.

[23] R. Nair, A. Alhudhaif, D. Koundal, R. I. Doewes, and P. Sharma, "Deep learning-based COVID-19 detection system using pulmonary CT scans," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 29, no. 1, pp. 2716–2727, 2021.

[24] D. J. Jwo, S. F. Chiu, S. Gupta et al., "Deep learning based automated detection of diseases from apple leaf images," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1849–1866, 2022.

Research Article

Research on the Relationship between Motion Performance and User Experience of Golf Virtual Simulation Putting Simulator

Qihua Zhang,[1] **Zhihui Li** ⬤,[1] **and Mingjun Liu**[2]

[1]*Hebei Sport University, Shijiazhuang, Hebei 050000, China*
[2]*Huaxin College of Hebei GEO University, Shijiazhuang, Hebei 050000, China*

Correspondence should be addressed to Zhihui Li; lizhihui@hepec.edu.cn

This paper designs and develops a virtual golf simulation putting simulator based on the existing computer technology and conducts in-depth research and analysis on the relationship between its motion performance and user experience. The network architecture of the distributed virtual golf simulation system and the scene data management model are established, based on which the server-side system design and the client-side network communication module design of the distributed virtual golf simulation system are carried out. In the requirement analysis, the functional requirements such as building VR scenes, data communication and recognition models, and the non-functional requirements such as system security and ease of use are analyzed; in the outline design, the hardware equipment and logical architecture of the automatic user experience optimization system are described; in the detailed design, the functional modules of the system are designed in detail, including VR induction experience, physiological signal dataset user experience identification, data communication, optimization strategy, and so on, and important class diagrams and flowcharts are given. The intervention effects of positive thinking training on sports performance and improving athletes' attention and receptivity have been verified and recognized by coaches and athletes. The putting simulator in the experimental class had higher hole-in-hole parameters than the control class, a highly significant difference; the putting simulator in the experimental class had higher hole-in-hole parameters than the control class, with a highly significant difference. These 3D models may contain more detailed information. In a virtual scene, the more detailed information a model contains, the more polygons the model needs, so that the computer needs to draw many polygons per frame, which has a great impact on the real-time performance of scene drawing. The parameters of the 5-yard chip-and-shoot in the experimental class were higher than those in the control class, and there was a very significant difference between the parameters of the 15-yard chip-and-shoot in the experimental class and those in the control class. The experimental results show that the model optimization processing method and rendering acceleration technology proposed in this paper can largely improve the rendering efficiency of 3D virtual scenes.

1. Introduction

In today's high-speed economic development, people's demand for sports is getting higher and higher, and golf, which was once an aristocratic sport, is slowly appearing in the vision of popular sports, becoming one of the important sports for the public to participate in sports activities, thus enriching people's daily life. Because golf is the fourth most difficult sport in the world and belongs to the mental skill type of sports, to play golf well, you need to have more

excellent golf-specific skills [1]. Golf technical movements are complex, especially the process of putting to hitting the ball, which seems simple but is complicated. To effectively master the basic techniques of this sport and accurately hit the ball to obtain the ideal sports performance, practitioners need to have good physical strength quality and effective movement control ability. Core strength training is especially important due to the technical force structure of golf. At present, most colleges and universities teach golf only in golf professional technical training and only pay attention to

the demonstration of movements, ignoring the fundamental reasons why students cannot complete standard movements due to lack of strength and unstable movements, resulting in the slow improvement of students' technical level. Therefore, the college golf course should focus on strengthening the core strength training of golf in addition to the professional training of students. After the golf-specific core strength training, the subjects have a better effect on the proficiency in golf sports skills; especially in the process of core power chain transmission, the control and performance of technical movements are improved more obviously, thus enhancing the stability of golf putting technology, improving the accuracy of hitting the ball, and improving the technical level of students [2]. The expert group has more obvious advantages and higher accuracy. Therefore, how to conduct effective core power training for golf students and how to select and evaluate the evaluation index of core power training in golf becomes the important and difficult point of this paper.

Putting is the most basic and important action in golf, and the special strength required is an important quality in the basic strength training of golfers. The accuracy of the putting stroke is very important in golf, and whether the putting stroke is standard or not is directly related to the distance and effect of the shot. The characteristics of putting action technology are mainly reflected in the high degree of coordination and unity of swing skill and individual swing skill. The entire process requires the organic cooperation of the overall rotation of the body and upper limb links [3]. Professional golfers are very concerned about their strength, balance, rhythm, and speed in the putting process, so the practice of putting action becomes the key to improving the level of golfers, and putting allows us to appreciate the leisure and relaxation of golf. Analyzing the players' putts only through their feelings or the coach's observation will cause many difficulties in golf training, and it is not easy to improve their putts [4].

Our golf trainer software combines laser infrared sensing technology, data acquisition technology, 3D image processing technology, and computer software technology and is used for golf putting and putting training, which can capture the angle of the stroke, the speed of the ball, and the rotation of the ball through laser infrared sensing. Players can adjust their next hitting force and angle of the ball in real time. The golf trainer software simulates the three-dimensional terrain of the golf course and processes the collected data to display the movement of the golf ball and the trajectory of the golf ball using three-dimensional technology and calculates the distance of the golf ball movement. It can visually display the player's ball striking index. To achieve a better training effect, the course should be as realistic as possible. The golf trainer software simulates a three-dimensional course so that the player can be integrated into the real golf course as much as possible.

2. Related Works

Positive thinking training combines Buddhist Zen thought with third-generation behaviorism as its basis. It advocates an appropriate understanding of current physical and mental conditions, acceptance of negative emotions, acceptance of such internal senses, and understanding of their meaning [5]. In the study of virtual simulation technology in physical education theory, Soltani and other scholars made a specific analysis of the feasibility of using virtual simulation technology in physical education based on a brief introduction of virtual reality technology [6]. It is proposed that the use of virtual reality systems in physical education can make up for the shortage of teaching venues, funds, and equipment (such as students' favorite shooting, rock climbing, diving, wilderness survival, trampoline, and so on), and students can practice in the virtual environment, and it is proposed that the technology can optimize the physical education process [7]. The experimental results show that golf experts can use their previous experience to adapt more quickly when facing different golf positions, to maintain good sports performance and higher putting accuracy. The conclusion that the technology can optimize the physical education process, improve the physical education environment, improve the level of teaching and training, and cultivate innovative talents has a far-reaching impact. Uhm believes that virtual physical education is an intelligent, distributed, interactive, and graphic teaching that is not limited by time and space and is completely different from traditional physical education methods [8]. González-Rodríguez et al. discussed the composition of the virtual system of sports scenes in exploring new methods and means of action technique analysis and sports training and were able to provide a flexible and safe environment for sports training [9].

The research on virtual golf simulation systems is mainly focused on data sensor technology and graphic simulation technology [10]. Data sensors are mainly used to collect accurate golf ball movement data or human hitting posture data in real time, and the main technologies used by foreign golf simulator manufacturers are radar sensor technology, infrared detection technology, and high-speed camera acquisition technology, among which 3D stereo high-speed camera acquisition technology is the most advanced [11]. Graphics simulation technology is mainly to simulate more realistic and natural golf natural scenes and more accurate golf physical movement effects, and there are many options for virtual golf simulation, such as the use of a self-developed graphics development engine or the use of a mature game engine [12].

The purpose of this paper is to study the implementation of a virtual golf course in which several different users located in different geographical locations can compete in the same virtual golf course through network interconnection. Since this paper must solve such problems as real-time interaction between different users, spatial and temporal consistency of virtual environment, management of massive scene data, network topology and network communication protocol of the system, simulation of virtual golf scene effects, optimization of large-scale virtual scene rendering, and so on, which are hotspots and difficulties of distributed virtual reality system, the research of this paper has important theoretical significance. At the same time, since the distributed golf simulation system overcomes the problem

that the stand-alone golf simulation system cannot be networked across regions for interaction, the distributed golf simulation system has higher use value and commercial value.

3. Golf Virtual Simulation Putting Simulator Design

Since the virtual golf simulation system is a realistic simulation of the golf ball flight scene, the golf simulation system should firstly have a highly realistic three-dimensional scenery simulation function; secondly, to truly display the golf player's shot level, the system needs to have the golf ball data collection function and the function to truly simulate the physical movement process of the golf ball in the virtual scene. In addition, to connect the virtual scene with the real interaction, the system must provide a friendly user interaction interface, and the data generated during the user interaction need to be saved and managed [13]. To enhance the user experience of the virtual golf simulation system, the system needs to add some additional functions, such as virtual 3D sound, multi-screen display, and other functions.

The data reception and parsing are mainly based on the API library of radar sensors provided by FlightScope. This part is mainly responsible for receiving and parsing the radar data and providing the initial data for the golf ball physical motion simulation module of the system.

The main function of the golf trainer software system is golf training, but of course, it can also play the role of entertainment. The system achieves training by establishing a driving range scene and a putting range scene. Putting data are collected in the practice range 3D scene through serial transmission, mainly including initial speed, angle, declination, and so on. Different changes can be made to the single-subject experiment to meet different experimental needs. A series of algorithms are used to simulate the putt data and display the trajectory of the ball after the putt. In this way, after each putt, the player can observe the distance of the ball to grasp the strength of his putt and the angle of the ball to grasp the standard of his putt. The putter data are collected through the serial port in the three-dimensional putting scene, mainly including putting speed, deflection angle, and so on.

Through the laser transmitter between the putter and the horizontal laser receiver tube, the receiver tube receives the data sent to the PC through the serial port, collects and processes the data and sends it to the golf instructor software for processing to calculate the ball index. The structure design of this system is shown in Figure 1. If the single-subject experimental method can be used reasonably, good research results can still be obtained after adaptive changes.

All the data collected by the laser receiver tube are transmitted to the PC through the serial port, and the software operation and use are realized on the PC [14]. Data acquisition is mainly using laser infrared technology, and there are three groups of laser transmitters and laser receiver tubes on the golf trainer: two groups located at the bottom of the golf trainer in the horizontal direction, and one group located in the vertical direction at an angle of 60° to the horizontal direction. In addition, this system leads to a sound acquisition controller, whose role is to collect the sound of hitting the ball. When the ball is hit, the two groups of lasers in the horizontal direction will collect the two groups of numbers and send them to the golf trainer software through the serial port for processing to calculate the speed and offset angle of the ball, while in the vertical direction, a group of numbers will be collected and sent to the golf trainer software through the serial port for processing to calculate the height of the ball. The data are processed and displayed in the 3D human-machine interface.

The end motion simulation platform is driven by three electric or hydraulic cylinders to simulate the position and attitude provided by the vision system, which can realize the lifting, pitching, tilting, acceleration and deceleration, bumps, and compound motion during the motion. The simulator consists of an end motion simulation platform, three electric cylinders, a base, three hooke hinges connecting the motion platform to the pushrod, and three Hooke hinges connecting the cylinders and the view system. By controlling the expansion and contraction of the three electric cylinders, the movement of the motion platform with three degrees of freedom of lift, pitch, and lateral tilt is realized. Combined with virtual reality technology and a somatosensory algorithm, various motion scenes are simulated to give the rider the feeling of being in the scene.

Unlike tandem robots, the end deformation of parallel robots not only comes from the deformation of its structure and the accumulation of deformation of other parts but also includes the displacement of the rigid body brought by the relative rotation or movement of the passive motion sub-assemblies [15]. The dangerous poses of the mechanism need to be determined, so the positions and poses of the mechanism need to be parameterized, as shown in Table 1.

As there are three different difficult holes on the general golf green, each hole is in a different position, some holes may belong to the green slope, and some holes may be in the depression of the green. When a golfer is not very far from the green, he may choose to hit the golf ball on the green in one shot, which can reduce the total number of golf shots. For the adjustment of golf ball speed, this system needs to simulate the real shot, so the flying speed of the golf ball should be as close to the reality as possible. In the virtual golf simulation system, it is difficult to perceive the elevation information of the green simply from the graphical point of view, so it is difficult to determine the ideal landing point of the golf ball when hitting the golf ball on the green from a long distance [16]. Therefore, this paper uses advanced coloring language technology to color the elevation of the golf course greens, aiming to provide users with sufficient information about the elevation of the green terrain.

The green topography elevation grid is designed to visualize the green elevation when a golf ball is dropped into the green. Each vertex of the green elevation grid has a roller ball, which can continuously roll from high to low according to the height of both vertices to visually represent the height information of the terrain, and the grid itself is also colored
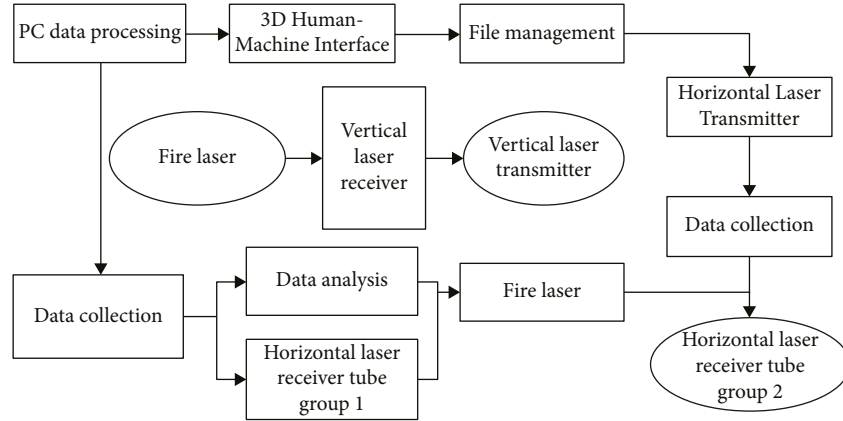
FIGURE 1: Golf trainer system putting training structure design.

TABLE 1: Position acquisition point settings.

X	Y	Z
0	0	0
1	2	−23
2	21	15
3	23	−17
4	5	−21
5	23	0
6	−20	9
7	6	−17

with red, green, and blue gradient according to the vertex height information.

In the virtual golf simulation system, the projection display of the 3D scene is a very important part. At present, the scene display mode of the virtual golf simulation system has three ways: single-screen display, three-screen display, and ring screen display. The scene display of the distributed virtual golf simulation system studied in this paper adopts the three-screen display mode, i.e., the rendered results are projected by three projectors onto three screens with different angles for the virtual scene display. The display system based on three projectors can give the user a 120° viewing angle, which largely enhances the immersion of the virtual reality environment.

The traditional multi-screen display technology adopts the C/S architecture, as shown in the figure below, i.e., one host is used as the server, the rest of the hosts are used as the clients to control one projector each, and the server sends information such as the scene perspective, upward vector, and projection matrix to each client to be displayed, and the client renders the scene output according to the received information and projects it through the projector. In this way, a wide-angle scene is formed on the screen, as shown in Figure 2.

However, for the display of three-dimensional scenes, the simple use of hardware to expand the scene screen is not enough to meet the requirements. Because the camera in the three-dimensional view system is mostly used in perspective projection, if the hardware is used to expand the scene screen to increase the width of the view, it will make both sides of

the scene screen appears a large degree of stretch. Therefore, the cone projection matrix and viewpoint direction need to be calculated for different projection planes.

Through the observation of the actual flying speed of the golf ball, we have calculated a series of speed formulas and calculated different speed formulas at different stages. In virtual reality systems, to improve the realism of virtual scenes, it is often necessary to construct various complex 3D scene models, which may contain more detailed information [17]. The more detailed information the model contains in the virtual scene, the more polygons the model needs, thus making the computer draw many polygons per frame, which has a great impact on the real-time scene drawing. Therefore, how to achieve real-time drawing of large-scale scenes is an important research topic.

However, in many cases, for the same hardware level, the drawing efficiency of computer graphics seems to be inversely proportional to the picture quality of the final drawn scene. When pursuing graphics drawing efficiency, it may be necessary to reduce the complexity of the scene to be drawn, i.e., to reduce the detailed information of the scene to be expressed, which inevitably has a significant impact on the scene's picture quality.

4. Putting Simulator Motion Performance and User Experience Relationship Analysis

Golf putting requires a high standard of action, and whether the putting action is standard or not is directly related to the distance and effect of the stroke [18]. For the training of the standard of the putting action, we mainly analyze the angle of the ball to determine whether the putting action is standard. The next stroke is then modified according to the trajectory of the golf ball.

It was found that prior experience in motor skills can be of great help in motor skill operation. Therefore, this paper uses advanced coloring language technology to colorize the elevation of the greens of golf courses, aiming to provide users with enough information about the elevation of the greens. The experimental study found that the expert group of golfers and the novice group were more stable in putting at different distances, and the expert group was more
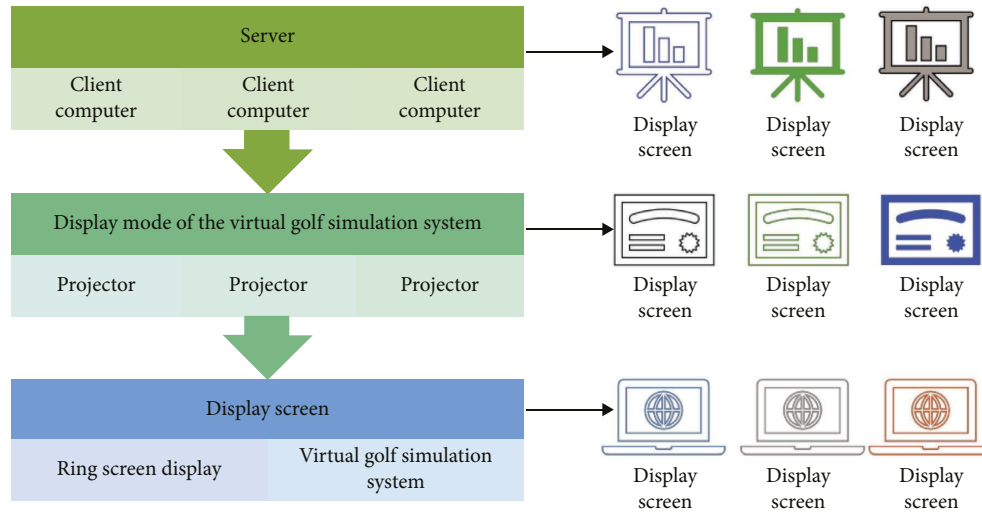
FIGURE 2: C/S mode three-screen display method.

accurate in putting with lower putter impact. There were also differences between the two groups when putting at different distances with differently weighted putters. The expert group used their existing experience to make faster and better judgments on the amplitude of the putting action, the ball position, and the speed of the putt, and the expert group had a more obvious advantage and a higher accuracy rate. The results of the experiment indicate that the golf expert team players can use their prior experience to make adaptive adjustments to different ball positions, thus maintaining good athletic performance and higher putting accuracy.

A-B design is the most basic experimental design for single-subject experiments, and the A-B form of single-subject experimental design has evolved to increase the validity of experimental effects and to adapt to different research needs. The reason for choosing the multi baseline level design is that the study believes that the multi baseline level tester in the single subject experiment can fully control the degree of independent variables, dependent variables, and irrelevant variables, and apply it to this study. In the A-B-A multiple baseline level design, in which the experiment is withdrawn in the study phase, no variables are used in the follow-up phase, and the behavior of the subjects and the collected data are observed to ensure the validity of the intervention experiment to eliminate irrelevant variables, and the data collection in the A2 follow-up phase is conducted and analyzed at the end of the intervention experiment to prove that the intervention experiment can affect the subjects to achieve the intervention effect [19].

At the end of the intervention, a follow-up test was conducted on the junior golfers, and the test was repeated using the Five-Factor Positive Thinking Inventory and the Brief Mood Inventory, and the data were collected seven times during the follow-up phase 2. After the data were obtained, we observed and analyzed whether the level of positive thinking and state of mind remained in a certain range or fluctuated and tested the sustainability of the positive thinking training effect. The specific time allocation of the data is shown in Table 2.

Single-subject experiments have been of great importance in the history of psychology. Single-subject experiments, also known as small-sample experimental designs, have a variety of designs and are flexible and highly operational, so that researchers can make different changes to single-subject experiments to suit different experimental needs and can still achieve good results after adaptive changes if the single-subject experimental method is used wisely. To effectively master the basic skills of this sport and accurately hit the ball to achieve ideal sports results, practitioners need to have good physical strength and effective movement control ability.

To ensure the objectivity and accuracy of the data collected from the experimental tests, statistical software such as Excel was used in this study to test whether there were significant differences between the data before and after the experiments to confirm the feasibility of the experimental protocol [20]. The data collected from the experiments were categorized and organized to produce views and tables of the relevant studies, which allowed for intuitive and clear analysis of the findings and visual interpretation of the data. Data visualization is often used in single-subject designs, which are characterized by the ability to make the most of experimental data and to analyze the effects of intervention data using visual graphics, which not only are easy to use but also make the results of the experiment clear briefly [21].

Any motor skill, although discussed in the context of sport, is, in general, a reaction to the habits of the individual. In golf putting instruction, it is also common to find that many golfers fail to devote all their energy to practicing what the instructor is teaching them, or their minds wander, or they chat about something else. By sorting out value-oriented habit styles, it helps individuals establish a motivational pattern of following a plan rather than an avoidance pattern. Make the event of the moment the goal of the action and focus on the task without distractions, as shown in Figure 3. To play golf well requires several good qualities such as strength, explosive power, endurance, flexibility, balance, and stability [22]. For a professional golfer, any one

Table 2: Time allocation for data collection.

Schedule	Frequency one	Frequency two	Time
Baseline period	7	9	6.14–7.13
Intervention expectations	8	7	7.14–8.13
Tracking period	9	8	8.14–9.13

of them is in a weak point, which will lead to a decrease in the quality of his or her swing and unstable shots, causing the performance to be negatively affected. Golf is different from other sports in that the most important thing is the stability and control of the golfer.

In the whole process of the golf swing, to ensure that the energy is not lost, the golfer needs to polish and improve the details of each technical action of the swing and the overall body movement chain is not able to have a large loss. The control and performance of technical movements are obviously improved, thereby enhancing the stability of golf putting technology, improving the accuracy of hitting the ball, and improving the technical level of students. In particular, the torso is the center of the whole body and the intersection of all the forces of the body, which plays an important role in the overall power transmission and integration [23]. If the stability of the torso is not sufficient, it will hinder the transmission of energy and cannot effectively control the force of all parts of the body, resulting in deformation of technical movements and deviation of the ball, which will hurt the sports performance.

5. Analysis of Results

5.1. Analysis of the Performance of the Virtual Simulation System. After each function of the system is realized, the system needs to be debugged to test whether each interface can be unified and whether each functional module meets the system requirements. System debugging is mainly divided into hardware debugging and software function debugging.

The software must have the corresponding hardware support, which is also a major highlight of the system design. Because the current domestic golf trainer is either only software or only hardware, rarely can the hardware and software be combined to develop. It can be difficult to determine where the golf ball will ideally land when hitting the green from a distance. The debugging of the swing training hardware is mainly the debugging of the laser transmitter and the laser receiver tube. Only the laser emitted from the laser transmitter hits the laser receiver tube to trigger the first, second, and third boards. For the first, second, and third boards, it must be triggered to collect data and realize the function of the data acquisition module. Whether the debugging is successful or not is reflected through the software interface.

The debugging of the swing training module in the driving range scene mainly includes debugging the received data and debugging the golf ball speed after the swing. The debugging of the received data is mainly through the serial debugging assistant to receive data and compare it with the received data of the system, if the result
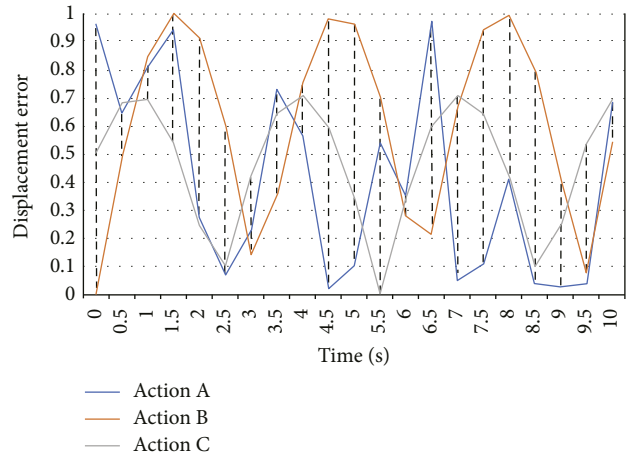
Figure 3: Control drive joint displacement error curve.

is consistent, it means that the received data are correct. For the debugging of golf ball speed, the system should simulate the real ball, so the speed of golf ball flight should be as close as possible to the actual. Through the observation of the actual ball flight speed, we counted a series of speed formulas, and different speed formulas were counted at different stages. After repeatedly debugging and modifying the parameters, the speed of the simulated high ball is made to match the actual high ball flight trajectory as much as possible. The same is true for the debugging of the putter training in the putting scenario. We connect the PC, monitor, and hardware for the overall debugging, as shown in Figure 4.

From the performance data shown in Figure 4, when the 1264 trees in the golf scene are not rendered, the frame rate of the scene is 35.23 fps, the number of private Vertices is (4431069), and the number of Primitives is (1477023). When the scene is rendered using the Instance technique, the data table shows that the number of private vertices and primitives in the scene is 218,412 and 72,804; however, the number of vertex instances and primitives is 0.0%. However, the number of vertex instances is 4431069 and the number of tuples is 1477023, the sharing rate of vertex instances is 95.07%, and tuple instances are 95.07%. This shows that Instance technology has a significant effect on the rendering performance improvement of the scene.

By observing the speed and distance of the golf ball in the practice scene, we can train the strength of our swing; by observing the trajectory and offset of the golf ball, we can train the standard of our swing. If the swing is not standard, the trajectory of the golf ball will be incorrect. We can switch the viewpoint in the scene to observe more intuitively, and there are three viewpoints set in the system. We can switch between them at will so that we can better observe the indicators of our swing, as shown in Table 3. In conclusion, Golf trainer has an important role in swing training. Through field debugging and a lot of experiments, this function meets the system requirements.

In the development process of the system 3D scene, there are still some problems in the articulation between various functional things, especially the articulation of
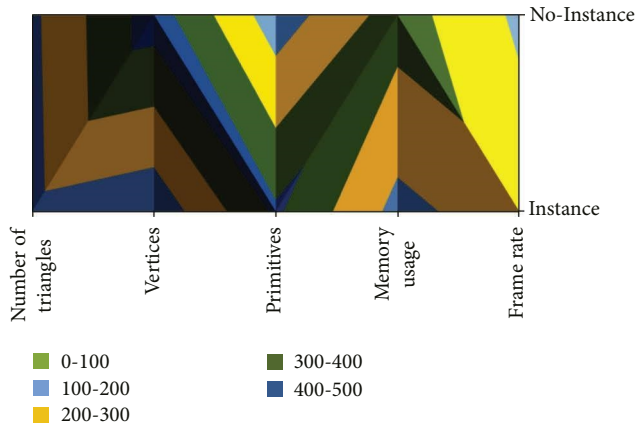
Figure 4: Rendering performance data.

Table 3: The difference test of pre-test and post-test.

Name	Pre-test		Post-test	
	M	SD	M	SD
Acceptance of action	88.11	83.78	77.36	67.47
Fluency	82.46	85.45	77.86	90.8
Cognitive anxiety	61.46	60.65	99.46	65.89
Somatic anxiety	95.38	99.25	98.09	95.25
State of confidence	64.07	76.95	75.46	60.44
Putt hole (big hole)	61.24	89.08	96.08	82.12
Putt hole (small hole)	94.12	82.09	78.8	98.81

the sky box still has defects. In the moment of hitting the ball, the scene sometimes flickers a bit. When the putter scene receives data to hit the ball, the putter scene will flicker a bit, which should be a problem when switching between data processing and scene rendering direct messages. Use three-dimensional technology to display the motion of the golf ball and the trajectory of the golf ball and calculate the distance of the golf ball. This is to be solved in future debugging. In the network online function module, the system can only support four players to join. This feature needs to be optimized and developed to allow a random number of players to train online. The implementation of this feature is subject to further development.

5.2. User Experience Relationship Results. When individuals face critical putting tasks, there is often an element of worry or fear at the root of anxiety in the mind, and training on breath awareness can help individuals identify feelings of safety and calmness. Second, the breath comes and goes freely, and attention to it can be achieved without undue distraction. That is, being aware of the breath allows one to observe one's thoughts more clearly when needed and to return one's attention calmly to the present moment, even during the briefest of breath meditations. Finally, the breath can be long or short, and can be adjusted whenever necessary, so that one can be aware of the state at this moment and obtain his own consciousness at this moment.

Clarity training focuses on improving the sharpness and vividness and realism of representations. It must be practiced by using all sensory experiences and performing the representation rehearsal as vividly and realistically as possible. It includes two elements: action recall training and scene recall training. For the action recall training, the putter is practiced before the exercise. Then, the subject is guided to close his eyes and imagine as much as possible the details of the movements of different parts of his body when he hits the ball. The tableau starts by seeing his ready position, the player looks at the hitting line with his eyes, sees and feels the action of his backswing to the apex, feels his body tense on his left and right legs, tilts his upper body

slightly, swings the club like a pendulum, and sees the ball launched, far and straight, just landing in the middle of the fairway.

Scene recall training allows the subject to feel the change of mental state and feelings during the examination and competition situation. It includes imagining the game scene, standing on the tee, seeing the surrounding game scene, seeing the shape and color of the green, seeing the length and width of the course, the flatness of the turf, feeling the flow of the breeze and the sunshine, and making all kinds of feelings into their imagination; the more accurate and detailed the imagination, the better, as shown in Figure 5.

The positive thinking group scored significantly lower than the representational group in the acceptance action factor, indicating a significant reduction in behavioral avoidance, according to the statistical effect measure showing $r = 0.15$, and in this experiment, about 15% of the overall variance came from the larger effect of the experimental treatment.

The between-group difference in the post-test of positive thinking state between the phenotype group and the positive thinking group was 0.01, which was significantly less than 0.05, indicating that there was a significant difference in the post-test scores of positive thinking state between the two groups after removing the effect of the pre-test of positive thinking state. To improve the user experience of the virtual golf simulation system, the system needs to add some additional functions, such as virtual three-dimensional sound effects, multi-screen display, and other functions. The post-test scores of the total score of positive thinking in the phenom group were significantly lower than the post-test scores of the total score of positive thinking in the positive thinking group. According to the statistical effect measure showing $r = 0.40$, about 40% of the overall variance in this experiment came from the experimental treatment.

Traditional physical training is mainly for whole-body muscle development and has relatively little effect on core strength intervention. Core strength training is an effective training system for core muscle groups such as lumbar and abdominal muscle groups, back muscle groups, hip muscle groups, and so on. These muscle groups all belong to the mid-axis region of the body, and their main function is to maintain spinal curvature and stability, so that the generation, transmission, and control of force can be optimized. Therefore, the improvement of students' stability after core strength training is more obvious, which proves that core strength training is more targeted and effective for core
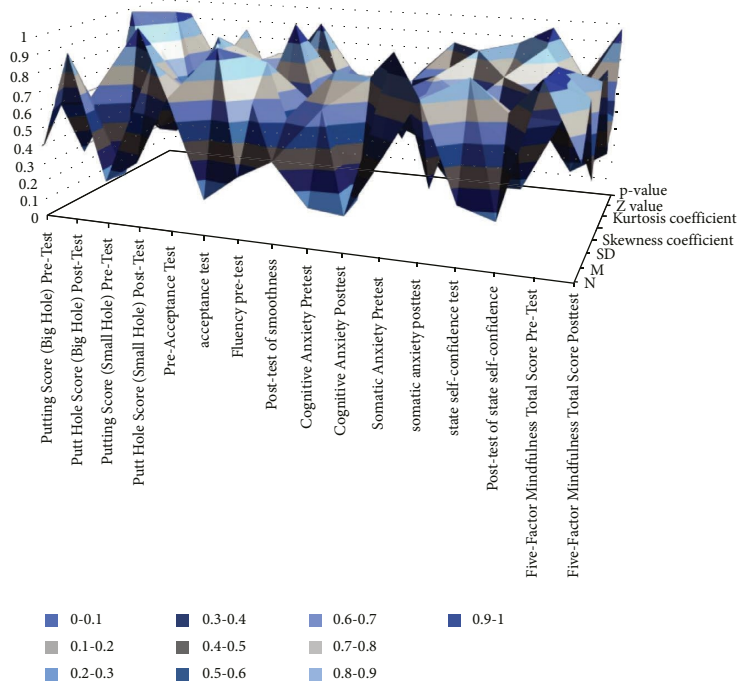
Figure 5: Normal distribution test chart of pre-test and post-test scores.

stability than traditional physical training, as shown in Table 4.

As shown in Table 4, after 18 weeks of training and golf instruction, both the experimental class and the control class improved their putting performance under the independent sample t-value test. The parameter (mean-± standard deviation) of the big hole in the experimental class after the experiment was 7.75 ± 0.20, and the parameter of the control class was 6.30 ± 0.24, with a significant difference of P value <0.01. The mean value shows that the rate of large hole entry in the experimental class was 70% higher than 60% in the control class. The small hole is smaller in diameter than the big hole, so the accuracy of the putter is more demanding.

In conclusion, the effect of golf core strength training is higher than traditional physical training, putting plays a vital role in the golf game, and although the final game score is determined by the total number of strokes in 18 holes, putting accounts for more than 65% of the total number of strokes, and putting has a great influence on the player's performance in playing golf. Master the standard of your putter by observing the angle of the ball. Therefore, we should pay more attention to the training of putting technique. The golf putting technique is to lead, hit, and send the ball. The action is focused on the shoulder turn of the backswing, the release of the ball after a stable hit, and the body as the central axis of the shoulder turn to power the ball. Core strength training is a professional training aimed at fixing muscle groups. Its main function is to improve the core stability. The purpose is to control the posture and center of gravity of the body stably during the movement, so that the strength of the body and limbs can be transferred to the ends of the limbs with maximum efficiency. At the same

Table 4: Comparison of putter accuracy parameters between the experimental class and the control class after the experiment.

Test items	Experimental class	Control class	T	P
Acceptance of action	83.6	73.63	90.67	71.87
Fluency	64.81	83.02	61.77	77.97
Cognitive anxiety	92.48	91.1	86.84	82.67
Somatic anxiety	69.81	74.58	70.08	64.19
State of confidence	97.95	71.22	80.55	71.97
Putt hole (big hole)	60.01	79.66	79.52	90.77
Putt hole (small hole)	97.16	88.61	92.18	62.14

time, it also has the power generation function. It can transfer the strength and control the center of gravity during the movement, so that the golfer can maintain body balance during high-speed movement.

6. Conclusion

The virtual golf grand simulation system is an application of virtual reality technology in sports training. It uses virtual reality technology to move the golf course indoors so that golf enthusiasts can break through the limitations of golf by space, time, and environment and get a realistic golf experience indoors. The design of the server-side system mainly includes the design of multi-threading, the design of data packets, the design of data protocols, and the distribution and management of interactive simulation data, while the design of the client-side network

communication module mainly includes the design of multi-threading, the coordination and management of local interactive simulation data, and the real-time sending and receiving of interactive simulation data. The functional requirements of the golf trainer software are completed. According to the system requirements, this thesis completes the training of golf swing function and golf putting function. Through the treatment of hitting speed, distance, hitting angle, and trajectory, the golf swing strength training and the standardization of swing action are completed. Through the processing of putt speed, distance, and putt angle, complete the golf putt strength and putt angle training. In the three-dimensional scene of the putter, the putter data are collected through the serial port, including the speed of the putter, the declination angle, and so on. After the overall debugging, analysis, and actual operating results of the system, the requirements of the system functions are met. It improves the level of the athlete's mind while improving the performance of sports behavior. When there is windy weather, athletes do not need to pay too much attention to the changes brought about by the weather, which affects the state of mind and affects sports performance, and athletes should always adjust their attention to the changes in the external environment so as not to judge or react, but to focus on the current putting behavior.

Data Availability

The dataset is available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] D. J. Harris, G. Buckingham, M. R. Wilson et al., "Exploring sensorimotor performance and user experience within a virtual reality golf putting simulator," *Virtual Reality*, vol. 25, no. 3, pp. 647–654, 2021.

[2] G. Wood, D. J. Wright, D. Harris, A. Pal, Z. C. Franklin, and S. J. Vine, "Testing the construct validity of a soccer-specific virtual reality simulator using novice, academy, and professional soccer players," *Virtual Reality*, vol. 25, no. 1, pp. 43–51, 2021.

[3] A. Akbaş, W. Marszałek, A. Kamieniarz, J. Polechoński, K. J. Slomka, and G. Juras, "Application of virtual reality in competitive athletes-a review," *Journal of Human Kinetics*, vol. 69, no. 1, pp. 5–16, 2019.

[4] M. Jiang, W. Lan, J. Chang et al., "A game prototype for understanding the safety issues of a lifeboat launch," *Virtual Reality*, vol. 22, no. 2, pp. 137–148, 2018.

[5] N. Kalkan, "Metaverse esbgybd," *Ulusal Spor Bilimleri Dergisi*, vol. 5, no. 2, pp. 163–174, 2021.

[6] P. Soltani, P. Figueiredo, and J. P. Vilas-Boas, "Does exergaming drive future physical activity and sport intentions?" *Journal of Health Psychology*, vol. 26, no. 12, pp. 2173–2185, 2021.

[7] J. Fayaz, S. Azar, M. Dabaghi, and F. Zareian, "An efficient algorithm to simulate hazard-targeted site-based synthetic ground motions," *Earthquake Spectra*, vol. 37, no. 2, pp. 876–902, 2021.

[8] J. P. Uhm, H. W. Lee, and J. W. Han, "Creating sense of presence in a virtual reality experience: impact on neurophysiological arousal and attitude towards a winter sport," *Sport Management Review*, vol. 23, no. 4, pp. 588–600, 2020.

[9] M. R. González-Rodríguez, M. C. Díaz-Fernández, and M. Á. Pino-Mejías, "The impact of virtual reality technology on tourists' experience: a textual data analysis," *Soft Computing*, vol. 24, no. 18, pp. 13879–13892, 2020.

[10] C. Choi, T. C. Greenwell, and K. Lee, "Effects of service quality, perceived value, and consumer satisfaction on behavioral intentions in virtual golf," *Journal of Physical Education and Sport*, vol. 18, no. 3, pp. 1459–1468, 2018.

[11] W. Dangxiao, G. Yuan, L. Shiyi, Z. Yuru, X. Weiliang, and X. Jing, "Haptic display for virtual reality: progress and challenges," *Virtual Reality & Intelligent Hardware*, vol. 1, no. 2, pp. 136–162, 2019.

[12] J. Toner and A. Moran, "Exploring the orthogonal relationship between controlled and automated processes in skilled action," *Review of Philosophy and Psychology*, vol. 12, no. 3, pp. 577–593, 2021.

[13] D. Westmattelmann, J. G. Grotenhermen, M. Sprenger, and G. Schewe, "The show must go on - virtualisation of sport events during the COVID-19 pandemic," *European Journal of Information Systems*, vol. 30, no. 2, pp. 119–136, 2021.

[14] M. Takac, J. Collett, R. Conduit, and A. De Foe, "Addressing virtual reality misclassification: a hardware-based qualification matrix for virtual reality technology," *Clinical Psychology & Psychotherapy*, vol. 28, no. 3, pp. 538–556, 2021.

[15] J. J. Liu, J. Newman, and D. J. Lee, "Using artificial intelligence to provide visual feedback for golf swing training," *Electronic Imaging*, vol. 2021, no. 6, pp. 321-1–321-6, 2021.

[16] J. Kwon and T. Chung, "Importance-performance analysis (IPA) of service quality for virtual reality golf center," *International Journal of Marketing Studies*, vol. 10, no. 3, pp. 30–40, 2018.

[17] L. J. Roberts, M. S. Jackson, and I. H. Grundy, "The effects of cognitive interference during the preparation and execution of the golf swing," *International Journal of Sport and Exercise Psychology*, vol. 19, no. 3, pp. 413–428, 2021.

[18] S. K. Reed, "Combining physical, virtual, and mental actions and objects," *Educational Psychology Review*, vol. 30, no. 3, pp. 1091–1113, 2018.

[19] B. Heusler and C. Sutter, "Gaze control and training for highstress situations in law enforcement: a systematic review," *Journal of Police and Criminal Psychology*, vol. 35, no. 4, pp. 401–413, 2020.

[20] M. Chang, D. Büchel, K. Reinecke, T. Lehmann, and J. Baumeister, "Ecological validity in exercise neuroscience research: a systematic investigation," *European Journal of Neuroscience*, vol. 55, no. 2, pp. 487–509, 2022.

[21] E. McNeill, A. J. Toth, A. J. Harrison, and M. J. Campbell, "Cognitive to physical performance: a conceptual model for the role of motor simulation in performance," *International*

Review of Sport and Exercise Psychology, vol. 13, no. 1, pp. 205–230, 2020.

[22] C. Chauvin, F. Said, P. Rauffet, and S. Langlois, "Analyzing the take-over performance in an automated vehicle in terms of cognitive control modes," *Le, Travail Humain*, vol. 83, no. 4, pp. 379–405, 2020.

[23] D. M. Ste-Marie, N. Lelievre, and L. St Germain, "Revisiting the applied model for the use of observation: a review of articles spanning 2011–2018," *Research Quarterly for Exercise & Sport*, vol. 91, no. 4, pp. 594–617, 2020.

Research Article

An Internet of Things (IoT) Based Image Process Screening to Prevent COVID-19 in Public Gatherings

Suhail H. Serbaya ⓘ

Department of Industrial Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Correspondence should be addressed to Suhail H. Serbaya; sserbaya@kau.edu.sa

The global community is now coping with such a significant issue as the Covid-19 virus, public gatherings are experiencing certain restrictions in order to stop the virus from spreading further. The issue takes on a bigger significance during religious pilgrimages such as the Hajj and the Umrah, when tens of thousands, if not hundreds of thousands, of people gather in holy cities to participate in religious rituals. During such a time period, it is quite difficult to single out an infected person from among the big crowd that is there. The current screening approach only includes a single element of identity, which means that there is a possibility that the screening process may fail because there will not be enough identification. The use of thermal imaging provides a higher level of accuracy when compared to more conventional ways of testing for viral infections in the detection of these symptoms in crowded locations. The primary method that is utilised to determine whether or not a person is infected with the virus is an image processing algorithm that is built in MATLAB. The first step in the process of acquiring an image is to divide the video that is being captured into individual frames. Following this step, the frames that have been focussed are processed in a number of ways. The temperature of a person's body may be estimated by taking a thermal image and then using the RGB separation feature on it. In order to categorise and sort the data, the k-means approach was used as part of the segmentation operation. In addition to eliminating the skin frequency, it also gets rid of the background noise, which often has a higher frequency than the skin frequency. The Viola–Jones technique, which may be used to identify the person's breathing rate, can be used to locate the end of a person's nose, specifically the tip of the nose. The Cascaded Adaboost Classifier is an option that may be used to finish the classification process after the operation has been completed. The suggested method has an accuracy rate of 89.23 percent and a simulation period of around 60 seconds, which guarantees the safety of huge groups of people's public health.

1. Introduction

It is necessary to shorten the duration of the blockchain in order to increase the pace at which the pandemic sickness is spread [1]. According to the Globe Health Organization, fever is one of the most prominent symptoms of the coronavirus and has been seen in a variety of places throughout the world to interrupt the cycle of transmission of the virus. Individuals are screened with the use of infrared thermography in a number of contexts, including airports, railway stations, and other public venues [2, 3]. Infrared thermography is one of the most efficient methods available for identifying fever symptoms in passengers. Despite the fact that IRT is a very powerful method, it is not going to be enough in all circumstances. The IRT technique could not

provide accurate results when used to patients who were on antipyretic medication [3]. The results of the IRT-based screening technique indicate that there are problems in the detection process [4]. The findings vary from 35 percent to 66 percent, which indicates that there is a significant range. It is strongly recommended that employees be checked utilising a wide range of symptoms in order to minimise this drawback. As a result of changes in the respiratory system, COVID-19 causes multiple changes in vital signs, the most notable of which are a slowing of the heart rate and a raising of the temperature of the body. A thermal surveillance camera is employed internally to check for the existence of these important signs [5].

Figure 1 demonstrates the process of contactless screening that will be used for the proposed task. During the
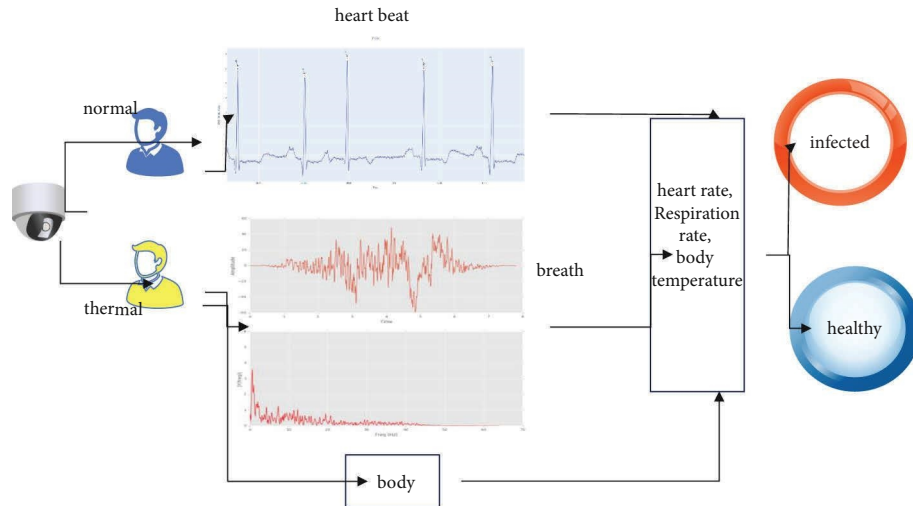
Figure 1: Contactless vital sign measurement systems for symptoms screening.

screening process, the RGB thermal camera observes the variation in the quantity of light that is being emitted by the subject in order to calculate the amount of blood pulse [6]. The light absorption factor in the recommended task, which was assessed using the model derived from the Covid-19 patient, was responsible for the various changes that can be seen in the employee's face, which can be seen in the photo that was taken, which includes the employee's facial region. [7] The inquiry that has been proposed requires that there be no human subjects or participants involved in order to identify a number of symptoms that are associated with the virus. Utilizing the RGB channel of the image input, the work that has been finished mostly concentrated on distinguishing between the blood pulse and the breathing rate [8].

When looking at ways to reduce death rates over the long term, placing a greater emphasis on early therapies may prove to be beneficial. Patients who are in critical condition are obliged to make use of ventilators and to be admitted to an intensive care unit (ICU) as soon as possible [9]. Since COVID-19 patients often arrive to the hospital after the illness is already in a more advanced state, medical professionals are frequently unable to provide an accurate prognosis for these patients at the time of their admission. It is possible that the progression of COVID-19 will occasionally take unexpected turns. One example of this is when the health of a patient who appeared to be in stable condition suddenly deteriorates to a critical stage. This is something that can take even the most experienced physicians by surprise. Artificial Intelligence (AI) models have the potential to be effective tools in improving clinical prediction because they are capable of finding complex patterns in large datasets, which is a feature that the human brain is not capable of achieving on its own. This is one of the reasons why AI models have the potential to be effective tools in improving clinical prediction. In the battle against COVID-19, technologies using artificial intelligence have been applied on a variety of scales, ranging from epidemiological modelling to personalised diagnosis and prognostic

prediction [10]. These applications of artificial intelligence include: Even though a number of different COVID-19 prognostic models have been developed, no exhaustive study has been carried out to analyse and compare the prognostic prediction capacity of noninvasive and invasive factors in the COVID-19 patient group.

In order for the researcher to accomplish these three goals, the first step is to collect routine clinical data from patients on the first day of their admission. The second step is to determine whether noninvasive patient characteristics can predict the COVID-19 mortality outcome. Finally, the third step is to conduct a direct comparison of the mortality prediction powers of noninvasive and invasive patient characteristics [11]. In the first instance of the aforementioned feature groups, noninvasive demographic and clinical factors were separated from invasive laboratory tests in order to facilitate an effective analysis of patient data. This was done because invasive laboratory tests required a greater level of physical access. According to the findings of many studies, the vast majority of COVID-19 patients had their first exacerbation episode between 24 and 48 hours after being admitted to the hospital [12].

A signal extraction method may be constructed from the RGB image that was used as input, which is what was used to accomplish the task of extracting the blood pulse rate. This method was employed. In order to investigate the respiration rate, we are decreasing the SNR value. In order to do this, we are relying on the signal quality index. An Excel sheet is used to compile the employee's medical history, after which it may be uploaded to a more comprehensive government database [13]. During the testing method, which takes place inside the company, there are a total of 20 people taking part: ten healthy volunteers and ten ill ones.

As the pandemic disease has become more widespread, several different approaches have been devised to combat it. Goldberg and colleagues came up with a novel approach to evaluating the levels of blood saturation by making use of a digital camera [14]. In addition to identifying position, this camera can also pick up on a person's temperature. The

image processing technique that was used in the suggested research, which was based on their initial hourglass network, is an essential component [15, 16]. They referred to it as the "Rapid Response Analysis Method," and Poh Ming and his colleagues came up with it. This method was developed to determine the rate of breathing and other vital sign readings while simultaneously recording video with RGB and using changes in the color of the skin. Martinez et al. [17] quantified the movement of the chest by using the red Dot Matrix while conducting a mobility experiment. In this particular instance, infrared imaging will be used so that the human posture estimation may be determined. It is possible to assess the radiation pattern of an object using thermal imaging technology if the person's body temperature is greater than zero degrees Celsius. Xiao, Xuan, and colleagues have developed a camera-equipped smartphone that can determine a person's respiration rate by tracking the movement of their chest and abdomen [18]. When doing this activity, the median number of errors often falls below 2 percent. Basu and colleagues developed a technique to identify changes in the rates of respiration across individuals by using thermal imaging that was based on infrared thermography [19]. He was also responsible for determining the heart rate by applying the method of continuous monitoring of the pulse. According to the results of a number of different survey investigations, it is possible to determine a person's heart rate by observing the blood circulation of that individual. On the other hand, a great number of investigations have shown that filters are applied to remove inaccuracies in the measurement of heart rate. Yan et al. made advantage of this circumstance in order to calculate a person's heart rate using a video of a red, green, and blue facial picture that was lighted by sunlight [13]. To analyse a person's heart rate using monochrome and infrared cameras, Gupta et al. developed a multi-model technique, which they reported in Science [12]. This approach was similar to the one described above. According to [16], the outputs of the experiment have a reasonable accuracy range when the procedure of guiding it is being carried out. The comprehensive research was conducted under natural conditions, and it was found that an infrared sensor could be utilised to monitor the subject's heart rate as well as their breathing rate [20]. Utilizing a dual imaging approach allows for the estimation of the blood oxygen saturation rate. It is not possible to identify the current pandemic illness based only on a single sign measurement [19]. This is due to the reasons discussed above, as well as the fact that multiple thermal screening methods are currently in use. As a result, it is of the utmost importance to devise a screening strategy that incorporates a number of distinct readings for the vital signs. In the event that an employee in the firm contracts the disease, the suggested course of action would be to bring the individual's health information that is stored in the organisation up to date.

A timely and accurate diagnosis of COVID-19 is possible because to effective screening for SARS-CoV-2, which also helps to reduce the load on healthcare delivery systems. The likelihood of infection may now be predicted using models that take into account a number of different factors simultaneously. These are intended to provide assistance to medical personnel all over the globe in the process of triaging patients, which is particularly important given the limited resources available in the healthcare sector. We developed a method based on machine learning that was taught using data from 51,831 people who were assessed (of whom 4769 were confirmed to have COVID-19). The results from the next week were included in the sample set to be examined (47,401 tested individuals of whom 3624 were confirmed to have COVID-19). Our model was able to accurately predict the outcomes of the COVID-19 test by utilising just eight binary characteristics, which were gender, age under 60 years, known contact with an infected person, the presence of five early clinical signs, and known exposure to the virus. Overall, we constructed a model that identifies instances of COVID-19 by using simple indicators that can be accessible by asking fundamental questions. This model was based on the countrywide data that was publicly disclosed by the Israeli Ministry of Health. When there are limited resources available for testing, our methodology may be utilised as one of the factors in the decision-making process to prioritise testing for COVID-19.

2. Materials and Methods

In accordance with the process that had been detailed before, a 10-fold cross-validation was carried out in order to train and tune the model using different examples from the training set [17]. These curves were produced using an algorithm that was not dependent in any way on the technique selecting criteria that were used in the building of the algorithm. When it comes to making accurate predictions, the performance of the noninvasive model seems to be on par with that of the joint model. There were determined to be no statistically significant differences between the two models after a further in-depth statistical investigation [18] discovered that there were. To determine whether or not there were statistically significant variations in the performance of the models and to test the resilience of the data [21] concerning sample sizes, a hold-out cross-validation with a percentage of 10 percent was carried out for a total of 20 iterations. This was carried out. [Cross-validation] At the end of each iteration, ten percent of the data was chosen at random and included into the process of training the model and evaluating the remaining data. The findings demonstrated that the models were insensitive to changes in the sample size, and they also revealed that there was no statistically significant difference between the invasive and noninvasive models [22]. Conclusions: test accuracy ratings for models that had joint, noninvasive, and invasive components were, respectively, 0.80 0.03, 0.77 0.04, and 0.75 0.4, with the joint model having the greatest accuracy out of the three types of models.

Infrared thermal imaging is used in conjunction with an RGB image combination to facilitate the monitoring of a subject's heart rate, respiration rate, and body temperature. Matlab software must be used in order to carry out image processing procedures, and the organization's server must be kept current and up to date with regard to the health-related

database. Both of these requirements must be satisfied. The thermal imaging camera model FLIR C2 is now being evaluated for potential use in the next research project. It has a temperature range of −14 to 302 degrees Fahrenheit and an infrared resolution of 80×60 (4800). In addition, it is able to function in this temperature range. The memory capacity of the camera in question is something in the neighbourhood of five hundred images at a bare minimum. The pixel resolution of the digital camera ranges from 640×480 all the way up to 960×1280, and the pace at which it can capture video is around 15 frames per second (640×480).

2.1. Face Detection and Breathing Rate Calculation.

Because the breathing rate is calculated based on the facial region, it is vital to determine the area around the face. The RGB thermal camera is used to make the determination of the age of the worker. There are flicker cameras on the market that make use of infrared thermography to determine a person's body temperature, as well as their heart rate and breathing rate. The RGB camera was used to record the changes in haemoglobin that occurred in the facial region when blood vessels were subjected to light absorption, and the results of this experiment were acquired by the camera. During inhalation and exhalation, respectively, the flickering camera is used to determine the variation in temperature that occurs in the nasal region. After the image had been processed, the Viola–Jones method was applied to it in order to isolate the facial area from the rest of the picture. K-means segmentation is used in order to complete the task of segmenting the region around the nose. An rise in a person's fever, abnormalities in their breathing rate, and irregularities in their heart rate are all signs that a person is sick, and these symptoms may be used to identify infected people. The research that has been conducted shows that there is a clear divide between those working in the industry who are healthy and those who are afflicted with sickness. Following the implementation of the Viola–Jones technique, the face detection procedure is carried out in the phases that are listed as follows.

Step 1. The attributes of the bigger window size are gathered as a sliding phase in the process, while the smaller window size is gathered as a subwindow.

By moving the sliding window in both the vertical and horizontal orientations in both directions, it is possible to compare the line and edge properties of the photo that was entered by the user.

When the sliding window has reached its maximum position in step three, the corresponding subwindow may be deleted.

Figure 2 depicts the technique that has been discussed.

In Figure 2, the Haar, edge, and line characteristics of the input pictures are represented by the letters F1, F2, and F3. Calculating the respiration rate begins with the collection of data from the input picture via the use of the K-means segmentation technique. It is possible to compute the breath rate by comparing changes in the RGB area between two frames and the nasal region segregated using the K-means

$$F(x) = a1f1\,(x) + a2f2\,(x) + a3f3\,(x)\ \ldots\ldots$$
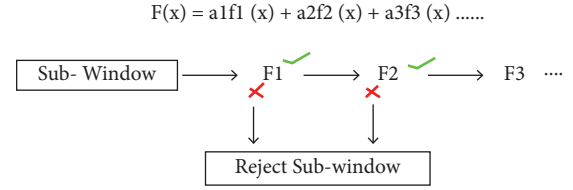
Figure 2: Viola–Jones algorithm.

algorithm. The following is the equation that was used to estimate the rate of breathing. Thermal image segmentation may be used to separate the parts of the nose and mouth that are visible. When the sampled thermal image u is transformed into a two-dimensional magnitude map of thermal-gradient u by the boundary detection technique, the breathing area will be divided into two halves. The breathing region will be divided by the boundary detection method.

$$he\,\Phi\,(x, y) = \sqrt{\left(\frac{\partial u\,(x, y)}{\partial x}\right) + \left(\frac{\partial u\,(x, y)}{\partial y}\right)}, \tag{1}$$

x and y parameters represent the coordinates of the x-y plane in the image matrix. The heat exchange during inhalation and exhalation determines the proposed breathing pattern of the victim. Consecutive frames are selected, and the average value of the spatial thermal distribution calculated for each frame.

$$v(t) = \int_{0}^{Tmin} S(t)dt\ \approx\ \sum_{i}^{t}\sum_{j}^{t} T\delta(t) - u\hat{\imath}j(t), \tag{2}$$

where $v(t)$ determines the thermal value of the nose region in the cross-sectional view whereas the temperature T, $u\hat{\imath}j$ represents the temperature with the absolute value in the tracked area, and $T\delta$ indicates the upper boundary level to the inner volume of the segmented thermal image that is above the temporal moving value ($n = 2$) in the spatial domain. The spatial temperature represents the global thermal changes within the image. If the K-means algorithm loses the nose region's data, $T\delta$ will element the present frame and move on to the next frame with the bounding box region.

2.2. Image Processing Techniques to Detect Vital Signs.

Figure 3 depicts the method used by the Viola–Jones approach to recognize a facial picture in a photograph. It is the Viola–Jones algorithm that is used to recognize faces in images. The letter k indicates that segmentation was performed in order to separate the tip of the nose. In order to retrieve the single feature of the picture, the pixel value of the nose area is computed and used. The respiration rate may be calculated using the pixel attributes of each frame, which can be found in a calculator. A Fast Fourier transform is used to a respiratory signal in order to determine the rate at which the person is breathing. Using the total number of pixels (px) in the segmented picture and the original image (py), the segmentation procedure is evaluated for its proximity to the original image.
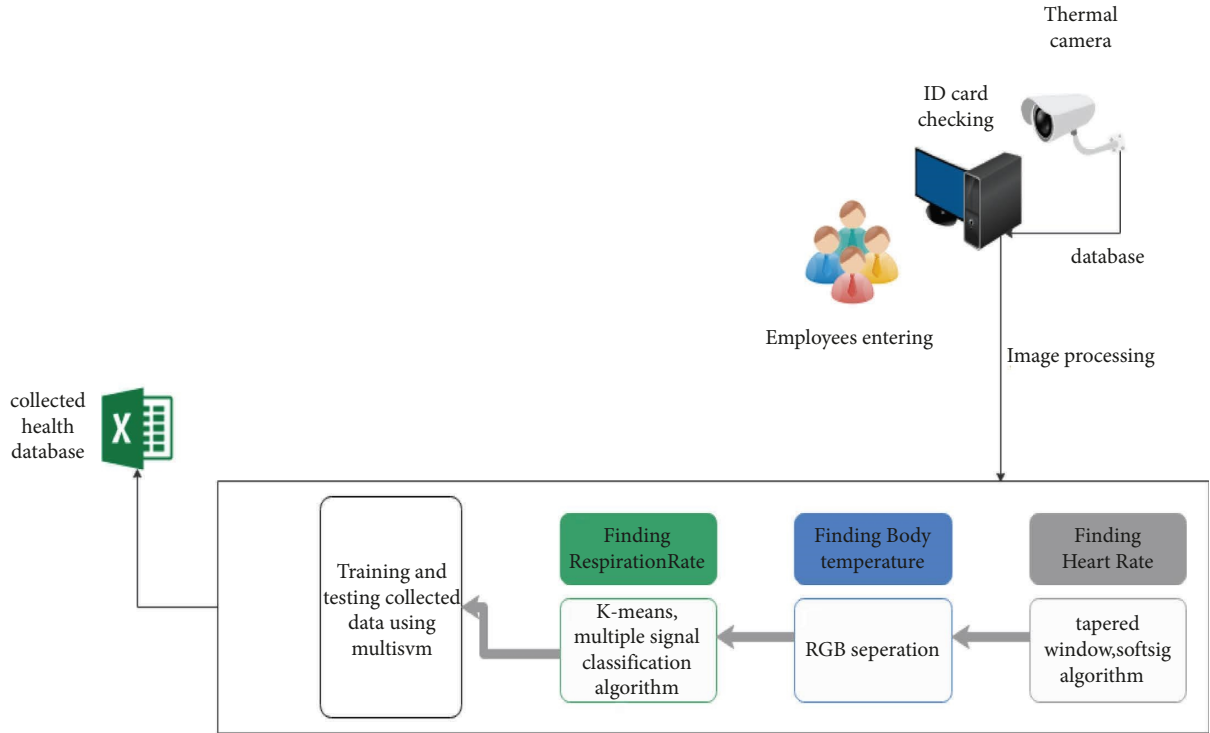
FIGURE 3: Block diagram of the proposed work.

$$\text{Input Image Closed Percentage} = \frac{A1}{A2} * 100. \qquad (3)$$

During the inhaling and exhalation processes, the color of the tip of the nasal area changes from red to blue to green. Changing the RGB region will result in corresponding changes in the pixel area. The Viola–Jones method is used to recognize the face area in a thermal picture captured with a thermal camera.

Figure 4 shows the nose area being segmented using the K-means segmentation method, which is achieved by the fusion of RGB and thermal imaging images. The respiration signals are retrieved from the identified locations with the use of pixels included within the pictures, which are used to extract the signals. The following equation may be used to calculate the mean temperature and the height of the segmented area.

$$x\text{mean}(t) = \frac{1}{mn} \sum_{x=0}^{m=1} \sum_{y=0}^{n=1} I(x, y, t) = I(x, y, t). \qquad (4)$$

The respiration signals are included in the functions $x$mean(t) and $x$min(t). With respect to the segmented image, $I(x, y, t)$ represent the pixel range in terms of temperature with the image coordinates of (x, y) in the segmented image and time t, m indicates the extracted signals, which are: $x$mean nose(t), $x$min nose(t), $x$mean (t), and $x$min nose(t) (t). Respiration signals are selected based on their temperature traces in the three different channels of the input picture, which are split into three groups. The frequency of the input image's respiration is the most important element in determining the rate at which it is breathing.

2.3. RGB Processing for HR Estimation Using Tapered Window.
To detect the Region of Interest facial a general window function is used which is nothing but a Tapered window (Figure 4(b)). The edge area is impacted by the latency caused by the face tracker in the facial Region of Interest, which is visible in the edge region. Instead, the center of the Region of Interest may be used to create steady tracking of the face skin, which is more accurate. To reduce the amount of noise lifted by the face tracking, we obtained a tapered window to a weighted Region of Interest and applied it to the facial tracking data.

$$\% \text{ taper}(i) = \{i = 0, 1, 2, \ldots m - 1 i = n - m, \ldots . n \text{ otherwise}\}. \qquad (5)$$

By using the divergence between absorption of RGB and signal, the purpose of signal reconstruction is to determine the identity of a reconstruction vector $V = r, g$, and b in order to recover the signal from the heartbeat using the absorption of RGB and signal. It is necessary to employ three RGB channels in order to optimize a linear function for enhancing the SNR value while reconstructing a BVP signal. The pulse reflection strength, according to earlier study, is referred to as the link in $G > B > R$, which categorizes the RGB channels based on their intensity levels. With the help of this link, the reconstruction of the signal may be represented as follows:

$$y(t) = V_r X_r(t) + V_g X_g(t) + V_b X_b(t). \qquad (6)$$

Vector reconstruction is represented by the variables vr, vg and vb. Unlike this strategy, which is based on the Softsig method, we enhance the process of identifying the
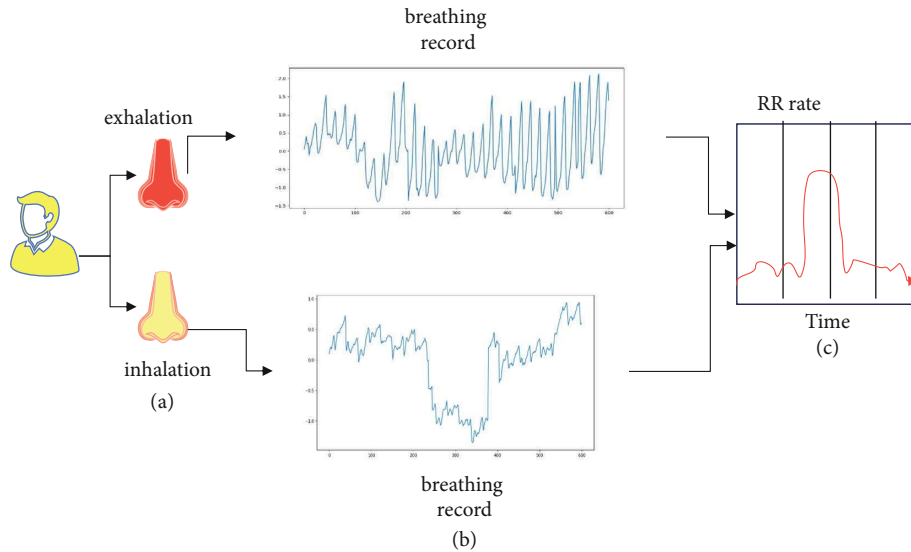
Figure 4: Block diagram of signal processing for respiration rate (RR) estimation. (a) Segmented Region. (b) Time-series extracted from RGB signal. (c) Power spectrum.

procedure for vector V. We picked V in order to increase the kurtosis of the spectra in the Heart Rate range of [0.87–4.97 Hz] in order to get a better signal of the heartbeat. Finally, the MUSIC system was established in order to interpret Heart Rate and Respiration Rate data in a short period of time. A high-resolution frequency estimate of Heart Rate and Respiration Rate is possible with this approach, which allows for a better comprehension of the frequency. The current technique for monitoring respiration using an IRT is based on the temperature change of the nasal passages.

3. Results and Discussion

As previously stated, the primary frequency noise component is included in a raw trace color of RGB and is distinguishable by their spectra, as the pulse oximeter gives 1.83 Hz during the measurement of Heart Rate, which is considered the ground truth. By using the suggested approach, it is possible to detect an evident peak in the component of Heart Rate frequencies shown in Figure 5. The benefits of having a planned Heart Rate evaluation are shown in this example.

For example, as shown in Figure 6, the raw green trace, which utilizes FFT and just one green channel, is compared to the suggested technique for evaluating signal reconstruction and the tapered window, which is shown by the dotted line in Figure 5. If you have an RGB camera and want to monitor the heart rate of a person, you may use the same fundamental method as the green trace. A pulse oximeter and an electrocardiogram are used to assess the heart rate of the patient. The heart rates of 128 pairs of individuals were measured four times 15 seconds apart next to healthy control participants. The subjects comprised 41 patients with flulike symptoms and 22 healthy control subjects in the study, for a total of 128 pairs of subjects.

Figure 5(a) depicts the green trace process of the scatter plot, with the Pearson correlation coefficient (R) at the time of analysis being 0.48 (green trace process). The proposed Bland–Altman strategy, shown in Figure 5(b), makes use of a variety of approaches, including signal reconstruction, Music, and tapered window techniques, among others. The 95 percent limits of agreement vary from −10.7 to 12.9 bpm (standard deviation = 5.89), with a root mean square error of 5.97. The 95 percent limits of agreement range from −10.7 to 12.9 bpm (standard deviation = 5.89). There is a mean of −10.7 between the 95 percent bounds of agreement. This method's scatter plot is shown in Figure 7(d); the Pearson correlation coefficient for this specific technique was 0.91, which indicates that it is very effective. The experimental findings reveal that the proposed strategy has the potential to lower the 95 percent limits of agreement from [23.9, 33.8] bpm to [10.7, 12.9] bpm, which is a statistically significant decrease. As a result of the clinic environment serving as a mimic of a real-world scenario, patients suffering from flulike disease (red circle) fared far better than the general population.

It was in December of the year 2019 that this pandemic crisis started in China, and it has since spread around the world. By the end of March 2020, this virus would have infected more than 110 countries all over the globe, according to estimates. Coughing, high fever, and irregularities in the body's physiology are among of the most visible signs of the dress pandemic illness, which has spread over the world. By using new technologies, the proposed research was able to enhance the multiscreening of the infection's vital signs. It is designed for use mainly in the screening of interpretive signs using a contactless technique, which is the primary application of the proposed technology. During this endeavor, data from both healthy and sick people will be collected and trained. There is a prototype model proposed in this study that is more compatible and accurate than the present model.

(a)

(b)
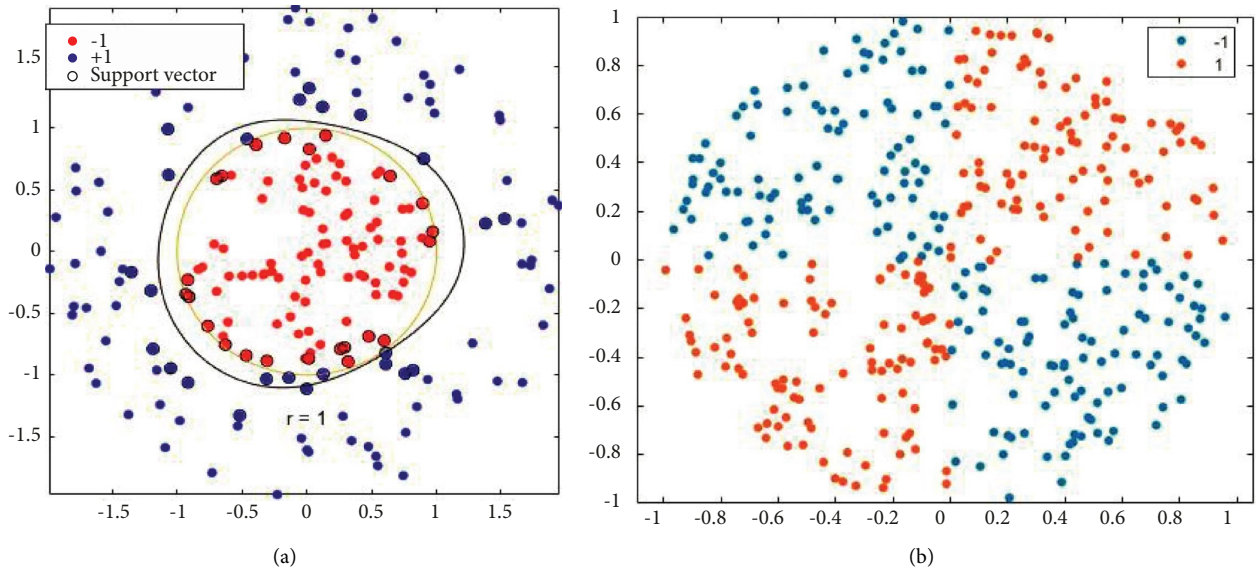
Figure 5: (a) Classifier training region. (b) Classifier tested region.
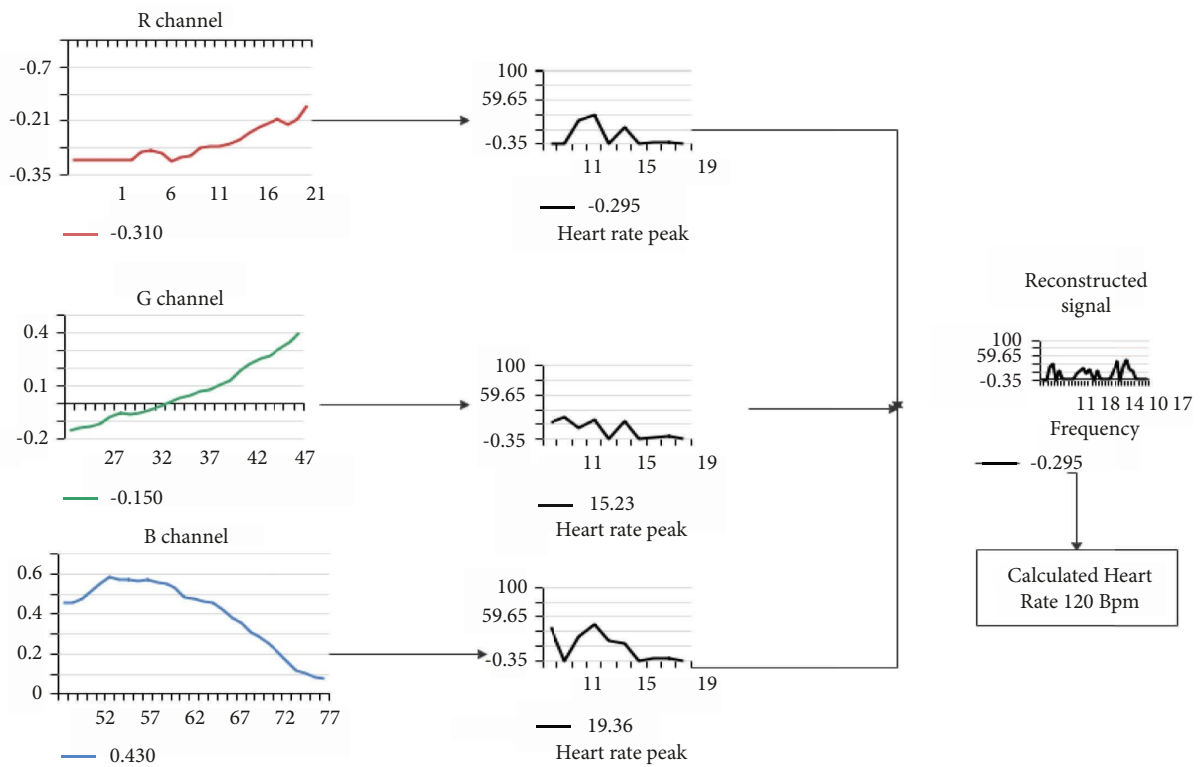
Figure 6: Reconstructed signal.

Techniques such as the tapered window method, k-means segmentation, and support vector machine classification are among those used in the suggested investigation. The Viola–Jones method takes as input a thermal image and extracts a face area from the face region using the face region as input. [10.12, 12.32] beats per minute, [2.123, 3.86] beats per minute, and temperature: [0.687, 2.84] degrees Celsius, the suggested study reveals the similarity between the heart rate of the infected individual and the reference device [23]. Using the suggested study, it will be possible to show the similarity between the heart rate of the infected individual and the reference device. The stability and reliability of our framework in terms of its ability to provide vital signs has been greatly enhanced.

The classifier predicts the output class as an infected or healthy person and writes up in the XL sheet as shown in

Figure 7: Classifier result.

Table 1: Calculated metrics from classifiers.

	TP	TN	FP	FN
Actual_class1	10	0	0	90
Actual_class2	25	0	0	75

Figure 8: Confusion matrix.

Figure 7. The predicted class gives the binary range where 0 as a healthy person and 1 as an infected person.

Table 1 represents the Performance metrics calculated from the predicted class which decides the accuracy range of the classification process.

As shown in Figure 8, the confusion matrix depicts the training accuracy of the gathered characteristics. The planned job will be submitted to a screening process that will include the affected individual inside the business as well as personnel in familiar locations.

4. Conclusion

The global community has been put in the position of dealing with a myriad of problems ever since the COVID-19 made its debut not too long ago. Keeping the propagation of viruses under control in crowded surroundings is one of the challenges that presents one of the greatest opportunities for frustration and failure. By monitoring a variety of measurements of the patients' vital signs, the researcher in this study aimed to identify individuals who were infected. Using real-time measurements, a brand-new system has been constructed, and it has the capacity to evaluate a huge number of essential sizes in real time. Real-time measurements have been used. A wide range of sign measures that are based on thermal imaging need the use of thermal image processing methods, and it is necessary to have these techniques available. The k-means methodology, which was used throughout the whole of the segmentation process, was utilised in order to segment the data. Particularly in this study, techniques of image processing were employed to amplify the minute temperature variations generated by facial-skin blood flow. The heart rate was computed based on the amplified temperature signal that was applied, and the results were presented in this work. The Viola–Jones technique is used in order to identify the tip of the nose, which is then utilised in the subsequent calculation of the individual's breathing rate based on the tip of the nose. At the end of the procedure, the Cascaded Adaboost Classifier is used to complete the classification process that was started earlier in the procedure. The findings are recorded in a database sheet that is maintained on a consistent basis and include information on the health of the staff members. The accuracy rate of categorization reveals that 82.25 percent of the classifications are accurate, while only 81.05 percent of the classifications are inaccurate. This ensures that the virus will be contained in public meetings. The results of the sensitivity and specificity tests show that the goals have been met with an accuracy of 82.25 and 81.05 percent, respectively. According to the most recent findings of research, it may be possible for it to limit the spread of the virus by interfering with the chain-breaking mechanism when the virus is in the reproduction phase. The implementation of artificial intelligence (AI)-based contactless screening has proven to be an efficient method for preventing the spread of viruses in public gatherings and assisting various governments in their efforts to manage crowds. Another goal of this method is to aid in the management of crowds.

Data Availability

The data that support the findings of this study are available from the author on request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] D. S. Hui, T. A. Madani, F. Ntoumi et al., "The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—the latest 2019 novel coronavirus outbreak in Wuhan, China," *International Journal of Infectious Diseases*, vol. 91, pp. 264–266, 2020.

[2] E. Y. Ng, G. Kawb, and W. Chang, "Analysis of IR thermal imager for mass blind fever screening," *Microvascular Research*, vol. 68, no. 2, pp. 104–109, 2004.

[3] M.-F. Chiang, P. W. Lin, L. F. Lin et al., "Mass screening of suspected febrile patients with remote-sensing infrared thermography: alarm temperature and optimal distance," *Journal of the Formosan Medical Association*, vol. 107, no. 12, pp. 937–944, 2008.

[4] M. Condoluci, G. Araniti, T. Mahmoodi, and M. Dohler, "Enabling the IoT machine age with 5G: machine-type multicast services for innovative real-time applications," *IEEE Access*, vol. 4, pp. 5555–5569, 2016.

[5] F. Liu, D. Chen, X. Zhou, W. Dai, and F. Xu, "Let AI perform better next time—a systematic review of medical imaging-based automated diagnosis of COVID-19: 2020–2022," *Applied Sciences*, vol. 12, no. 8, p. 3895, 2022.

[6] P. Jain, W. F. Alsanie, D. O. Gago et al., "A cloud-based machine learning approach to reduce noise in ECG arrhythmias for smart healthcare services," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[7] H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Zaguia, and H. Hamam, "Diabetic retinopathy diagnosis from fundus images using stacked generalization of deep models," *IEEE Access*, vol. 9, Article ID 108276, 2021.

[8] Y. Pathak, P. K. Shukla, A. Tiwari, S. Stalin, and S. Singh, "Deep transfer learning based classification model for COVID-19 disease," *Irbm*, vol. 43, 2020.

[9] A. A. Bakhsh, A. Rizwan, A. B. Khoshaim, E. H. Abualsauod, and G. C. Altamirano, "Implications of COVID-19 on student learning satisfaction (sls): a remedial framework for universities," *International Journal of Engineering Education*, vol. 37, no. 6, pp. 1582–1593, 2021.

[10] E. Costan, G. Gonzales, R. Gonzales et al., "Education 4.0 in developing economies: a systematic literature review of implementation barriers and future research agenda," *Sustainability*, vol. 13, no. 22, Article ID 12763, 2021.

[11] A. Procházka, M. Schatz, O. Vysata, and M. Valis, "Microsoft kinect visual and depth sensors for breathing and heart rate analysis," *Sensors*, vol. 16, no. 7, p. 996, 2016.

[12] O. Gupta, D. McDuff, and R. Raskar, "Real-time physiological measurement and visualization using a synchronized multi-camera system," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Las Vegas, NV, USA, June 2016.

[13] Y. Yan, X. Ma, L. Yao, and J. Ouyang, "Noncontact measurement of heart rate using facial video illuminated under natural light and signal weighted analysis," *Bio-Medical Materials and Engineering*, vol. 26, no. s1, pp. S903–S909, 2015.

[14] S. Goldberg, S. Heitner, F. Mimouni, L. Joseph, R. Bromiker, and E. Picard, "The influence of reducing fever on blood oxygen saturation in children," *European Journal of Pediatrics*, vol. 177, no. 1, pp. 95–99, 2018.

[15] R. Krishnamoorthi, S. Joshi, H. Z. Almarzouki et al., "A novel diabetes healthcare disease prediction framework using machine learning techniques," *Journal of Healthcare Engineering*, vol. 2022, 10 pages, 2022.

[16] G. Sun, T. Matsui, T. Kirimoto, and Y. Yao, "Applications of infrared thermography for noncontact and noninvasive mass screening of febrile international travelers at airport quarantine stations," in *Application of Infrared to Biomedical Sciences*Springer, Berlin, Germany, 2017.

[17] A. Martinez-Möller, D. Zikic, R. M. Botnar et al., "Dual cardiac–respiratory gated PET: implementation and results from a feasibility study," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 34, no. 9, pp. 1447–1454, 2007.

[18] X. Xiao, Z.-C. Wu, and K.-C. Chou, "A multi-label classifier for predicting the subcellular localization of gram-negative bacterial proteins with both single and multiple sites," *PLoS One*, vol. 6, no. 6, Article ID e20592, 2011.

[19] A. Basu, A. Routray, and R. Mukherjee, "Infrared imaging based hyperventilation monitoring through respiration rate estimation," *Infrared Physics & Technology*, vol. 77, pp. 382–390, 2016.

[20] H. Z. Almarzouki, H. Alsulami, A. Rizwan, M. S. Basingab, H. Bukhari, and M. Shabaz, "An internet of medical things-based model for real-time monitoring and averting stroke sensors," *Journal of Healthcare Engineering*, vol. 20219 pages, 2021.

[21] A. Motwani, P. K. Shukla, and M. Pawar, "Novel framework based on deep learning and cloud analytics for smart patient monitoring and recommendation (SPMR)," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–16, 2021.

[22] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Optics Express*, vol. 18, no. 10, Article ID 10762, 2010.

[23] A. O. Al-Youbi, A. Al-Hayani, A. Rizwan, and H. Choudhry, "Implications of COVID-19 on the labor market of Saudi Arabia: the role of universities for a sustainable workforce," *Sustainability*, vol. 12, no. 17, p. 7090, 2020.

Research Article

Computer Vision and the IoT-Based Intelligent Road Lane Detection System

R. Shashidhar [ID],[1] **B. N. Arunakumari** [ID],[2] **A. S. Manjunath,**[3] **Neelu Jyoti Ahuja,**[4]
Vinh Truong Hoang [ID],[5] **Kiet Tran-Trung** [ID],[5] **and Assaye Belay** [ID][6]

[1]*Department of Electronics and Communication Engineering, JSS Science and Technology University, Mysuru 570006, India*
[2]*Department of Computer Science and Engineering, BMS Institute of Technology and Management, Bengaluru 560064, India*
[3]*Department of Computer Applications, JSS Science and Technology University, Mysuru 570006, India*
[4]*Department of Systemics, School of Computer Science, University of Petroleum & Energy Studies, Dehradun, India*
[5]*Faculty of Computer Science, Ho Chi Minh City Open University, Vietnam, 97 Vo Van Tan, Ward Vo Thi Sau, District 3,
Ho Chi Minh City 70000, Vietnam*
[6]*Department of Statistics, Mizan-Tepi University, Tepi, Ethiopia*

Correspondence should be addressed to Assaye Belay; abstat23@gmail.com

Many technical improvements have recently been made in the field of road safety, as accidents have been increasing at an alarming rate, and one of the major causes of such accidents is a driver's lack of attention. To lower the incidence of accidents and keep safe, technological innovations should be made. One way to accomplish this is with IoT-based lane detection systems, which function by recognizing the lane borders on the road and then prompting the turning of the road. Because of the various road conditions that one can encounter when driving, lane detection is a difficult problem. An image processing-based method for lane detection has been proposed in this paper. In this regard, each frame from the video is extracted and image processing techniques are applied for the detection of lanes. The frame which is extracted from the video is then subjected to a Gaussian filter for the removal of noise. Subsequently, color masking has been used to process the frame to detect only the road lanes, whose edges are obtained by applying the canny edge detection algorithm. Afterward, the Hough transform has been applied to the region of interest to extrapolate the lines. Finally, the path is plotted along the lines, and turns are predicted by using the concept of vanishing points.

1. Introduction

In today's world, numerous technical advancements are coming in the field of the automobile industry, and everything is getting automated. One of the examples is self-driving cars (cars that can navigate without drivers). One of the ways to achieve the same is to interface the IoT-based road lane detection system with the cars, which works by recognizing the lane borders or boundaries on the road and helps the car to navigate and guides it to take turns.

Many systems, such as lane departure warning, adaptive cruise control, lane change aid, turn to assist, time to lane change, and fully autonomous driving systems, have been actively researching IoT and vision-based road lane recognition. Although specialized systems for recognizing specific road types have made great progress, little work has been achieved in presenting a general method to detect a range of road conditions. In straight and curved, white and yellow, single and double, solid and broken, and pavement or highway lane limits, an effective lane identification system will navigate autonomously or aid drivers. Intelligent vehicles and smart infrastructure work together in intelligent transportation systems to create a safer environment and better traffic conditions. Lane detection is a critical component of intelligent vehicle systems. As a result, we want to demonstrate a reliable road lane marker detecting system.

However, a more compelling incentive to develop intelligent vehicles is to increase safety by automating or

partially automating driving responsibilities. Among these functions, road detection is critical in driving assistance systems that offer data on the lane structure and the vehicle position relative to the lane. The most compelling reason to equip vehicles with autonomous capabilities is to meet safety requirements. As a result, a system that alerts the driver about the danger has the potential to save a significant number of lives. Computer vision, which has become a powerful tool for perceiving the surroundings and has been widely used in many applications by intelligent transportation systems, is one of the primary technologies involved in these (ITS). Lane detection is often described as the localization of certain primitives such as road markings on the surface of painted roads in several proposed systems. This restriction simplifies the detection process; nonetheless, two conditions can obstruct it: the presence of other vehicles in the same lane that partially obscure the road markings ahead of the vehicle as well as the existence of shadows cast by trees, buildings, and other structures. This study proposes a vision-based method for detecting and tracking structured road boundaries with a slight curvature in real-time even in the presence of shadows.

Lane borders can be identified in real-time by using a camera fixed on the front of the vehicle, which captures the view of the road. In this study, we will use a prerecorded video of the road as an input, extract all of the frames as a set of images, and then extract the characteristics from each of those images, which will be used to detect lanes and predict turns.

This system will not depend much on external factors. It only needs a clear picture of the road. This system is not justmeant for self-driving cars; we can even implement it in normal cars which will give a warning to the driver when he crosses the line on a highway because of drowsiness or some other reasons. It will also help the driver to automatically turn on the indicator if he wishes to cross the line. The major objective of the proposed work is to extract frames from a video and apply image processing techniques to detect the lanes. Join all the lanes and form two lines at the boundaries of the path. Plot the path between the lanes. Predict the turn-based on the length of the boundary lines. Display the turnover of the image. Display the possibility of overtaking based on the type of road lane. Contributions of the proposed work is to develop and implement a road lane detection algorithm using the basic image processing techniques and plot a path between the road lanes in which an autonomous vehicle has to move by predicting the turn direction and also predict the possibility of overtaking based on the type of road lane.

The major contributions of the proposed work are as follows: (i) Conventional image processing techniques such as edge detection, Hough transform, and line extrapolation to detect the road lanes are very accurate and easily implementable than machine learning methods and also comparatively faster; (ii) we proved an IoT and vision-based approach capable of reaching a real-time performance in the detection and tracking of structured road boundaries with a slight curvature, which is robust enough in the presence of shadow conditions; and (iii) proved machine learning as high error-susceptible than in our work.

The remaining sections are as follows: Section 2 described the literature survey of the exciting work; Section 3 describes the proposed methodology in detail; Section 4 mentioned the result and comparison table; and Section 5 concludes the proposed work with future scope.

2. Related Work

Vehicle wrecks continue to be the largest source of death and injury around the globe, costing the lives of tens of thousands of people each year and hurting millions more. The majority of these transportation-related deaths and injuries happen on America's roadways. As autonomous cars are ready to fill up the roads, the chances of the crash rate will increase more, which is a most concerning issue. This paper proposes a vision-based solution to real-time recognition and tracking of structured road boundaries with small curvature that is strong enough in the face of shadow situations.

In complicated driving settings, a robust lane recognition model based on vertical spatial features and contextual driving information is proposed [1–4]. Two developed blocks such as the feature merging block [5] and the information exchange block illustrate a more effective utilization of contextual information and vertical spatial features, allowing detection of more unclear and occluded lane lines. Hough transform applied to detect lane markers after picture noise filtering devised utilizing an image processing computer language platform is another efficient way [6]. Using a pair of hyperbolas that are fit to the boundaries of the lane and extracted using the Hough transform, a vision-based lane identification strategy capable of approaching real-time operation with robustness to illumination change and shadows [7] is also described. Two approaches to recognizing automobiles on the road are appearance-based and feature-based [8–10]. Optical flow, backdrop subtraction, and frame subtraction are the most often utilized methods for vehicle detection [11]. Background subtraction is the most well-known and widely utilized approach today [9]. Furthermore, tail light pairs can be used for night-time vehicle detection [11]. Other methods for recognizing automobiles on the road include texture descriptors [10], Haar filters [12], and others. A deep architecture [13] that can run in real-time while giving accurate semantic segmentation also exists. Our architecture is built around a new layer that employs residual connections and factorized convolutions. In addition, a multitask deep convolutional network [14] recognizes both the presence of the target and its geometric features with the region of interest. For structured visual identification, a recurrent neural layer [15] is used to deal with the spatial distribution of observable cues relating to an object whose shape or structure is difficult to characterize directly. Machine learning is not flawless despite its many benefits and widespread appeal. Data acquisition, time and resources, and result interpretation are some of the factors, which make

machine learning high error-susceptible. In the modern era, IoT is proposed for the traffic system. Several sensors are integrated using a microcontroller for a smart traffic system in an emergency [16–20].

To solve this problem, we intend to use predictable image processing techniques such as edge detection, Hough transforms, and line extrapolation to detect the road lanes, which is very accurate and easily implementable than other methods, and also comparatively faster. The most important point is that the above-mentioned techniques only detect the lanes but the latter also predicts the turns by extrapolating the lanes.

3. Material and Methods

3.1. Dataset. The source of image frames given to the system is data collected from the camera located between the front windscreen and the rear-view mirror. The taken image frames can be separated into foreground and background fields when the camera lens is parallel to the ground. By selecting an adequate ROI, you may reduce not just the search area in photos but also the interference from un-wanted items. We have so far relied on datasets that were freely available on the Internet. https://www.kaggle.com/soumya044/advanced-lane-detection.

3.2. Methods. The four main core methods used for this work are as follows: First is color masking; second is edge detection; the third is region of interest selection; and the fourth method is Hough line transform extrapolation.

Different methods modalities are found useful for the task of road lane detection, each having its pros and cons. This section elucidates a simple yet effective strategy for doing so, explains the major characteristics and usual applications, and breaks down the road lane detection task into functional modules, enumerating the approach to each module's implementation.

3.2.1. Proposed Work. To implement the proposed method, MATLAB's Image Processing Tool has been used, and the block diagram of is shown in Figure 1. To detect the lanes and extract the features of the lanes in every frame of the video, we use color masking, edge detection, ROC selection, Hough transform selection, and other techniques.

The frame read from the video is subjected to a Gaussian filter to remove the noise. Then it is masked with yellow and white colors to detect the road lanes properly. Then the edges of the lanes are obtained from the masked images by applying the canny edge detection technique. Then the region of interest, i.e., road lanes will be selected and then the Hough transform will be applied to extrapolate the lines. Finally, the path will be plotted along the lines and turns will be predicted by observing the vanishing points from the extrapolated lines.

If the right vanishing point is greater than the left vanishing point, then the system will predict that there is a right turn ahead and vice versa. The source of image frames given to the system is data collected from the
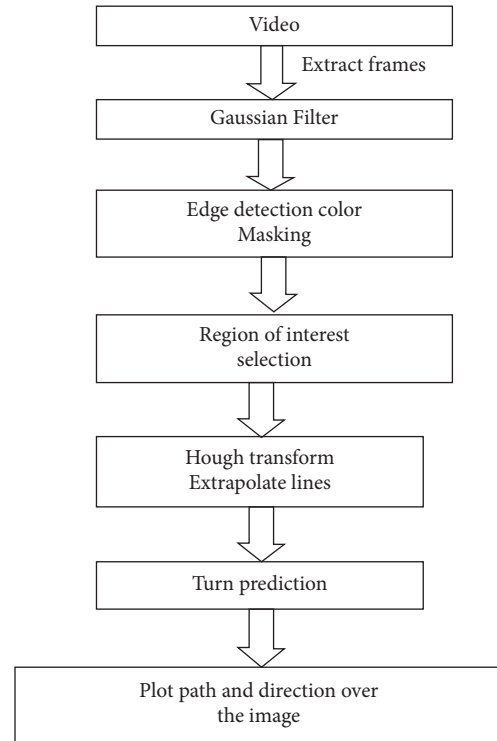
FIGURE 1: Block diagram.

camera located between the front windscreen and the rear-view mirror. The taken image frames can be separated into foreground and background fields when the camera lens is parallel to the ground. By selecting an adequate ROI, you may reduce not just the search area in photos but also the interference from unwanted items. As of now, we have used the datasets which were available online. We have surfed and searched for the dataset videos which would perfectly match the appropriate application which we are thinking of.

The four main core concepts that we intend to use in this method are color masking, edge detection, region of interest (ROI) selection, and Hough line transform extrapolation.

3.2.2. Gaussian Filter. A linear filter is a Gaussian filter. It is frequently used to blur or minimize the noise in an image. For "unsharp masking," two of them are employed and subtracted (edge detection). The Gaussian filter blurs edges and reduces contrast by itself. Figure 2shows the frame of the image after the application of the Gaussian filter to Figure 3.

The Gaussian filter function is defined as

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-x^2+y^2/2\sigma^2}. \qquad (1)$$

The amount of filtering is inversely proportional to the sigma (variance) value. More frequencies are muted when sigma is less, and vice versa. The coordinates of the pixel are denoted by x and y.
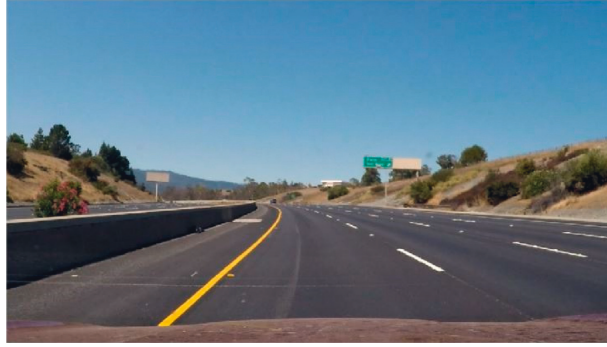
FIGURE 2: Picture frame after applying the Gaussian filter.
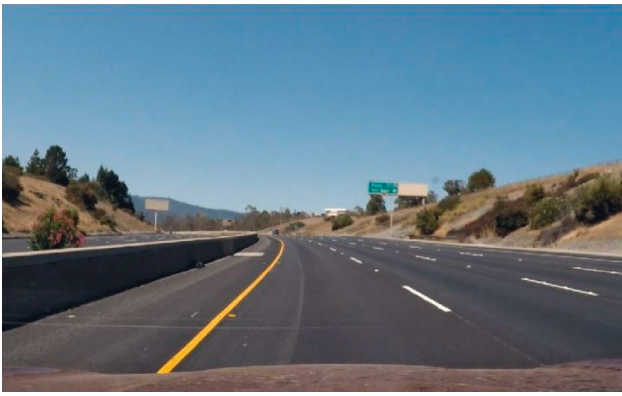
FIGURE 3: Original picture frame.
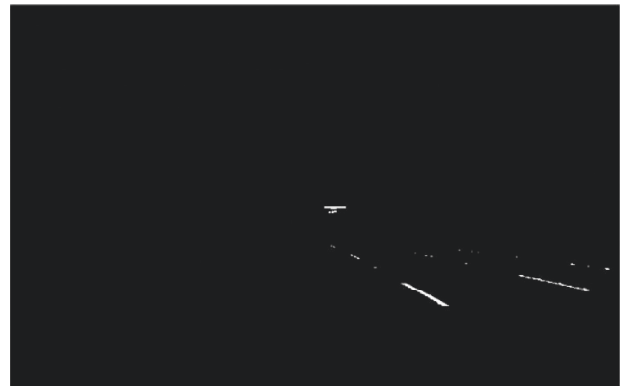
FIGURE 4: Picture frame after applying color masking.

3.2.3. Color Masking. Color masking is the conversion of a particular color pixel in an image or relative color pixels of an image into white and other colors into black by altering the pixel values, as shown in Figure 4. It is similar to thresholding, but here, instead of a grey scale image, we will take a color image.

For example, to detect the white color in an image, the RGB value of white color (R, G, B) = (255,255,255) is applied. If (R, G, B) = (250 − 255, 250 − 255, 250 − 255), then the color mask of the white color is given by

$$P(x, y) = \begin{cases} 1, & if\ C(x, y, z) = (R, G, B), \\ 0, & otherwise. \end{cases} \tag{2}$$

3.2.4. Edge Detection. The term "detection" refers to the process of identifying rapid changes in the intensity of a pixel in a picture. In this work, the Canny edge detection method is used to identify the edges, whose output is as shown in Figure 3 when applied over Figure 4.

The Canny edge detection algorithm is a gradient-based edge detection technique; to use this method, first we have to optimize the signal-to-noise ratio using a Gaussian filter or any other filtering technique Algorithm 1.

Canny edge detection algorithm steps:

3.2.5. Region of Interest Selection. There are different methods for a region of interest selection; in this work, the approach of selecting the polygon in the image formed from a certain set of predetermined points is followed, as shown in Figure 5.

3.2.6. Hough Transform. The Hough transform is an image processing technique for isolating features of a specific shape in an image. The Hough transform is most typically used to detect regular curves like lines, ellipses, circles, and so on.

The Hough transform can be used to find lines in a picture using the parametric representation of a line:

$$\rho = x * \cos(\theta) + y * \sin(\theta), \tag{3}$$

where ρ is the distance along a vector perpendicular to the line from the origin to the line, and θ is the angle formed between the vector and the x-axis.

3.2.7. Turn Prediction. Fitting a trapezium to the visual data gathered in a single frame is commonly used to direct turn prediction from the top down. Both lanes, which are modeled as a 2D path with left and right borders, employ similar model-fitting approaches. By adopting a smooth path model with limits on its breadth and curvature, the noisy bottom-up path detection is enhanced. In the temporal

Step 1:
(i)   Smoothing: Using the Gaussian filter
(ii) $f_x = f * G_x$, where $G_x = (-x/\sigma^2).G(x, y)$
(iii) $f_y = f * G_y$, where $G_y = (-y/\sigma^2).G(x, y)$
Step 2:
(i)   Gradient magnitude: Sobel or Prewitt
      $|G| = \sqrt{G_x^2 + G_y^2},$
      $\approx |G_x| + |G_y|.$
Step 3:
(i)   Edge direction
      $\theta = \tan^{-1}(G_y G_x)$
Step 4:
(i)   Resolve Edge Direction
Step 5:
(i)   Nonmaxima suppression: keep all local maxima in the gradient and remove everything else
(ii) Gives a thin edge line
Step 6:
(i)   Double or hysteresis thresholding

ALGORITHM 1: Canny edge detection Algorithm.

integration stage, this path representation is often fine-tuned by comparing it to earlier frames. A path is often described by its boundary points or its centerline and lateral extension at each centerline position, which defines the borders uniquely.

The last step is to predict the lane heading for the vehicle. This is calculated using the position of the vehicle for lanes at any given point. The pixel coordinates of the extrapolated lane lines are stored in the form of a matrix. First, the thick 2D plane of both lanes is converted to a 1D line at the center of the plane, using the concept of projection, i.e., projection of $\overline{b}$ on $\overline{a} = (\overline{a}.\overline{b}/|\overline{b}|)$.

First, the central point of the lane is calculated using the position of both the left and right lanes. Then, these central coordinates are stored in another matrix. Further, the perpendicular distance between the two lanes is calculated by measuring the projection of one lane over the other. This gives the central point of the road. Then, the vanishing point is found by taking a cross product of the obtained two vectors.

4. Algorithm

The algorithm for the proposed work is given as Algorithm 2.

The radius of curvature is calculated using the derivatives of the polynomial equation estimated using the polyfit parameters. Both the first and the second derivatives of the curve are used for calculating the radius of curvature for both the right and the left lanes separately. The radius of curvature is given as follows:

$$R_{\text{curve}} = \frac{\left[1 + (dx/dy)^2\right]^{3/2}}{\left|d^2 x/dy^2\right|}. \tag{4}$$

Then, this radius of curvature and the position of the center of the vehicle, that is, the center of the image is used to predict the lane turn. The difference between the lane center

FIGURE 5: Picture frame after selecting the ROI.

and the image center is calculated. This is calculated for both lanes. The vector cross product is calculated to obtain a vanishing point.

$$\overline{A} \times \overline{B} = |\overline{A}||\overline{B}|\sin \theta. \tag{5}$$

The ratio of the vanishing point and the frame size results in a new parameter called the vanishing ratio. This value of the vanishing Ratio is compared to predict the turns as follows:

If vanishing ratio is 0.5, then "Straight" is predicted.

If $0.47 \leq$ vanishing ratio $M < 0.49$, then the "Left" turn is predicted.

If the vanishing ratio> 0.5, then the "Right" turn is predicted.

5. Result and Discussion

This section shows how the proposed road lane recognition system performs in real-world scenarios. The left and right lane markers are detected and extracted by the algorithm. The suggested algorithm is put to the test using video pictures captured with an onboard vision sensor.

Step 1: A couple of videos are used as data to detect the lanes in the videos. The video is loaded to read each frame using the function *VideoReader* of Matlab.
Step 2: The Gaussian filter is applied to every frame in the image to remove noises in the image using the function *imgaussfilt3*.
Step 3: Yellow and white color masking is done to the frame.
Note: Lanes on the road are present in yellow and white colors.
Step 4: Then the *Canny* edge detection algorithm is applied to identify the edges of the yellow and white color regions in the image.
Step 5: The unwanted small group of pixels is removed using the function *bare open*.
Step 6: The region of interest is selected using the function *roipoly* in every frame of the video. This yields the region of the image where only lanes are present.
Step 7: The Hough transform is applied to the interested region to detect the lanes using the function *hough*. Further, Hough peaks function is used to determine the parameter space's peak values. The potential lines in the input image are represented by these peaks.
Step 8: The Hough lines function is used to locate the endpoints of line segments in the Hough transform that correspond to peaks. This function fills in minor gaps in line segments automatically.
Step 9: The slope of the left and right lanes is calculated, which is further used to extrapolate the broken road lanes if any.
Step 10: The cross product of both the vector lines is calculated to obtain the vanishing point, and its ratio to the frame size yields the vanishing ratio.
Step 11: The turn is predicted based on the obtained magnitude of the vanishing ratio (VR).
(i) Straight if VR ≃ 0.5.
(ii) Left if $0.47 \leq VR < 0.49$.
(iii) Right if $VR > 0.5$.

ALGORITHM 2: Hough transform method.
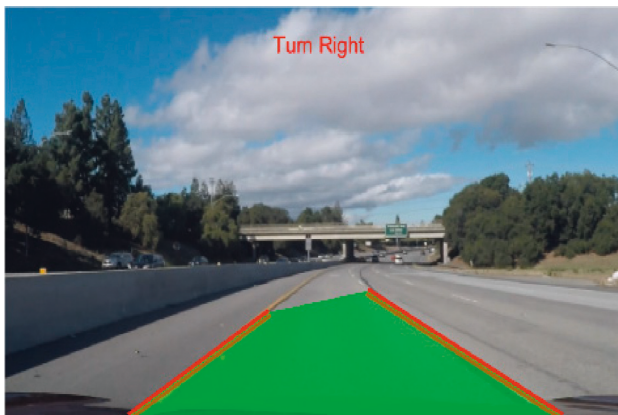
FIGURE 6: Picture frame predicting the right turn.

FIGURE 8: Picture frame predicting the straight direction.

5.1. Evaluation Metrics

5.1.1. Accuracy. The accuracy of the proposed method is 92%. In the proposed technique, we used a custom dataset of around fifty frames; in fifty frames, forty-six frames showed the correct path, so the accuracy of the proposed method is 92%.

5.1.2. Error Rate. The error rate of the proposed method is 8%.Out of fifty frames, forty-six frames show correct output, so we consider the reaming part as an error rate.

The capacity to evaluate algorithms is necessary for comparing and identifying the performance of various

FIGURE 7: Picture frame predicting the left turn.

TABLE 1: Comparison of the results of different proposed methods.

	Liu et al. [1]	Kim et al.[12]	Proposed method
Dataset	Custom dataset	Custom dataset	Custom dataset
Accuracy	91.1%	84.54%	92%
Error rate	8.9%	15.45%	8%

algorithms and methodologies. However, because of a lack of standardized test methodologies, performance measures, and publicly available datasets, this issue is extremely troublesome in the lane and road detection literature. Many works result in terms of efficiency only, but the complete breakdown of the number of frames tested, the number of frames successful, and the number of failure frames are not quoted for the existing methods. The lanes in front of a vehicle are successfully detected, and the path between those lanes is plotted in green color and lanes are marked in red color. It is clear from Figure 6 that the right sideline (plotted) is ahead of the left side lane, which indicates that there is a slight right curve in the road, so the steering of the vehicle should turn slightly towards the right side. This is predicted by finding the vanishing points and vanishing ratios of the two lanes.

If the left side lane (plotted) is ahead of the right side lane, it can be concluded that the road ahead has a curve towards the left, as in Figure 7, and hence, the vehicle steering has to turn in a left direction. Else if both the left and right lanes are equal in length, as in Figure 8, it can be inferred that the vehicle needs to proceed straight, with the steering in the steady position, since the vanishing ratio is 0.5. The direction in which the car has to travel is displayed over each frame. Table 1 shows the comparative result of the existing result and proposed result.

6. Conclusion

Lane detection is often complicated by varying roads, markings, clutter from other vehicles and complex shadows, lighting, occlusion from vehicles, and varying road conditions. In this work, we have presented a solution to the lane detection problem that shows robustness to these conditions.

This work focuses on road lane detection using basic image processing techniques. The lane detection algorithm presented has been tested on both straight and curved lanes. The algorithm passes through a series of low-level image processing to make useful information easier to be extracted. Then the road features are extracted with Canny edge detection. Hough transform is applied to find relevant lines that can be used as the left and right lane boundaries. Finally, the collected left and right lane boundaries are averaged to get the desired result of left and right boundaries and are shown on the original image. To reduce the high computational cost, the image is reduced to a smaller region of interest. The smaller threshold value is applied to edge detection for more complex environments. Experiments show that this algorithm can achieve fairly good performance in different kinds of road conditions.

The future scope of the work is that this model can be updated and tuned with more efficient mathematical modeling, whereas the classical OpenCV approach is limited and no upgrade is possible as the approach is not efficient. It is unable to give accurate results on the roads which do not have clear markings present on the roads. Also, it cannot work for all climatic conditions.

This technology is increasing the number of applications such as traffic control, traffic monitoring, traffic flow, and security. This can be integrated with a vision-based obstacle-detection algorithm, for example [8], for a collision-warning system. This can be further expanded to incorporate piecewise estimates of trajectory and curvature. Also, this system can be combined into a vehicle surround analysis system to create an intelligent vehicle driver assistance system. This lane detection algorithm can be applied with position tracking to reduce computational load. With real-world positioning applied to the image, distance and time to departure can be obtained.

With the combined parameters, a new lane departure warning system can be introduced.

Data Availability

This study used the data available here: https://www.kaggle.com/soumya044/advanced-lane-detection.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] W. Liu, F. Yan, J. Zhang, and T. Deng, "A robust lane detection model using vertical spatial features and contextual driving information," *Sensors*, vol. 21, no. 3, 708 pages, 2021.

[2] D. A. Goyal, D. M. Singh, and A. Srivatava, "Lane detection on roads using computer vision," *Vision International Journal of Engineering and Advanced Technology (IJEAT)ISSN*, vol. 9, pp. 1200–1205, October 2019.

[3] J. Jung and S. H. Bae, "Real-time road lane detection in urban areas using LiDAR data," *Electronics*, vol. 7, no. 11, 276 pages, 2018.

[4] C. Y. Low, H. Zamzuri, and S. A. Mazlan, "Simple robust road lane detection algorithm," in *Proceedings of the 5th International Conference on Intelligent and Advanced Systems (ICIAS)*, pp. 1–4, Kuala Lumpur, Malaysia, June 2014.

[5] A. Bar Hillel, R. Lerner, D. Levi, and G. Raz, "Raz advanced technical center Israel, general motors R&D, 7 HaMada st., 46725 herzliya, Israel, "recent progress in road and lane detection: a survey"," *Machine Vision and Applications*, vol. 25, no. 3, pp. 727–745, 2014.

[6] K. Ghazali, R. Xiao, and J. Ma, "Road lane detection using H-maxima and improved hough transform," in *Proceedings of the Fourth International Conference on Computational Intelligence, Modelling and Simulation*, pp. 205–208, Kuantan, Malaysia, September 2012.

[7] A. Assidiq, O. O. Khalifa, M. R. Islam, and S. Khan, "Real time lane detection for autonomous vehicles," in *Proceedings of the International Conference on Computer and Communication Engineering*, pp. 82–88, Kuala Lumpur, Malaysia, May 2008.

[8] Z. Kim, "Real-time obstacle detection and tracking based on constrained Delaunay triangulation," *Proceedings of the 2006 IEEE Intelligent Transportation Systems Conference*, pp. 548–553, Toronto, ON, Canada, September 2006.

[9] J. C. McCall and M. M. Trivedi, "An integrated, robust approach to lane marking detection and lane tracking," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 533–537, Parma, Italy, June 2004.

[10] D. Li, D. Zhao, Q. Zhang, and Y. Chen, "Reinforcement learning and deep learning based lateral control for autonomous driving [application notes]," *IEEE Computational Intelligence Magazine*, vol. 14, no. 2, pp. 83–98, 2019.

[11] Y. Xing, C. Lv, L. Chen et al., "Advances in vision-based lane detection: algorithms, integration, assessment, and perspectives on ACP-based parallel vision," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 3, pp. 645–661, 2018.

[12] Z. W. Kim, "Robust lane detection and tracking in challenging scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 16–26, 2008.

[13] E. Romera, J. M. Alvarez, L. M. Bergasa, and R. E. Arroyo, "Efficient residual factorized convnet for real-time semantic segmentation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 1, 2017.

[14] J. Li, X. Mei, D. Prokhorov, and D. Tao, "Deep neural network for structural prediction and lane detection in traffic scene," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 3, 2017.

[15] W. Yang, X. Zhang, Q. Lei, D. Shen, P. Xiao, and Y. Huang, "Lane position detection based on long short-term memory (LSTM)," *Sensors*, vol. 20, no. 11, 3115 pages, 2020.

[16] G. R. Nagarjuna, R. Shashidhar, S. B. Puneeth, and B. N. Arunakumari, "IoT enabled smart traffic system for public and emergency mobility in smart city," in *Proceedings of the 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 53–59, Palladam, India, October 2020.

[17] D. Koundal, "Texture-based image segmentation using neutrosophic clustering," *IET Image Processing*, vol. 11, no. 8, pp. 640–645, 2017.

[18] H. Yar, T. Hussain, Z. A. Khan, D. Koundal, M. Y. Lee, and S. W. Baik, "Vision sensor-based real-time fire detection in resource-constrained IoT environments," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–15, Article ID 5195508, 2021.

[19] G. Madhu, B. Lalith Bharadwaj, R. Boddeda et al., "Intelligent disease diagnosis model for energy aware cluster based IoT healthcare systems," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1189–1203, 2022.

[20] K. Bhalla, D. Koundal, S. Bhatia, M. Khalid Imam Rahmani, and M. Tahir, "Fusion of infrared and visible images using fuzzy based siamese convolutional network," *Computers, Materials & Continua*, vol. 70, no. 3, pp. 5503–5518, 2022.

Retraction

Retracted: The Use of Artificial Intelligence for Smart Decision-Making in Smart Cities: A Moderated Mediated Model of Technology Anxiety and Internal Threats of IoT

Mathematical Problems in Engineering

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] A. Alloulbi, T. Öz, and A. Alzubi, "The Use of Artificial Intelligence for Smart Decision-Making in Smart Cities: A Moderated Mediated Model of Technology Anxiety and Internal Threats of IoT," *Mathematical Problems in Engineering*, vol. 2022, Article ID 6707431, 12 pages, 2022.

Research Article

The Use of Artificial Intelligence for Smart Decision-Making in Smart Cities: A Moderated Mediated Model of Technology Anxiety and Internal Threats of IoT

Abdelmenem Alloulbi [ID], Tolga Öz [ID], and Ahmad Alzubi [ID]

Business Administration Department, Karpas Mediterranean University, Institute of Graduate Research and Studies, TRNC, Mersin 10, Turkey

Correspondence should be addressed to Ahmad Alzubi; ahmad.alzaubi@akun.edu.tr

With the rapid development of artificial intelligence (AI), AI for smart decision-making is attracting a lot of attention, but research on this topic in smart cities is not yet comprehensive. Thus, the current research aimed to examine the direct and indirect (via technology anxiety) relationships between artificial intelligence (AI), technology anxiety, and smart decision-making (SDM). The research article also examines the moderating role of internal threats of IoT on AI and technology anxiety and the relationship between AI and smart decision-making. 614 cross-sectional data gathered from participants from public and private sectors in Turkey were utilized to investigate the aforementioned relationships. The results indicated that AI had a positive influence on smart decision-making. AI contributes negatively to technology anxiety. Technology anxiety has a negative effect on smart decision-making. Technology anxiety partially mediated the direct effect of AI on smart decision-making. The results revealed that internal threats of IoT moderated the negative relationship between AI and artificial intelligence, such that the negative relationship is further strengthened when internal threats of IoT are high. The results also indicated that internal threats of IoT moderated the positive direct relationship between AI and smart decision-making, such that the positive relationship is weakened when internal threats of IoT are high. The findings present crucial practical implications for government and local authorities in building smart cities.

1. Introduction

The use of artificial intelligence (AI) has been claimed to provide transformational potential in different areas, ranging from smart cities, its influence on governance, and innovative discipline, and even promote human capabilities [1–4]. With the advancement of technology, AI has taken a significant leap becoming an essential part of everyday life [5]. From this standpoint, AI and its transformational potential had become hot topic for discussion in both practice and literature in the modern era [1, 6]. In both government and private sectors, data generation through AI is possible in exploring innovative ways to comprehend our world. With the rapid advancement of big data technologies and their availability, AI is being revitalized while enabling informed

decision [7]. AI promotes smart cities' decision-making because SDM involves the systematic technique to data collection and its application of logical decision-making methods rather than generalizing from experience, intuition, or the use of trial and error [8].

"Smart cities" has been defined in various ways by several scholars because it is a multifaceted concept. Nevertheless, the necessary requirement for smart cities is to promote quality of life and attain sustainable development using AI and information and communication technology (ICT) [9]. Based on this, Atlam et al. [10] defined a smart city with respect to the technology aspect as "a technologically interconnected city" or Internet of things (IoT) utilizing big data to attain intelligent and efficient handling of cities resources. Albino et al. [11] research regarding SDM in

smart cities through big data conceptualized three-dimensional layers distinguishing feature of a smart city, namely, instrumentation, interconnection, and intelligence. Smart cities in the implementation stage use AI and IoT for the purpose of data collection through the usage of sensors, meters, cameras, and social media for fast feedback. The data gathered through the use of AI, IoT, and the above-mentioned sources are then incorporated and transformed into even-related message in the interconnection stage to produce a deeper understanding for smart city decision-making. Consequently, the collected information through the data is then projected to comprehend the city's needs, demands, requirements, and needed policies, thus promoting well-rounded and smart decision-making [11, 12].

With smart cities impacting various areas of human's life, AI and IoT system may be subjected to information attacks, leading to security problems such as user data disclosure. Based on this, Roman et al. [13] argued that one of the main issues that must be overcome concerning IoT is security. People's perception and attitude are crucial determinants of success for any smart services [14]. The social cognitive theory postulates that anxiety is a crucial determining factor in behavioral intentions [15]. Technology anxiety is a major factor influencing the use of self-service technology [16]. Technology anxiety relates to "the fear, apprehension, and hope that individuals feel when considering use or actually using certain technology" [16]. This reenergizes the idea that technology anxiety could play a crucial role in the relationship between AI and SDM. Consequently, technology anxiety is introduced as a mediator in the relationship between AI and SDM.

Prior studies have acknowledged and discussed the impact of AI on smart decision-making process. However, several factors can influence decision-making in smart cities [17]. Furthermore, our understanding of the precise usage of AI on smart decision-making in smart cities is still limited, especially in Turkey. From a research point of view and to the best of our knowledge, no study has examined the four constructs in this study simultaneously. Therefore, in furthering the existing literature, this study aimed to test a moderated mediation model where both the direct and indirect relationships between the uses of AI in smart decision-making are mediated by technology anxiety and moderated by internal threats of IoT (Figure 1).

1.1. Theoretical Background and Hypotheses

1.1.1. AI.
In the public sector, e.g., municipalities, Mikalef et al. [18] noted that the deployment levels of AI are still at the early stage. Leveraging AI in such context is subjected to variety of various forces and is hampered by factors such as legal, political, and policy challenges [3]. Studies related to AI in the public organization have attracted limited attention, although lately there has been a rise in AI-related research and technical reports (see [4, 19, 20]). These studies emphasized the importance of the application of AI in the public organizations across various branches of governance, indicating that the massive potential of AI is being diminished by organizational, technical, and policy difficulties.

Prior studies have substantiated the crucial impact of AI in rational decision-making procedures, facilitating a high-quality life and making a city smart in practice. AI-enabled governance has been characterized as enabler of collaboration among cities to establish smart services that could have been impossible for a single city to create [21]. AI-enabled governance in smart cities promotes collection of data through sensors and other sources in enhancing urban safety governance [22]. Most recently, the South Korean government used AI in response to the coronavirus pandemic to provide an exchange of information to its citizens, helping them in comprehending the situation and applying government-released safety measures [23].

AI using information technology is the cornerstone of smart cities [24]. As the world population grows and increases urbanization, AI technologies are increasingly being used to improve quality of life: smart traffic [25], intelligent information management [26], and smart health care [27]. Such evolution typically consists of various AI approaches that are widely adopted by city's decision-makers who establish the governing system [28]. AI through intelligent information processing or data analysis can improve data cleaning, data collection, and storage to gain more insight from generated data. Such a crucial attribute is a vital element of automated learning and decision-making process by complex mathematical depiction of the issues.

In government, AI can help to enhance power structure through smart decisions by providing new and smart services. Such qualities in decision-making process can offer solution to common mistakes in administrative decision-making procedures such as improper forecast of administrative tasks [1].

1.1.2. Artificial Intelligence and Decision-Making in Smart Cities.
AI has well been documented to aid decision-making in several disciplines ranging from private and public organizations to smart cities. AI has been reported to be very useful in reinventing business models and ecosystems, thus promoting decision-making [1]. The usefulness of new generation of AI systems to forecast strong-impact weather is widening human's ability to explore huge amount of data in order to gain insights and serves as a proper guide for analysts and decision-makers [29]. Several private organizations and governmental organizations across the world provide open datasets online that can be unutilized for data analysis and decision-making. From this standpoint, Marr [30] argued that the actual value of big data is not in the huge volume of data but in the advancement of new generation of AI systems such as machine learning ability to analyze complex and massive dataset beyond anything we could ever do before. Such evolution would provide rich insights for decision-making.

The building of a smart city is a complex, systematic, and extremely advanced integrated project that includes sensing devices, information collection, and monitoring of infrastructure to enhance decision-making [31]. Furthermore, prior studies pointed out that processing and the interpretation of big data through AI are a huge step aimed at
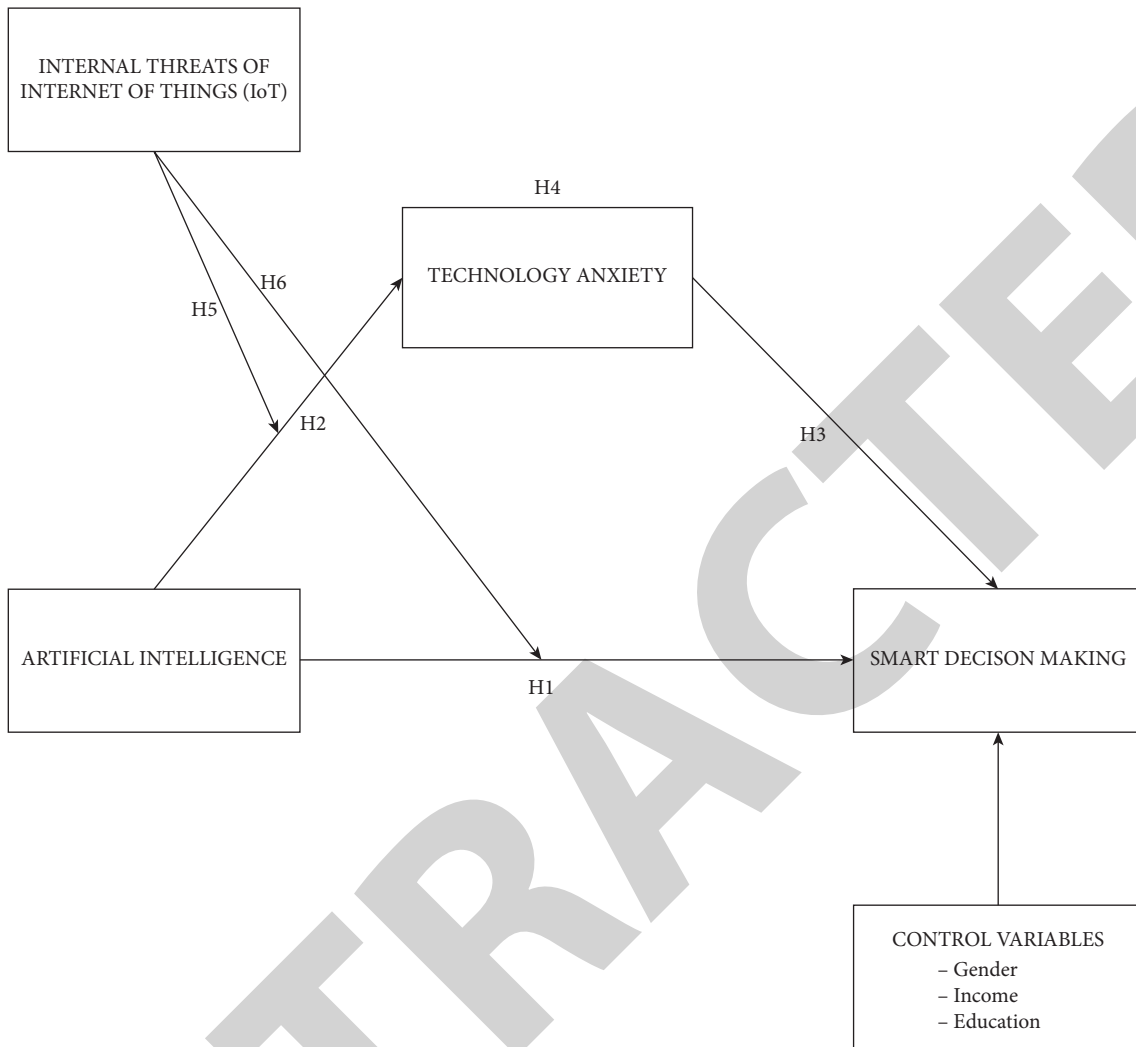
Figure 1: Conceptual model.

improving decision-making in smart cities [8, 32]. Digital ideas associated with AI technologies are enabling cities to become smart as the push for modern concept of "smart" is being made increasingly possible through the collection of real-time data and interpretation of such data to gain insight into how cities transform, adjust, and react to diverse environments [32]. Thus, we hypothesized that

H1: artificial intelligence using big data is positively related to SDM in smart cities

1.1.3. AI and Technology Anxiety. With the accelerated growth of AI, people have slowly begun to show concerns about AI [33, 34]. From this standpoint, numerous studies have indicated that AI technologies had overtaken humans in several aspects; AI performs substantially better than humans in the game of Go [35] and defeated 99.8% of human players in StarCraft [36]. Additionally, several research and professional assessments have also stimulated anxiety. About 400 to 800 million jobs were forecasted by the McKinsey Global Institute to be replaced by A1 by 2030 [37].

A large number of expert opinions in addition to facts indicated that technology anxiety through the use of AI has become a global phenomenon that will highly affect people's life paths, future study, and work [33, 38]. Furthermore, AI has been suggested to create technology anxiety and a succession of other social issues in smart cities [39]. Based on the above evidence and reasoning, we hypothesized that

H2: AI contributes negatively to technology anxiety in smart cities

1.1.4. AI, Technology Anxiety, and SDM. A smart city utilizes information system-centric procedures through intelligent usage of ICT within an interactive infrastructure to offer its citizen innovative and upgraded amenities, therefore influencing the quality of life of the people [40]. Technology has changed the way individuals work, the balance between personal and work-life [41]. However, in several fields of modern life, individuals are anxious about the possible existence of superintelligent technologies and the anxiety tends to be growing more [42]. From this standpoint,

technology anxiety plays a crucial role in the endorsement of various smart services [14] and thus smart decision-making.

As AI becomes more and more important in public organizations, the question of how AI should be integrated into decision-making process is becoming increasingly more important. In the context of technology adoption, technology context or technology anxiety has been suggested to play a crucial role in the adoption of disruptive technologies, for example, big data [43] and cloud computing [44]. This offers an essential orientation point for research of AI in public organizations. Furthermore, the technology context describes the impacts of perceptions of technology. The qualitative study of Schaefer et al. [45] pointed out some of the challenges pertaining to AI tool adoption via interviews of municipal employees in Germany. The study highlighted how self-service systems are constantly optimizing and evolving themselves, which indeed can help municipality employees to streamline and automate processes, thus smart decision-making. However, AI anxiety or technology anxiety in particular entails the independent self-evolution of machine and the iteration speed of AI is a lot quicker compared with that of humans [34]. Additionally, AI can make independent decisions and operate autonomously of humans [46], which can rest in unpredictable harm [47]. For this reason, people with high technology anxiety are more unlikely to use it [16], thus influencing their smart decision-making. As argued by Little et al. [48], many direct relationships rely on contextual factors, so in this study we examine the role of technology anxiety on the direct relationship. In accordance with the above empirical evidence and argument, we posit that

> H3: technology anxiety negatively influences smart decision-making in smart cities
>
> H4: technology anxiety plays a mediating role between AI using big data and smart decision-making

1.1.5. Internal Threats of IoT as a Moderator.

1.1.5. Internal Threats of IoT as a Moderator. IoT relates to the network of interconnected physical devices across the globe that are equipped with connectivity to collect gather, transfer, and share data [26, 49]. IoT development is important because it helps cities, buildings, and management services make smarter plans for change through ICT [50]. Many private companies have increased the number of remote and sensor monitors used in managing an organization's environment, which encouraged the government to adopt new technologies in smart cities to promote energy efficiency, traffic congestion reduction, and improve water and air quality. The IoT devices such as electronic communications, social networks, machinery, digital hearing aid, Fitbit, or GPS create streams of data by connecting and monitoring people [51–53]. However, these devices or systems commonly make use of wireless communication, sometimes with open characteristics that bring convenience to users, threatening the system security and users' privacy [54]. IoT encounters challenges such as internal threats due to security challenges of the systems because they can be easily hijacked by cybercriminals [54]. From this standpoint, prior research has indicated that technology anxiety, especially within the context of computer-related system and information

TABLE 1: Demographic information.

Characteristics ($n = 614$)	n	Percentage (%)
Gender		
Male	338	63.2
Female	226	36.8
Income (Turkish Lira)		
Less than 5000	165	26.9
5000–10000	195	31.8
10001–15000	130	21.2
15001–20000	95	15.5
20001 and above	29	4.7
Education		
High school	123	20.1
Vocational degree	260	42.3
Bachelor degree	147	23.9
Master degree and above	84	13.7
Marital status		
Married	321	52.3
Single	207	33.7
Divorced	80	13.0
Widowed	6	1.0

services, is quite common [55, 56]. Consequently, individuals with greater concerns regarding internal threats of IoT are more likely to stick to the services they are used to, rather than switching to some new generation AI technologies. Thus, we argue that high internal threats from IoT can further strengthen the negative relationship between AI and technology anxiety.

Roman et al. [13] suggested that privacy and security concerns are the major determinants that will impact the adoption of IoT in the development of a smart city. The study further suggested that should the problems be too complex and the advantages too little, individuals may stick with the traditional services they are accustomed to. Furthermore, the main idea behind IoT is global connectivity ("access anyone") and accessibility ("access anyhow, anytime"), and such rationale makes the threats that can affect IoT systems enormous [57, 58]. Such threats regarding how IoT systems work have been suggested to impact decision-making process in adopting smart services [59]. Together, we expect that internal threats from IoT may mitigate the influence of artificial intelligence on SDM in smart cities. Thus, we hypothesize the following:

> H5: internal threats of IoT moderate the negative relationship between AI using big data and technological anxiety, such that the negative relationship is strengthened when internal threats of IoT are high.
>
> H6: internal threats of IoT moderate the positive relationship between AI using big data and SDM, such that the positive relationship is weakened when internal threats of IoT are high.

2. Methods

2.1. Sample and Procedures. The participants of this study were Turkish adults between the ages of 21 and 65 years. A questionnaire survey was sent electronically through email containing the link to survey to the targeted audience to achieve

the study's objectives. The participants were asked to complete the questionnaires and forward it to us via the email. All the measurement items adopted in the current research were primarily in English; however, our sample is from Turkey. Hence, with the help of two different professional experts, the items were translated into Turkish and then translated back into English to ensure double-check for precision. A total number of 723 questionnaires were distributed; 614 valid responses were recovered with a response rate of 84.92%. The participants were requested to give their opinions regarding the use of AI in smart cities, their perceptions about technology anxiety, and internal threats of IoT associated with public services and smart decision-making. All the constructs in the current research were rated on a 5-point Likert scale (where 1 = strongly disagree to 5 = strongly agree).

The demographic information is illustrated in Table 1. 338 (63.2%) of the participants were males, and 226 (36.8) were females. 449 (73.2%) earn above 5,000 Turkish Lira, and a large proportion of the participants (491 (79.9%)) had at least vocational degree. In terms of marital status, the majority of the participants were married (321 (52.3%)).

2.2. Measurement Items.
Artificial intelligence was measured using 5 items adopted from [60, 61]. The items were worded to extract participants' perceptions about AI trustworthiness and the influence on the society.

Technology anxiety was measured using 5 items from [14]. It measures participants' perceptions of smart services in smart cities.

Internal threats of IoT were measured using 3 items from [62]. A sample item is "most IoT devices operate unattended by humans; thus, it is easy for an attacker to physically gain access to them."

Smart decision-making was measured using 5 items from [63] regarding the use of new technology for decision-making.

2.3. Control Variables.
Gender, income, and education were controlled in this study.

2.4. Statistical Methods.
Data collected were analyzed using SPSS 28 and AMOS 26 software. In particular, AMOS 26 was used for confirmatory factor analysis to estimate the measurement model of all the variables in this study. SPSS 28 was used for descriptive statistics, common method bias, Pearson's correlation analysis, and PROCESS (plug-in) compiled by Hayes. PROCESS models 4 and 8 were chosen to examine the mediation model and moderated mediation model, respectively [64]. 5000 bootstrap resamples with confidence interval (CI) of 95%; that is, where the 95% CI excludes zero indicates a significant conditional indirect effect.

3. Results

3.1. Common Method Variance.
Data were collected via self-report measure. To rule out the possibility of common method variance (CMV) associated with self-report method,

the respondents were told to rate the items anonymously. Furthermore, the Harman single factor test was adopted. The results showed that the first component for variance interpretation rate accounted for 23.61%; this is below the critical threshold of 50%, suggesting that CMV was not a major concern in this research [65]. In addition, the variance inflation factor (VIF) results for correlation between our study's constructs were below 3, implying that there are no substantial collinearity problems in the results [66].

3.2. Measurement Model.
Testing for data distribution, Lei and Lomax [67] indicated that skewness and kurtosis indexes should not surpass $|2.3|$. As demonstrated in Table 2, the skewness values are from 0.027 to 1.235 and kurtosis values from 0.013 to 1.553. Thus, the measured items are within the acceptable thresholds, indicating that the collected data are normally distributed.

Next, we test the reliability and validity of the measurement items. The factor loadings for all items were found to be higher than 0.6. To fulfill convergent validity, the average variance extracted (AVE) of each variable is estimated [68]. The minimum threshold for AVE should be 0.5 [69]. To estimate composite reliability, the construct reliability (CR) was measured for each variable. Its lowest threshold should be 0.7 [70]. The results as illustrated in Table 2 revealed that all factor loading ranged from 0.602 to 0.989, AVE ranged from 0.611 to 0.799, and CR ranged from 0.871 to 0.952, thus suggesting that items are appropriate, and the variables are reliable and consistent.

Fornell and Larcker [68] recommended that for discriminant validity, the square root of each AVE is larger than the surrounding correlations, and then, discriminant validity is established. The results showed that the square of AVEs (illustrated in bold parenthesis) is larger than the surrounding correction, indicating evidence of discriminant validity as shown in Table 3.

Table 4 shows the confirmatory factor analysis (CFA) of the adopted model (research model). The model fit indices were estimated by several statistics: IFI, CFI, NFI, RMR, TLI, and RMSEA. IFI and TLI should be higher than 0.8, CFI should be higher than 0.9, RMR should be higher than 0.1, and RMSEA should be lower than 0.08 [71]. All fall within the recommended thresholds, indicating the research model, shows an acceptable fit with the data collected.

3.3. Testing for Mediation Model.
To examine hypotheses, H1, H2, H3, and H4, the PROCESS macro-model 4 compiled by Hayes was used. The relationship between AI and SDM was supported predicted in H1 ($\beta = 0.367$, $p = 0.000$, 95% CI exclude zero, 0.276 to 0.458) as illustrated in Table 5. This result revealed that AI using big data has a positive significant effect on SDM. The relationship between artificial intelligence and technology anxiety was validated as predicted in H2 ($\beta = -0.799$, $p = 0.000$, 95% CI = -0.276, -0.113). The results revealed that artificial intelligence has a negative significant effect on technology anxiety. Also, there was a negative significant relationship

TABLE 2: Reliability and validity assessments.

Constructs	Measurement items	Factor loading (λ)	Normal distribution	
			Skewness	Kurtosis
Artificial intelligence	($\propto$ = 0.953; CR = 0.952; AVE = 0.799)			
	AI1	0.989	−0.467	−1.212
	AI2	0.960	−0.489	−1.371
	AI3	0.800	−0.897	−0.556
	AI4	0.840	−0.825	−0.795
	AI5	0.860	−0.113	−1.505
Technology anxiety	($\propto$ = 0.907; CR = 0.889; AVE = 0.621)			
	TA1	0.980	−0.381	−1.410
	TA2	0.710	−1.089	0.013
	TA3	0.770	−0.015	−1.549
	TA4	0.640	−0.895	−0.248
	TA5	0.790	−0.008	−1.545
Smart decision-making	($\propto$ = 0.813 CR = 0.871; AVE = 0.611)			
	SDM1	0.900	0.107	−0.812
	SDM2	0.690	−0.275	−1.370
	SDM3	0.691	0.484	−1.416
	SDM4	0.680	0.027	−1.458
	SDM5	0.790	0.036	−1.553
Internal threat of Internet of things	($\propto$ = 0.869; CR = 0.873; AVE = 0.688)			
	IOT1	0.602	−1.235	1.042
	IOT2	0.980	−0.665	−0.920
	IOT3	0.950	−0.700	−0.812

Note. (1) AI = artificial intelligence; TA = technology anxiety; SDM = smart decision-making; IOT = internal threat of Internet of things; (2) AVE = average variance extracted; CR = composite reliability; $\propto$ = Cronbach's alpha.

TABLE 3: Descriptive statistics, Pearson's correlation matrix, and discriminant validity.

Construct	M	SD	AI	TA	SDM	IOT	Gender	Income	Education
AI	3.711	1.409	**0.799**						
TA	3.490	1.312	0.649**	**0.621**					
SDM	3.246	1.304	0.619**	0.522**	**0.611**				
IOT	2.866	1.990	0.538**	0.602**	0.591**	**0.688**			
Gender	1.370	0.483	0.662**	0.612**	0.566**	0.611**	—		
Income	2.394	1.171	0.544**	0.622**	0.540**	0.569**	0.573**	—	
Education	2.313	0.944	0.883**	0.827**	0.788**	0.622**	0.571**	0.608**	—

Note. (1) M = mean; SD = standard deviation; (2) correlation is significant at **p < 0.01 (two-tailed); (4) boldface in parentheses indicates that the square root of AVEs is greater than the off-diagonal correlations.

TABLE 4: Model fit statistics.

Goodness-of-fit index	Acceptable limit	Model value
CMIN/DF	<3 excellent fit	1312.759/633 2.073
IFI	>0.9	0.941
CFI	>0.9	0.939
NFI	>0.9	0.900
RMR	>0.10	0.129
TLI	>0.9	0.937
RMSEA	<0.08	0.049

between technology anxiety and smart decision-making (β = −0.322, p = 0.000, 95% CI = −0.246, −0.099).

With H1, H2, and H3 all being supported, for mediation analysis, 5000 bootstrapping was used. The bootstrap technique has become the accepted method for mediation analysis because of its ease and accuracy (see Hayes [72]). The distinction behind bootstrapping is the use of a form of resamples of the data available to make inference and gain deep insight into the underlying population [73]. Hence, to provide the most reliable result for mediation analysis, a bootstrap resamples of 5000 were used in this study. The results revealed that there was a significant negative indirect relationship between artificial intelligence and smart decision-making via technology anxiety (β = −0.293, 95% CI = −0.203, −0.101), and the confidence interval excludes zero. Therefore, hypothesis H4 was validated.

3.4. Testing for Moderation Model. To examine the moderating role of internal threats of IoT in the relationship between artificial intelligence and technology anxiety (H5) and the relationship between AI and SDM (H6), PROCESS macro (model 8) was utilized with gender, income, and education confounded for as covariates. The moderated mediation analysis is demonstrated in Table 6.

TABLE 5: Testing for direct and mediation effects.

Mediation results: technology anxiety partially mediates the relationship between artificial intelligence and smart decision-making (PROCESS model 4, 95% CI)
Bootstrap 95% CI

	β	SE	t	p	LLCI	ULCI	R^2
Model 1: mediator variable model		Outcome: TA					
Artificial intelligence	−0.799	−0.016	−9.827	0.000	−0.276	−0.113	0.413
Model 2: outcome variable model smart decision-making							
Technology anxiety	−0.322	−0.041	−7.803	0.000	−0.246	−0.099	0.571
Artificial intelligence	0.367	0.047	7.880	0.000	0.276	0.458	
Bootstrapped results for the indirect effect (the indirect effect of AI on SDM via TA)	−0.293	0.048			−0.203	−0.101	

Note. $n = 614$; bootstrap resample = 5000; LLCI = lower level of confidence interval; ULCI = upper level of confidence interval.

TABLE 6: Testing for moderated mediation: internal threat of Internet of things moderated the direct and indirect relationship between artificial intelligence and smart decision-making (model 8, 95% CI).

Bootstrap 95% CI

	β	SE	t	p	LLCI	LLCI	R^2
Model 1: mediator variable model		Outcome: technological anxiety					
Artificial intelligence	−0.752	−0.028	−26.465	0.000	−0.696	−0.407	0.418
Internal threats of Internet of things	−0.119	−0.037	3.245	0.001	−0.087	−0.016	
Artificial intelligence X internal threats of Internet of things (interaction)	−0.066	0.016	−4.244	0.000	−0.136	−0.097	
Co: gender	−0.606	0.088	−6.912	0.000	−0.778	−0.434	
Co: income	0.314	0.029	10.598	0.000	0.256	0.372	
Co: education	0.244	0.063	3.518	0.001	0.108	0.379	
The conditional direct effect of artificial intelligence on technology anxiety							
Internal threats of Internet of things (−1SD)	0.665	0.049	13.418	0.451	−0.210	0.411	
Internal threats of Internet of things (+1SD)	−0.741	−0.042	−17.864	0.000	−0.068	−0.167	
Model 2: dependent variable model-dependent: smart decision-making							
Artificial intelligence	0.149	0.021	4.188	0.001	0.249	0.330	0.255
Technology anxiety	−0.314	−0.045	−6.923	0.000	−0.425	−0.201	
Internal threats of Internet of things	−0.488	−0.046	−10.516	0.001	−0.597	−0.379	
Artificial intelligence X internal threats of Internet of things (interaction)	0.089	0.055	2.700	0.001	0.108	0.258	
Co: gender	−0.538	0.102	−5.297	0.000	−0.737	−0.339	
Co: income	−0.137	0.036	−3.802	0.000	−0.208	−0.066	
Co: education	0.673	0.078	8.622	0.000	0.520	0.827	
The conditional direct effect of AI on SDM							
Internal threats of Internet of things (-1SD)	0.248	0.057	4.334	0.001	0.360	0.535	
Internal threats of Internet of things (+1SD)	0.051	0.063	1.816	0.001	0.175	0.372	
Bootstrapped results for indirect effect (via technology anxiety)							
Index of moderated mediation	0.020	0.006			0.010	0.032	
The conditional indirect effect of artificial intelligence on smart decision-making (via technology anxiety)							
Internal threats of Internet of things (−1SD)	−0.249	0.051			−0.125	−0.091	
Internal threats of Internet of things (+1SD)	−0.205	0.044			−0.101	−0.083	

Note. $n = 614$; bootstrap resample = 5000; LLCI = lower level of confidence interval; ULCI = upper level of confidence interval; Co = control variable.

The results showed that after confounding for demographic covariates, the effect of the interaction term between AI using big data and internal threats of IoT on technology anxiety was significant ($\beta = −0.066$, $p \leq 0.000$), and this provides support for H5 indicating that internal threats of IoT moderated the negative relationship between AI using big data and technology anxiety. The significant effect of the interaction was further investigated through simple slope analysis. The interactions were plotted at +1 and −1 SD from the mean of internal threats of IoT (demonstrated in Figure 2). We constructed a simple slope to check the strength of the relationship between AI using big data and technology anxiety at high and low levels of internal threats of IoT. The

results of the conditional direct effect of AI on technology showed that the negative relationship was stronger when internal threats of IoT were high ($\beta = −0.741$, $p \leq 0.000$), while the relationship was not statistically significant ($\beta = −0.665$, $p \geq 0.05$) at low level of internal threats of IoT, therefore further validating H5.

The results also revealed that the interaction effect of AI using big data and internal threats of IoT on SDM was significant ($\beta = −0.089$, $p \leq 0.001$), indicating that internal threats of IoT moderate the positive relationship between AI using big data and SDM. The interactions were plotted at +1 and −1 SD from the mean of internal threats of IoT (demonstrated in Figure 3). A simple slope analysis was
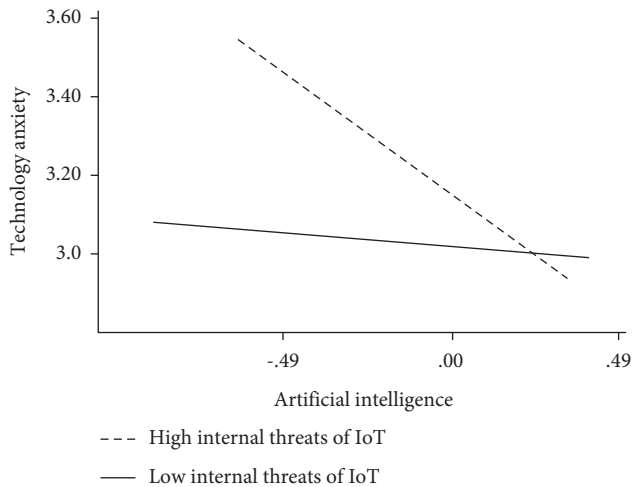
Figure 2: Association between AI and technology anxiety at different levels of internal threats of IoT: (1) low internal threats of IoT (1 SD below the mean) and (2) high internal threats of IoT (1 SD above the mean).
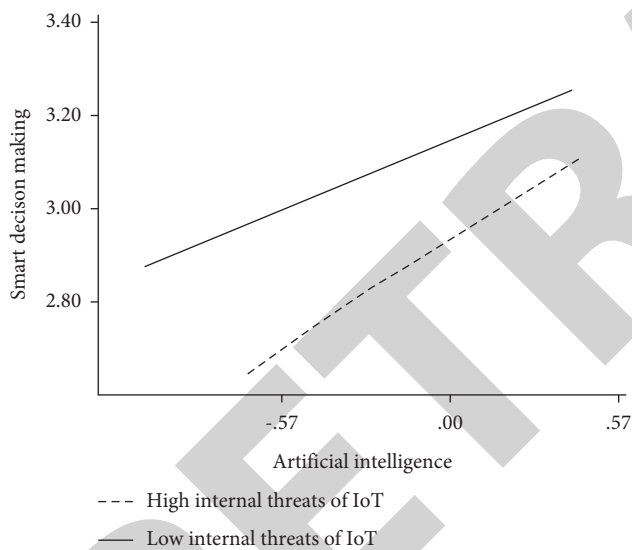
Figure 3: Association between AI and SDM at different levels of internal threats of IoT: (1) low internal threats of IoT (1 SD below the mean) and (2) high internal threats of IoT (1 SD above the mean).

constructed to examine the strength of the relationship between AI using big data on SDM at high and low levels of internal threats of IoT. The results of the conditional direct effect indicated that the relationship was stronger ($\beta = 0.205$, $p < 0.001$) when internal threats of IoT were low, while the relationship was weak ($\beta = 0.051$, $p \leq 0.001$) at high level of internal threats of IoT, therefore validating hypothesis 6.

Moreover, the results of the conditional indirect showed that internal threats of IoT moderated the indirect relationship (through technology anxiety) between AI using big data and SDM (bootstrapped estimate = 0.023, bias-corrected CI = 0.010 to 0.032). As illustrated in Table 6, the two

conditional indirect effects were statistically significant, plus bootstrap (confidence intervals) supported the results.

Finally, the results supported a moderated mediated model where technology anxiety is the partial mediator of AI using big data on smart decision-making.

4. Discussion

Based on the sample collected in Turkey, this study investigated a moderated mediation model and uncovered the underlying mechanism in the relationships between artificial intelligence using big data and SDM in the Turkish context. First, the results revealed that AI using big data has a positive impact on SDM in smart cities. This result aligns with the prior studies of [17, 32]. The consistencies of this pattern of result could suggest that AI technologies integrated with big data, complex algorithms, large storage, and processing capacity are influencing decision-making in smart cities. Second, AI was revealed to be a determinant of technology anxiety. This result aligns with Amodei [39] and supports the conclusions of [33, 38]. This particular result revealed that technology anxiety contributes negatively to technology in smart cities. Third, technology anxiety was found to be a significant and negative determinant of smart decision-making. This result provides empirical support for Meuter et al. [16], who suggested that the high level of technological anxiety will influence people's smart decision-making, by discouraging them not to use smart services in smart cities. Fourth, technology anxiety partially mediated the relationship between AI using big data and SDM.

Fifth, internal threats of IoT moderated the negative relationship between artificial intelligence and technological anxiety, such that the negative relationship is further enhanced when internal threats of IoT are high. Finally, the results showed that internal threats of IoT moderated the positive relationship between AI using big data and SDM, such that the positive relationship was weakened when internal threats of IoT were high.

4.1. Theoretical Implications. This study provides important theoretical implications. This study promotes our understanding of why and when artificial intelligence using big data is positively related to smart decision-making in smart cities. Bokhari and Myeong [17] made a research call that for theoretical advancement and practical success of AI applications in smart cities, it is crucial to recognize and investigate the indirect factors that may affect the good or adversarial relationship between AI and decision-making. Particularly, this research established that technology anxiety is a crucial intervening mechanism and opens up the black box for the relationship between AI and smart decision-making. Based on social cognitive theory [15] and self-service technology (SST) usage [16], our findings suggest that using AI technologies to improve both life and work in a substantial way may enrich smart decision-making; however, the anxiety related to its use can discourage people from using smart services, thus negatively influencing their smart making decision.

In the present era where sophisticated technologies offer a wide array of novel services, this study constructs a comprehensive model and empirically tested the relationships among the constructs leading to a substantial as-yet nascent study investigating the connection between AI and SDM in the smart city space. Our study revealed the internal threats of IoT as a boundary condition in the relationship between AI and technology anxiety, together with the direct relationship between AI and smart decision-making. The outcome of the current study indicated that the incongruence between interests in AI and users with high internal threats of IoT can impact the extent to which they make use of AI technologies, thus explaining the variation in the users' attitude regarding AI technologies and their smart decision-making. The difference in perspectives between interests in AI systems and users with high internal threats of IoT can be more ardent such that it further strengthens the negative influence of AI on technology anxiety and weakens the positive strength of the direct relationship.

In summary, our findings suggest that based on the levels of internal threats of IoT, users may show different levels of interests in AI technologies and therefore varying levels of AI impact on technology anxiety, and their smart decision-making can be observed. Based on this, this study presents a nuanced explanation by theorizing and presenting evidence that internal threats of IoT act as a boundary condition in the relationship between AI and SDM, therefore significantly extending the existing literature in AI and smart decision-making in smart cities research.

4.2. Practical Implications. The findings of the current research present important practical implications for decision-makers such as government and municipalities from a policy-making standpoint. It is important that government, local authorities, and technology providers create social awareness campaign that can encourage users that artificial intelligence applications, systems, and services respect data privacy and transparent and value individuals' choices. The decision-makers should understand the mental model of users regarding smart services that are not the same, thus must offer advanced personalization options. This can be achieved by simplifying the interfaces to facilitate a condition where smart service users feel no technology anxiety. By doing this, users would be encouraged to adopt smart services to improve their quality of life and thus smart decision-making.

Depending on the levels of internal threats of IoT, users may show different levels of interests in using smart services, and hence, we may experience different levels of AI system impact on technology anxiety and their smart decision-making. Understanding the crucial role of internal threats of IoT is particularly important, as this can help government and local authorities in designing social awareness campaign that is aimed at promoting users' adoption of smart services.

As stated by prior research that the major challenges in adoption and implementation of smart city applications are privacy and security issues [54, 74], it is important for government and local authorities to upgrade and fortify the security system used in protecting smart service. Such action would reassure users they are protected from privacy and security breach.

5. Conclusion

Encouraged by the dynamic field of AI and decision-making in smart cities research, this research article presents important findings to better the comprehension of users' attitudes and perceptions of AI systems for smart decision-making. The rapid growth of IoT has come with some associated threats; a number of these threats are attributed to IoT device vulnerabilities arising from inappropriate use of system resources and cybercrime by hackers. Consistent with the findings of this study, it is important that IoT is constructed in a way that promotes safe and easy usage control. Users need confidence in order to fully embrace AI systems for smart decision-making to enjoy their benefits and avoid privacy and security risks. Therefore, decision-makers in smart cities should take the necessary steps to avoid such threats. This can be achieved by dealing with IoT devices' vulnerabilities through a smooth policy implementation process supported by strong procedures. Such understanding is not just crucial from policy-making perspective but also to promote smart services adoption and sustainability. As AI has gained popularity due to the use of big data, advanced algorithms, and enhanced processing storage and power, AI is increasingly being incorporated into our everyday life and substantially influences SDM. The results presented in this research significantly advance the comprehension of AI technologies and decision-making in both theory and practice [2].

5.1. Limitation and Future Research Direction. The present article offers some limitations that future studies can take advantage of. First, the sample used in the current research was constrained to Turkey; therefore, generalization to other nations is needed in future research. A cross-national survey can also be conducted by adopting the model in this study. Second, longitudinal research design is encouraged to make causal inferences. Third, limited research exists from users' perspective regarding intervening mechanisms in the association between AI systems and decision-making in smart cities research; therefore, future study could benefit by examining the work-life interface in the relationship. Iyiola and Rjoub [75] reported that trust is very crucial in building relationships among parties, and future studies could also benefit by examining the role of trust in the relationship. In particular, when IoT devices in smart cities are connected to the Internet, there is likelihood of attacks on these devices. From this standpoint, Saeed et al. [76] suggested that trust management is an important way to protect data from attacks; trust design models such as scalability, privacy, integrity, reliability, and accuracy associated with security mechanisms for secure communication in IoT devices could be investigated by future studies as possible mechanisms in the relationship between AI and smart decision-making in smart cities. Finally, still from the end users' perspective

future studies could benefit by examining the boundary condition (moderating) role of quality of life by adapting the model in the study.

Data Availability

The data used to support the findings of this study are available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. Duan, J. S. Edwards, and Y. K. Dwivedi, "Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda," *International Journal of Information Management*, vol. 48, pp. 63–71, 2019.

[2] P. Mikalef, K. Lemmer, C. Schaefer et al., "Enabling AI Capabilities in Government Agencies: A Study of Determinants for European Municipalities," *Government Information Quarterly*, vol. 39, Article ID 101596, 2021.

[3] Y. K. Dwivedi, L. Hughes, E. Ismagilova et al., "Artificial Intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," *International Journal of Information Management*, vol. 57, Article ID 101994, 2021.

[4] C. Collins, D. Dennehy, K. Conboy, and P. Mikalef, "Artificial intelligence in information systems research: a systematic literature review and research agenda," *International Journal of Information Management*, vol. 60, Article ID 102383, 2021.

[5] S. Raisch and S. Krakowski, "Artificial intelligence and management: the automation–augmentation paradox," *Academy of Management Review*, vol. 46, no. 1, pp. 192–210, 2021.

[6] F. J. Martínez-López and J. Casillas, "Artificial intelligence-based systems applied in industrial marketing: an historical overview, current and future insights," *Industrial Marketing Management*, vol. 42, no. 4, pp. 489–495, 2013.

[7] J. R. Alam, A. Sajid, R. Talib, and M. Niaz, "A review on the role of big data in business," *International Journal of Computer Science and Mobile Computing*, vol. 3, no. 4, pp. 446–453, 2014.

[8] L. Berntzen, M. R. Johannessen, and R. El-Gazzar, "Smart cities, big data and smart decision-making-understanding "big data" in smart city applications," in *Proceedings of the ICDS 2018, the Twelfth International Conference on Digital Society and eGovernments*, Rome, Italy, March 2018.

[9] A Vesco, *Handbook of research on social, economic, and environmental sustainability in the development of smart cities*, IGI Global, Hershey, Pennsylvania, 2015.

[10] H. F. Atlam, R. J. Walters, and G. B. Wills, "Fog computing and the internet of things: a review," *Big Data and Cognitive Computing*, vol. 2, no. 2, p. 10, 2018.

[11] V. Albino, U. Berardi, and R. M. Dangelico, "Smart cities: definitions, dimensions, performance, and initiatives," *Journal of Urban Technology*, vol. 22, no. 1, pp. 3–21, 2015.

[12] C. Harrison, B. Eckman, R. Hamilton et al., "Foundations for smarter cities," *IBM Journal of Research and Development*, vol. 54, no. 4, pp. 1–16, 2010.

[13] R. Roman, J. Zhou, and J. Lopez, "On the features and challenges of security and privacy in distributed internet of things," *Computer Networks*, vol. 57, no. 10, pp. 2266–2279, 2013.

[14] M. D. Lytras, A. Visvizi, P. K. Chopdar, A. Sarirete, and W. Alhalabi, "Information Management in Smart Cities: turning end users' views into multi-item scale development, validation, and policy-making recommendations," *International Journal of Information Management*, vol. 56, Article ID 102146, 2021.

[15] D. R. Compeau and C. A. Higgins, "Application of social cognitive theory to training for computer skills," *Information Systems Research*, vol. 6, no. 2, pp. 118–143, 1995.

[16] M. L. Meuter, A. L. Ostrom, M. J. Bitner, and R. Roundtree, "The influence of technology anxiety on consumer use and experiences with self-service technologies," *Journal of Business Research*, vol. 56, no. 11, pp. 899–906, 2003.

[17] S. A. A. Bokhari and S. Myeong, "Use of artificial intelligence in smart cities for smart decision-making: a social innovation perspective," *Sustainability*, vol. 14, no. 2, p. 620, 2022.

[18] P. Mikalef, S. O. Fjørtoft, and H. Y. Torvatn, "Artificial Intelligence in the public sector: a study of challenges and opportunities for Norwegian municipalities," in *Conference on E-Business, E-Services and E-Society*, pp. 267–277, Springer, Cham, Berlin, Germany, 2019a.

[19] A. Kankanhalli, Y. Charalabidis, and S. Mellouli, "IoT and AI for smart government: a research agenda," *Government Information Quarterly*, vol. 36, no. 2, pp. 304–309, 2019.

[20] M. M. Young, J. B. Bullock, and J. D. Lecy, "Artificial discretion as a tool of governance: a framework for understanding the impact of artificial intelligence on public administration," *Perspectives on Public Management and Governance*, vol. 2, no. 4, pp. 301–313, 2019.

[21] H. J. Scholl and S. AlAwadhi, "Creating smart governance: the key to radical ICT overhaul at the city of munich," *Information Polity*, vol. 21, no. 1, pp. 21–42, 2016.

[22] A. Meijer and M. Thaens, "Quantified street: smart governance of urban safety," *Information Polity*, vol. 23, no. 1, pp. 29–41, 2018.

[23] Y. J. Park, Y. J. Choe, O. Park et al., "Contact tracing during coronavirus disease outbreak, South Korea, 2020," *Emerging Infectious Diseases*, vol. 26, no. 10, pp. 2465–2468, 2020.

[24] L. Qiao, Y. Li, D. Chen, S. Serikawa, M. Guizani, and Z. Lv, "A survey on 5G/6G, AI, and Robotics," *Computers & Electrical Engineering*, vol. 95, Article ID 107372, 2021.

[25] H. Makino, K. Tamada, K. Sakai, and S. Kamijo, "Solutions for urban traffic issues by ITS technologies," *IATSS Research*, vol. 42, no. 2, pp. 49–60, 2018.

[26] Y. Ma, K. Ping, C. Wu, L. Chen, H. Shi, and D. Chong, "Artificial Intelligence Powered Internet of Things and Smart Public Service," *Library Hi Tech*, vol. 38, 2019.

[27] C. H. Lien and Y. Cao, "Examining WeChat users' motivations, trust, attitudes, and positive word-of-mouth: evidence from China," *Computers in Human Behavior*, vol. 41, pp. 104–111, 2014.

[28] R. Giffinger and H. Kramar, *Benchmarking, Profiling and Ranking of Cities: The 'European Smart Cities' Approach. Place-Based Performance Metrics In Building Sustainable Cities*, Routledge, Oxfordshire, UK, 2021.

[29] A. McGovern, K. L. Elmore, D. J. Gagne et al., "Using artificial intelligence to improve real-time decision-making for high-impact weather," *Bulletin of the American Meteorological Society*, vol. 98, no. 10, pp. 2073–2090, 2017.

[30] B. Marr, *Big Data: Using SMART Big Data, Analytics and Metrics to Make Better Decisions and Improve Performance*, John Wiley & Sons, Hoboken, NJ USA, 2015.

[31] A. Bennaceur, T. T. Tun, A. K. Bandara, Y. Yu, and B. Nuseibeh, "Feature-driven mediator synthesis: supporting collaborative security in the internet of things," *ACM Transactions on Cyber-Physical Systems*, vol. 2, no. 3, pp. 1–25, 2018.

[32] Z. Allam and Z. A. Dhunny, "On big data, artificial intelligence and smart cities," *Cities*, vol. 89, pp. 80–91, 2019.

[33] D. G. Johnson and M. Verdicchio, "AI anxiety," *Journal of the Association for Information Science and Technology*, vol. 68, no. 9, pp. 2267–2270, 2017.

[34] D. L. Waltz, "Evolution, sociobiology, and the future of artificial intelligence," *IEEE Intelligent Systems*, vol. 21, no. 3, pp. 66–69, 2006.

[35] S. R. Granter, A. H. Beck, and D. J. Papke Jr, "AlphaGo, deep learning, and the future of the human microscopist," *Archives of Pathology & Laboratory Medicine*, vol. 141, no. 5, pp. 619–621, 2017.

[36] O. Vinyals, I. Babuschkin, W. M. Czarnecki et al., "Grandmaster level in StarCraft II using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, 2019.

[37] McKinsey Global Institute, J. Lost, and J. Gained, "Workforce Transitions in a Time of Automation," Retrieved from https://www.mckinsey.com/global-themes/future-of-organizations-and-work/what-the-future-of-work-will-mean-for-jobs-skills-and-wages, 2017.

[38] M. U. Scherer, "Regulating artificial intelligence systems: risks, challenges, competencies, and strategies," *Harv. JL & Tech.*vol. 29, p. 353, 2015.

[39] D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, and D. Mané, "Concrete Problems in AI Safety," 2016, https://arxiv.org/abs/1606.06565.

[40] E. Ismagilova, L. Hughes, Y. K. Dwivedi, and K. R. Raman, "Smart cities: advances in research—an information systems perspective," *International Journal of Information Management*, vol. 47, pp. 88–100, 2019.

[41] G. E. Kreiner, E. C. Hollensbe, and M. L. Sheep, "Balancing borders and bridges: negotiating the work-home interface via boundary work tactics," *Academy of Management Journal*, vol. 52, no. 4, pp. 704–730, 2009.

[42] U. Köse, "Are we safe enough in the future of artificial intelligence? A discussion on machine ethics and artificial intelligence safety," *BRAIN: Broad Research in Artificial Intelligence and Neuroscience*, vol. 9, no. 2, pp. 184–197, 2018.

[43] C. Bremser, G. Piller, and F. Rothlauf, "How to start with big data-a multiple case study," in *Proceedings of the Bled eConference*, p. 40, Bled, Slovenia, June 2017.

[44] J. W. Lian, D. C. Yen, and Y. T. Wang, "An exploratory study to understand the critical factors affecting the decision to adopt cloud computing in Taiwan hospital," *International Journal of Information Management*, vol. 34, no. 1, pp. 28–36, 2014.

[45] C. Schaefer, K. Lemmer, K. Samy Kret, M. Ylinen, P. Mikalef, and B. Niehaves, "Truth or dare?–how can we influence the adoption of artificial intelligence in municipalities?" in *Proceedings of the 54th Hawaii International Conference On System Sciences*, p. 2347, Grand Wailea, Maui, Hawaii, January 2021.

[46] L. Muehlhauser and N. Bostrom, "Why we need friendly AI," *Thinking*, vol. 13, no. 36, pp. 41–47, 2014.

[47] R. Clarke, "Why the world wants controls over Artificial Intelligence," *Computer Law & Security Report*, vol. 35, no. 4, pp. 423–433, 2019.

[48] T. D. Little, N. A. Card, J. A. Bovaird, K. J. Preacher, and C. S. Crandall, "Structural equation modeling of mediation and moderation with contextual factors," *Modeling contextual effects in longitudinal studies*, vol. 1, pp. 207–230, 2007.

[49] N. Ahmed, R. Amin, H. Aldabbas, D. Koundal, B. Alouffi, and T. Shah, "Machine Learning Techniques for Spam Detection in Email and IoT Platforms: Analysis and Research Challenges," *Security and Communication Networks*, vol. 2022, Article ID 1862888, 2022.

[50] Q. Zhu, S. W. Loke, R. Trujillo-Rasua, F. Jiang, and Y. Xiang, "Applications of distributed ledger technologies to the internet of things: a survey," *ACM Computing Surveys*, vol. 52, no. 6, pp. 1–34, 2020.

[51] G. Madhu, B. Lalith Bharadwaj, R. Boddeda et al., "Intelligent disease diagnosis model for energy aware cluster based IoT healthcare systems," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1189–1203, 2022.

[52] W. Salehi, G. Gupta, S. Bhatia, D. Koundal, A. Mashat, and A. Belay, "IoT-Based Wearable Devices for Patients Suffering from Alzheimer Disease," *Contrast Media & Molecular Imaging*, vol. 2022, Article ID 3224939, 2022.

[53] A. Sestino, M. I. Prete, L. Piper, and G. Guido, "Internet of Things and Big Data as enablers for business digitalization strategies," *Technovation*, vol. 98, Article ID 102173, 2020.

[54] Z. Lv, L. Qiao, A. Kumar Singh, and Q. Wang, "AI-empowered IoT security for smart cities," *ACM Transactions on Internet Technology*, vol. 21, no. 4, pp. 1–21, 2021.

[55] T. F. Kummer, J. Recker, and M. Bick, "Technology-induced anxiety: manifestations, cultural influences, and its effect on the adoption of sensor-based technology in German and Australian hospitals," *Information & Management*, vol. 54, no. 1, pp. 73–89, 2017.

[56] A. L. Powell, "Computer anxiety: comparison of research from the 1990s and 2000s," *Computers in Human Behavior*, vol. 29, no. 6, pp. 2337–2381, 2013.

[57] S. Babar, P. Mahalle, A. Stango, N. Prasad, and R. Prasad, "Proposed security model and threat taxonomy for the Internet of Things (IoT)," in *Proceedings of the International Conference on Network Security And Applications*, pp. 420–429, Springer, Berlin, Heidelberg, July 2010.

[58] N. Saif Almuraqab and S. Jasimuddin, "Factors that influence end-users' adoption of smart government services in the UAE: aconceptual framework," *Electronic Journal of Information Systems Evaluation*, vol. 20, pp. 11–23, 2017.

[59] C. Graham, "Fear of the unknown with healthcare IoT devices: an exploratory study," *Information Security Journal: A Global Perspective*, vol. 30, no. 2, pp. 100–110, 2021.

[60] V. C. Müller and N. Bostrom, "Future progress in artificial intelligence: a survey of expert opinion," in *Fundamental Issues of Artificial Intelligence*, pp. 555–572, Springer, Cham, Berlin Germany, 2016.

[61] Y. P. Ongena, M. Haan, D. Yakar, and T. C. Kwee, "Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire," *European Radiology*, vol. 30, no. 2, pp. 1033–1040, 2020.

[62] M. Abomhara and G. M. Køien, "Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks," *Journal of Cyber Security and Mobility*, vol. 4, no. 1, pp. 65–88, 2015.

[63] H. Klug and A. Kmoch, "Operationalizing environmental indicators for real time multi-purpose decision making and action support," *Ecological Modelling*, vol. 295, pp. 66–74, 2015.

[64] A. F. Hayes and M. Scharkow, "The relative trustworthiness of inferential tests of the indirect effect in statistical mediation

Retraction

Retracted: Retinex Algorithm and Mathematical Methods Based Texture Detail Enhancement Method for Panoramic Images

Mathematical Problems in Engineering

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Kang, "Retinex Algorithm and Mathematical Methods Based Texture Detail Enhancement Method for Panoramic Images," *Mathematical Problems in Engineering*, vol. 2022, Article ID 6490393, 8 pages, 2022.

Hindawi

Research Article

Retinex Algorithm and Mathematical Methods Based Texture Detail Enhancement Method for Panoramic Images

Yingxi Kang ⓘ

College of Electrical and Information Engineering, Hunan Institute of Engineering, Xiangtan 411104, China

Correspondence should be addressed to Yingxi Kang; 19404276@masu.edu.cn

A panoramic image texture detail enhancement method based on the Retinex algorithm is proposed to work on the nature of panoramic images. Firstly, the panoramic image is collected, then the panoramic image is preprocessed through brightness enhancement, the brightness of the preprocessed panoramic image is normalized, and the panoramic image is optimized and increased by the improved Retinex algorithm. Finally, the simulation test of the panoramic image was carried out. The results show that the improved Retinex algorithm works on the sign to-commotion proportion of the panoramic image. Furthermore, the time required to enhance the panoramic texture detail is short, which can meet the practical requirements of panoramic image subsequent processing.

1. Introduction

Panoramic image displays portray however much of the general climate as could be expected through wide-point articulation, painting, photographs, recordings, three-layered models, etc., 360 display, or at least, catching the image data of the whole scene with an expert camera or delivering the image with demonstrating programming, utilizing the product to gather the image and playing it with a unique player, that is to say, changing the plane photograph or PC displaying the picture into a 360 degree display for computer generated reality perusing and recreating the two-layered plan into a genuine three-layered space and introducing it to the watcher. It uses a wide-angle means of expression, a new way to show the surrounding world from a new perspective and can show all the surrounding scenery as much as possible, so it has attracted more and more attention [1].

Compared with general renderings and 3D animation, the panorama has the following advantages: it avoids the effect that the general plane effect drawing has a single perspective and can not bring an all-round feeling. The picture effect is exactly the same as the general effect drawing when playing on this machine; it is highly interactive and can be manipulated by customers to interactively observe the scene from any angle, just like immersive and truly feel the final design result, which is also different from the lack of interactive 3D animation; the price is only slightly higher than the general effect drawing. Compared with the three-dimensional animation of hundreds of yuan per second, it is economical and short production cycle. All directions: all the scenes within the 360 degree sphere are displayed in an all-round way; in the example, you can press and drag with the left mouse button to view all directions of the scene; scene: real scene, real scene. Most of the three-dimensional real scenes are images assembled on the basis of photos, which retain the authenticity of the scene to the greatest extent; 360 : 360 degree look around effect. Although the photos are planar, the 360 real scene obtained after software processing can give people a three-dimensional spatial feeling, making the viewer feel like he is in it. Notwithstanding, the nature of panoramic images gathered under unfortunate lighting conditions, such as night and shadow, degrades, primarily in the aspects of uneven overall brightness, low contrast, and dark color of the image, which genuinely affects the enhanced visualization of panoramic images and acquaints a few troubles with

ensuing PC vision handling undertakings (for example, image division, target following, target acknowledgment, etc.) [2, 3]. As a result, research into the low illumination panoramic texture detail enhancement algorithm is critical for the field of machine vision.

The study of [4] proposed a fractal image texture detail enhancement method based on Newton iterative algorithm; Newton iterative algorithm is used to obtain the extreme value of fractal image coordinates and set the iterative initial point value. The rotation transformation of the image around the origin 120° completes the zero point attraction domain mapping; the RGB color channel expression is constructed by using iterative distance parameters, initial parameters, initial color value, color gradient parameters, and relevant operation symbols, and the number of meshes covering the target part of the image is calculated. The relationship between compression mapping and compression factor is obtained by calculating shrinkage affine transformation, and the texture detail enhancement is completed by iterative calculation. The study of [5] separates caries-related highlights from a computerized display got by vatech 400 gear utilizing image texture examination in view of the dim level coevent framework (GLCM) algorithm. The advanced image is first placed into the PC, then changed over completely to grayscale, the extraction cycle is handled and denoised, the measurable highlights of the GLCM lattice are extricated, and lastly the best elements are picked. Image enhancement is acknowledged, therefore. A powerful panoramic image delivering and smooth progress algorithm for contiguous perspectives was proposed in reference [6]. The element points of contiguous perspective images are extricated first, then the nearby point matches are laid out utilizing a powerful matching algorithm, and the matching triangle is shaped utilizing a similar name focuses. Then, at that point, a unique change model is made by at the same time controlling the shape and texture of every triangle. Finally, the dynamic transition model is superimposed on top of the 360-degree panoramic roaming scene to achieve a smooth transition between adjacent viewpoints. The consequences of the analyses show that this technique enjoys clear benefits in visual execution and visual movement.

The panoramic image can show all the scenes within a 360 degree sphere as a new way to show the new perspective of the surrounding world. In the process of panoramic image acquisition, because of the weather's influence, illumination, and other factors, the quality of the acquired panoramic image becomes worse, which is difficult to meet people's visual requirements and brings difficulties to the subsequent panoramic image processing operations. The existing methods mainly achieve the purpose of image enhancement through modern digital technology. There are two main purposes for panoramic texture detail enhancement: one is to improve the clarity of the image [7], and the other is to work on the special visualization of the image, so as to make interesting information or useful image detail information in the image more prominent [8]. In order to achieve a better image visual effect, a panoramic image texture detail enhancement method based on the Retinex algorithm is

proposed, and its performance is tested by comparative experiments.

2. Panoramic Texture Detail Enhancement Method Based on Retinex

In terrible circumstances, for example, climate and light, the brilliance and differentiation of the gathered panoramic image are generally low. To work on the general brilliance of the low-light panoramic image, make the detailed information of the object in the image clearer and make it a good visual effect, the research on high-quality image enhancement algorithm has attracted the attention of many scholars at home and abroad, it has as of late turned into a hotly debated issue in the field of image handling. Retinex theory was first proposed by land, also known as color constancy theory, which is widely used in image enhancement. In recent years, based on the theory of the Retinex algorithm, many improved Retinex algorithms have been proposed by scholars at home and abroad. However, a large number of experimental results show that the Retinex theory still has the defects of "halo artifact," image over enhancement, and low algorithm efficiency. In order to further make the Retinex algorithm have a better image enhancement effect and better retain the object detail information in the image, aiming at the shortcomings of the current Retinex algorithm, a panoramic texture detail enhancement method based on improved Retinex is proposed in this paper.

2.1. Panoramic Texture Detail Enhancement Process. The illumination intensity has a great impact on the panoramic image. Strong and weak illumination intensity will reduce the accuracy of image environment information [9]. The flow chart of panoramic texture detail enhancement under different illumination intensity is shown in Figure 1. The CIEXYZ color space method is used to identify the light power information in the image. After the recognition of the light power information in the image is completed, the low illumination part and high illumination part in the image are normalized [10], so as to solve the influence of the change of light power on the panoramic image when the light intensity is insufficient or too strong. The normalized panoramic image is enhanced by the improved Retinex algorithm.

2.2. Image Brightness Enhancement. Although noise removal can effectively improve image quality, multisource data are information transmitted in different ways, which will be limited by terminal hardware and reduce its own contrast. Therefore, this paper uses the multisource big data analysis method to analyze the color unevenness of small feature points in the image and enhance the brightness of the panoramic image on the basis of single-scale Retinex [11] for followup work.

In general, a panoramic image can be described as follows:

$$S(x, y) = L(x, y) \times R(x, y), \tag{1}$$

where $S(x, y)$ represents the initial image, $L(x, y)$ represents the data image after adaptive median filtering [12], and
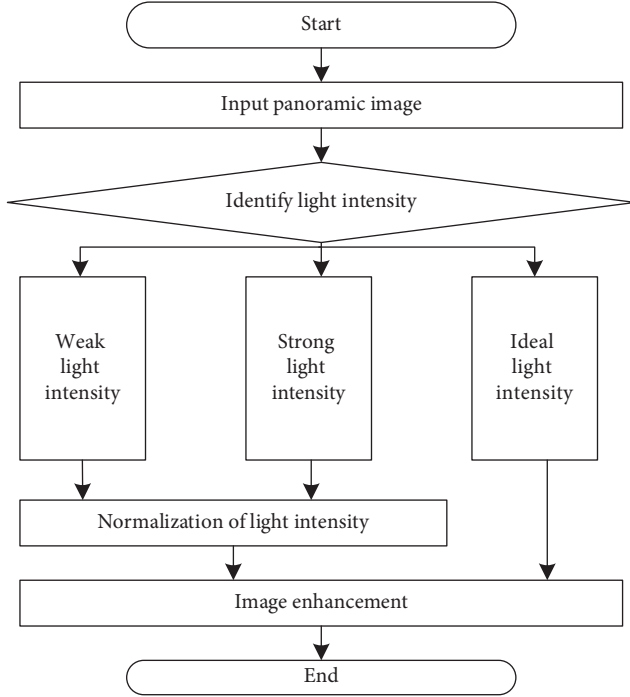
Figure 1: Flow chart of panoramic texture detail enhancement.

$R(x, y)$ represents the reflection image. Retinex can be obtained by merging and converting the above-given three images into the corresponding number field:

$$lbR(x, y) = lbS(x, y) - lb[F(x, y) \otimes S(x, y)], \quad (2)$$

where lb is the conversion coefficient and $F(x, y)$ is the center surround function. The center surround function can be further obtained according to (2):

$$F(x, y) = \lambda \exp\left[-\frac{(x^2 + y^2)}{c^2}\right]. \quad (3)$$

where c represents the scale of the surrounding function, and λ represents a scale in the scale of the surrounding function.

According to the above-given results, the Gaussian model [13, 14] is substituted into the traditional model to obtain the average brightness, enhance the panoramic image and improve the image edge feature comparison. According to the model, the color space conversion can be completed to obtain the brightness component I:

$$I(x, y) = \max\{I_R(x, y), I_G(x, y), I_B(x, y)\}, \quad (4)$$

where (x, y) represents the color component in the panoramic image, and $I_R(x, y)$, $I_G(x, y)$, and $I_B(x, y)$ address the RGB (red, green, and blue) color component. The average brightness $\overline{I}(x, y)$ can be further obtained according to bilateral filtering:

$$\overline{I}(x, y) = G_r(x, y)G_u \lg \frac{1 + I(x_i, y_i)}{\sum_{i=1}^{m} G_r(x, y)G_u} \quad (5)$$

where m represents the color of a region in the image, $G_u(x, y)$ represents the traditional Gaussian kernel function,

and $G_r(x, y)$ represents the kernel function close to the Gaussian state in the color space [15], which can be obtained:

$$\begin{aligned} G_r(x, y) = {} & \lambda_1 \exp\left\{-\left(x^2 + y^2\right)/\left(2\sigma_1^2\right)\right\} \\ & - \lambda_2 \exp\left\{-\left(x^2 + y^2\right)/\left(2\sigma_2^2\right)\right\} \\ & + \lambda_3 \exp\left\{-\frac{\left(x^2 + y^2\right)}{\left(2\sigma_3^2\right)}\right\}, \end{aligned} \quad (6)$$

where r represents the radius distance of the calculated filter, λ_1 and λ_2 represent the peak value coefficients around the middle and center points, respectively, λ_3 represents the edge peak value, and σ_1, σ_2, σ_3 represents the scale coefficient.

During the enhancement process, the scale value and quantity can be related to the final result [16]. Assuming that the surround function and the center function only enhance one-sided areas and do not enhance all pixel areas, the result will be unsatisfactory. The gain constant is the standard constant in the formula. The value is calculated after the fusion of the center function and the surround function, which can avoid one-sided enhancement. In order to ensure a uniform luminance pixel value, the constant a is substituted into the kernel function in the calculation to obtain the optimal value. The enhanced image $R_j(x, y)$ is

$$R_j(x, y) = \lg\left[a_j \times \frac{I(x, y)}{F_j(x, y) \times I(x, y)}\right], \quad j = 1, 2, 3. \quad (7)$$

The enhanced luminance component I is expressed as follows:

$$I(x, y) = \frac{1}{3}\left\{a_1 a_2 a_3 \frac{I(x, y)^3}{G_1(x, y) \times G_2(x, y) \times G_3(x, y)}\right\}, \quad (8)$$

where a_1, a_2, a_3 represents the gain constant with difference.

2.3. Brightness Normalization. For the image whose recognition result is that the illumination intensity is too high or too low, the panoramic image can be equalized by the illumination normalization method [17, 18], which is convenient for the subsequent texture detail enhancement.

(1) Gamma correction. Traditional gamma correction usually changes the brightness of the input image, and it is difficult to achieve a good enhancement effect in both too dark and too bright areas at the same time. The bilateral gamma correction curve designed in this paper can improve the visual effect of over dark area and over the bright area in the input image at the same time. Multiscale image decomposition can make the image brightness and detail processed separately. The image's brightness information is primarily concentrated in low frequency, while detail information is primarily concentrated in high frequency. Based on the above-given two points, this paper realizes image gamma correction by editing the gamma curve of the panoramic image [19, 20], edits the nonlinear tone of the image,

separates the light color part and dark part in the image signal, and uses gamma correction to compress and expand the high gray value and low gray value of the image, respectively, to improve the image contrast, reduce the influence of illumination intensity on panoramic image quality. The gamma correction formula is as follows:

$$I_{\text{out}} = \left(\frac{I_{\text{in}}}{\max \text{gray}}\right)^{(1/\lambda)} \max \text{gray}, \qquad (9)$$

where I_{in} and I_{out}, respectively, represent the original gray value and the corrected gray value of the input panoramic image; maxgray and λ represent the maximum gray value and correction coefficient of the input image, respectively.

(2) Gaussian differential filtering. Gamma correction is used to improve the contrast of the panoramic image, and the shadow area of the image edge still needs further processing. The edge of the panoramic image is extracted by band-pass filter and Gaussian difference filter [21, 22], so as to reduce the influence of illumination intensity on the edge region of the panoramic image. The panoramic image $I(x, y)$ is filtered by a Gaussian difference filter, also, the equation is as per the following:

$$G(x, y) = \frac{1}{\sqrt{2\pi}\varphi_1} \exp\left[-\frac{x^2 + y^2}{2\varphi_1^2}\right] - \frac{1}{\sqrt{2\pi}\varphi_2} \exp\left[-\frac{x^2 + y^2}{2\varphi_2^2}\right], \qquad (10)$$

where φ_1 and φ_2 are Gaussian differential filter coefficients, and $\varphi_1 < \varphi_2$.

$$g(x, y) = G(x, y) * I(x, y). \qquad (11)$$

Formula (10) and formula (11) are Gaussian difference filter and filtering processing results, respectively.

2.4. Improved Retinex Texture Detail Enhancement Algorithm. For the normalized image, the improved Retinex algorithm is used to enhance the texture details. The panoramic image is down sampled by Gaussian pyramid, and the sampling is completed by even line sampling and convolution sampling. The sampling formula is as follows:

$$G_{h+1}(i, j) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} R(m, n) G_h(2i - m, 2j - n), \qquad (12)$$

where m addresses the quantity of lines, n addresses the quantity of sections, and h and $R(m, n)$ represent pyramid algebra and Gaussian convolution kernel respectively.

The Gaussian weight is obtained by pixel difference and spatial distance. The points with a large difference in peripheral pixel values and the center point is the minimum or maximum are noise points. The noise points are processed by reducing the center weight. Filter the x-axis direction and y-axis direction of the image, respectively and obtain the filter function as follows:

$$l(x) = z^{-1}(x)(o(\psi, x) + s(f(\psi), f(x)))d\psi, \qquad (13)$$

where ψ and $z(x)$, respectively, represent the sum of spatial pixels and the weight normalized to the result, and o and s, respectively, represent the distance function and similarity function.

After bilateral filtering, the image contrast is compressed to avoid image degradation [23]. The bicubic difference algorithm is used to restore the image, obtain accurate interpolation graphics through the four neighborhood pixel information of the image, and enhance the image magnification effect. The bicubic difference formula is as follows:

$$D(X, Y) = \sum_{i=0}^{3} \sum_{j=0}^{3} w_{ij} W(i) W(j), \qquad (14)$$

where w_{ij} represents the weight coefficient; $W(i)$ and $W(j)$ represent abscissa weight and ordinate weight, respectively.

After the bicubic difference processing is completed, the handled image is deducted from the first image in logarithmic space, and the color constancy and detail enhancement of the image are realized based on the constant scale.

3. Simulation Experiment

In order to test the texture detail enhancement method of panoramic image based on the Retinex algorithm studied in this paper and optimize the effectiveness of panoramic image, 10 panoramic images under different illumination were selected from a website, and the enhancement effects of multiple low illumination panoramic images are compared on the MATLAB experimental platform (CPU Intel (R) core (TM) i5-2320 3.0 GHz); The effectiveness of texture details is enhanced through subjective visual effect and objective quality evaluation.

In order to verify the advantages of texture detail enhancement of this method, Newton iterative algorithm (reference [4]) and GLCM algorithm (reference [5]) are selected as comparison methods. Brightness, information entropy, clarity, mean square error (MSE), and peak signal-to-noise ratio (PNSR) are used as objective test indicators. The specific evaluation indicators are shown in Table 1.

The parameters for setting the target image are shown in Table 2.

According to the above parameter settings, the texture detail enhancement simulation is carried out. The original image is shown in Figure 2.

Taking the image of Figure 2 as the research sample, the panoramic image texture detail enhancement processing is carried out. Through the comparative test of the methods in reference [4] and reference [5], the panoramic image texture detail enhancement results are obtained, as shown in Figure 3. Figure 3 shows that the output quality of panoramic image texture detail enhancement by this method is better than the traditional method, the definition is higher, and the image's brightness and contrast are more balanced.

The results of detailed visual communication are shown in Figure 4. According to the analysis of Figure 4, the

TABLE 1: Evaluation index content.

Objective evaluation index	Index content
Brightness	The brightness of the picture, an image gives people an intuitive feeling.
Information entropy	It addresses the typical measure of data in the image. The higher the entropy, the more information there is.
Definition	It not just mirrors the image's little detail difference and texture change qualities, but it also reflects the image's clarity. The greater the size, the better.
Mean square error (MSE)	The greatness of the mean square worth of the pixel contrast between the first image and the contorted image decides the bending level of the twisted image.
Peak signal to noise ratio (PNSR)	It is typically used to compare the quality of an image after compression to the original image. The lower the distortion after compression, the higher the PSNR.

TABLE 2: Experimental parameters.

Parameter	Numerical value
Light intensity	125 dB
Intensity of image pixels	249×350
Resolution of image spatial sampling	360×360
Feature matching coefficient	0.21
Pixel spatial gain	0.02

FIGURE 2: Panoramic image.
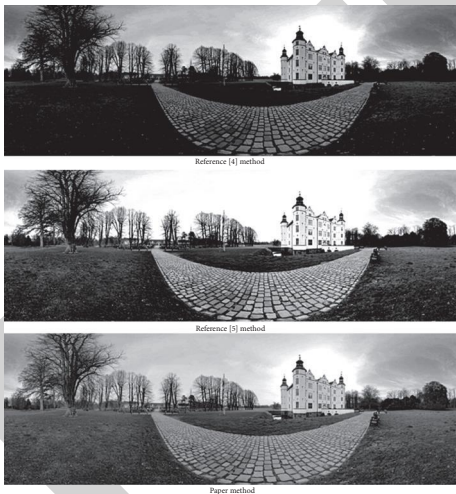
FIGURE 3: Comparison of texture detail enhancement results of panoramic image.

panoramic image texture detail enhancement method in this paper has good visual communication ability and high output quality of image imaging. The brighter parts in the original image still maintain rich texture details, while the parts with weak reflected light at the rear of the wheel are compensated by illumination, and the contrast is also improved.

Based on the above-given experiments, the fitness of the three methods after enhancement in Figure 4 is compared, as shown in Figure 5.

As can be seen from Figure 5, the image enhanced by the method in reference [5] has the slowest convergence speed and the lowest fitness, and the average fitness fluctuation is relatively gentle, but the fitness convergence is limited; The average fitness convergence speed of the image enhanced by the method in reference [4] is higher than that of the method in reference [5], but it is still slower than that of the method in this paper. The average fitness curve fluctuates greatly and is in an extremely unstable state; the average fitness of this strategy is significantly greater than reference [4] and reference [5], and the fluctuation of average fitness is relatively flat.

The image quality of panoramic texture details enhanced by three methods is detected from the three indexes of image brightness, information entropy, and definition. The statistical results are shown in Table 3. As per the comparison results in Table 3, the brightness of the panoramic image optimized by this method is between 120 cd/M and 130 cd/m, the information entropy is higher than 7.4 bit and the definition is higher than 5; The brightness of the panoramic image optimized by the comparison method is between 90 cd/M and 140 cd/m, the information entropy is less than 7 bit and the definition is less than 5. The comparison results show that the brightness of the panoramic image optimized by this method is very uniform, there is no high or low brightness, has high information entropy and clarity, and the optimized image quality is high. It is verified that this method has a high optimization effect [18, 24, 25].

Taking 10 min as a unit time, the changes of the number of texture details in the panoramic image are recorded in 6 unit times respectively. After being optimized by three groups of algorithms.

Table 4 shows that as the experimental time is increased, the number of texture details in this method maintained a significant upward trend in the first two unit times, maintained a small upward trend from the third unit time, and continued to remain stable in the fifth and sixth unit times, with the maximum value of the full name reaching 10.3×1011T; the number of texture details in reference [4] and reference [5] methods always keeps the change trend of rising and falling alternately, the phased maximum value continues to decrease, the phased minimum value basically remains unchanged, and the global maximum value reaches 8.9×1011t and 7.6×1011t, much lower than the method in this paper. To sum up, the application of visual image antiocclusion moving target

Figure 4: Visual communication results of texture detail enhancement.
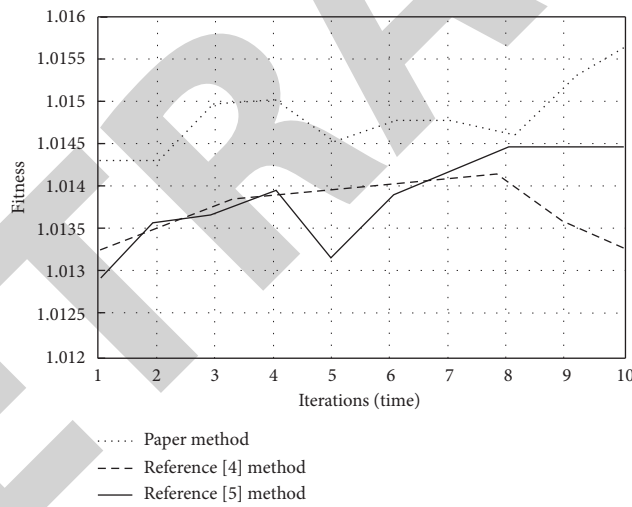
Figure 5: Fitness comparison chart.

Table 3: Comparison of image quality after optimization by different methods.

Image serial number	Paper method			Reference [4] method			Reference [5] method		
	Brightness	Information entropy	Definition	Brightness	Information entropy	Definition	Brightness	Information entropy	Definition
1	128.6	7.8	5.8	90.5	6.5	2.8	91.5	5.6	3.6
2	125.6	7.4	5.6	135.6	6.8	3.5	94.5	5.8	4.5
3	127.5	7.5	5.7	105.6	6.7	3.4	115.4	5.7	3.8
4	123.5	7.6	5.6	98.5	6.4	3.6	139.5	5.6	3.9
5	124.5	7.5	5.9	138.5	6.2	3.5	135.4	5.8	3.7
6	129.5	7.4	5.4	139.5	5.8	2.9	107.4	5.7	3.5
7	124.5	7.5	5.2	110.4	6.4	3.4	106.8	5.9	4.8
8	123.5	7.6	5.9	98.5	6.5	3.5	105.4	6.1	4.1
9	122.4	7.8	5.4	97.5	6.1	4.7	96.5	6.5	3.5
10	126.5	7.5	5.6	96.5	5.9	4.1	94.5	6.4	3.9

TABLE 4: Comparison of texture details per unit time.

Paper method			Reference [4] method			Reference [5] method		
Unit time group	Experimental time/(min)	Number of texture details/ (×10^{11}T)	Unit time group	Experimental time/(min)	Number of texture details/ (×10^{11}T)	Unit time group	Experimental time/(min)	Number of texture details/ (×10^{11}T)
1	5	6.9	1	5	7.6	1	5	6.3
	10	8.5		10	8.9		10	7.6
2	15	9.7	2	15	7.7	2	15	6.4
	20	9.8		20	8.8		20	7.5
3	25	9.9	3	25	7.6	3	25	6.3
	30	10		30	8.7		30	7.4
4	35	10.1	4	35	7.6	4	35	6.3
	40	10.2		40	8.6		40	7.3
5	45	10.3	5	45	7.6	5	45	6.3
	50	10.3		50	8.5		50	7.2
6	55	10.3	6	55	7.7	6	55	6.4
	60	10.3		60	8.4		60	7.1

tracking algorithm in augmented reality can ensure a stable increase in the number of texture details per unit time.

To sum up, this method uses the multi-scale decomposition method to construct the patch visual information detection model of the panoramic image, so as to realize the visual communication and information enhancement of the panoramic image. Compared with the traditional methods, this method improves the detection and recognition rate, contrast, brightness, illumination compensation, and detail enhancement of the panoramic image. It can produce an enhanced image with a good visual effect and improve the processing ability of texture detail enhancement of panoramic image [26–28].

4. Conclusion

Image enhancement innovation is one of the essential assignments of present day PC vision and advanced image handling. The implementation of this operation is to further improve the human sensory ability to observe images, better analyze and process the contents contained therein and enhance the readability of images. Retinex is an image enhancement theory based on scientific testing and scientific analysis. It simulates the way the human visual system perceives brightness and color under differentiated lighting conditions. Retinex theory holds that the reflection ability of an object to light determines the color of the object. Its essence is to obtain the reflection component of the object in the image without the illumination component, to reestablish the first appearance of the item. In contrast with other customary image enhancement strategies, the image enhancement model and algorithm in light of the Retinex hypothesis work on the image's edge, constant color, and large dynamic compression range, so that the image can obtain the maximum color stability without change and distortion, also, genuinely reestablish the first appearance of the article in the image.

The force of the light essentially affects the delivering impact of a panoramic image. The panoramic image enhancement under different illumination intensities is studied, the illumination intensity of the panoramic image is recognized, the illumination normalization is implemented for the images with high and low illumination intensity, and the improved Retinex algorithm is utilized to understand the panoramic image enhancement handling and complete the image enhancement. The exploratory outcomes demonstrate the way that the examination technique can streamline top notch images in a short measure of time. The optimized image can retain more image details and has a high image enhancement effect. The image is optimized because it has a high peak signal-to-noise ratio and a low image mean square blunder, which can meet the actual needs of panoramic image enhancement.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

References

[1] O. Kulce and L. Onural, "Generation of a polarized optical field from a given scalar field for wide-viewing-angle holographic displays," Optics and Lasers in Engineering, vol. 137, 2021.

[2] Z. D. Liu, L. N. Zhang, F. Ding, and H. H. Cao, "Two-phase flow texture feature extraction of digital image in immersion environment," Computer Simulation, vol. 38, no. 6, pp. 133–137, 2021.

[3] B. Cheng, T. Sun, J. Cheng, and T. Liu, "Multi-camera panoramic imaging system based on adaptive brightness

Research Article

Mathematical Methods for IoT-Based Annotating Object Datasets with Bounding Boxes

Abdelhamid Zaidi [ID]

Department of Mathematics, College of Science, Qassim University, P. O. Box 6644, Buraydah 51452, Saudi Arabia

Correspondence should be addressed to Abdelhamid Zaidi; a.zaidi@qu.edu.sa

Object datasets used in the construction of object detectors are typically annotated with horizontal or oriented bounding rectangles for IoT-based. The optimality of an annotation is obtained by fulfilling two conditions: (i) the rectangle covers the whole object and (ii) the area of the rectangle is minimal. Building a large-scale object dataset requires annotators with equal manual dexterity to carry out this tedious work. When an object is horizontal for IoT-based, it is easy for the annotator to reach the optimal bounding box within a reasonable time. However, if the object is oriented, the annotator needs additional time to decide whether the object will be annotated with a horizontal rectangle or an oriented rectangle for IoT-based. Moreover, in both cases, the final decision is not based on any objective argument, and the annotation is generally not optimal. In this study, we propose a new method of annotation by rectangles for IoT-based, called robust semi-automatic annotation, which combines speed and robustness. Our method has two phases. The first phase consists in inviting the annotator to click on the most relevant points located on the contour of the object. The outputs of the first phase are used by an algorithm to determine a rectangle enclosing these points. To carry out the second phase, we develop an algorithm called RANGE-MBR, which determines, from the selected points on the contour of the object, a rectangle enclosing these points in a linear time. The rectangle returned by RANGE-MBR always satisfies optimality condition (i). We prove that the optimality condition (ii) is always satisfied for objects with isotropic shapes. For objects with anisotropic shapes, we study the optimality condition (ii) by simulations. We show that the rectangle returned by RANGE-MBR is quasi-optimal for the condition (ii) and that its performance increases with dilated objects, which is the case for most of the objects appearing on images collected by aerial photography.

1. Introduction

The construction of an object detector generally goes through a learning phase, followed by a testing phase, and ends with a tuning phase. Each phase requires an independent annotated object dataset. Annotating an image IoT-based signal and image processing applications consists of locating all the objects present in this image and determining their categories. The way to locate an object varies from detector to detector. For example, mask R-CNN detector uses segmentation mask to locate objects [1], CoKe detector uses key points and landmarks to locate objects [2], and poly-YOLO detector represents objects using polygons [3]. However, the rectangle is considered to be the simplest and most used polygonal shape for locating or representing an object in many computer vision applications for IoT-based [4]. Locating an object using a rectangle consists of drawing a rectangle surrounding this object. The annotation of an object dataset is performed using free or commercial software designed for a particular annotation choice.

There are two types of annotations with rectangles depending on the orientations of the objects in the images for IoT-based. To locate objects, the first type uses horizontal bounding rectangles (HBR), while the second type uses oriented bounding rectangles (OBR). Annotation with horizontal rectangles is suitable for natural scenes, where the photographer is usually in front of the object and adjusts their camera so that the objects appear aligned with the horizontal edges of the image. By contrast, for aerial photography, where images are captured by Earth observation

satellites or other flying devices, objects often appear in the image with arbitrary orientations.

The optimality of an annotation results in the satisfaction of two conditions: (i) the bounding rectangle must cover the whole of the object and (ii) the area of the bounding rectangle must be minimal. For a horizontal object, it is easy to reach the optimal HBR that satisfies conditions (i) and (ii) mentioned above. On the other hand, for an oriented object, it is difficult to reach the optimal OBR if the annotation method does not render account of optimality conditions (i) and (ii).

In the literature, we distinguish two types of object detectors. The first type includes horizontal object detectors, which detect objects using HBR. The second type includes oriented object detectors, which detect objects using OBR. Any detector of the first class is trained, tested, and tuned on horizontal object datasets (HOD). On the other hand, some oriented object detectors are trained and tuned on HOD, such as OAOD [5] and BBAVectors [6], but others are trained tuned on to oriented object datasets (OOD), such as RoI transformer [7], R-RoI [8], RRPN [9], R2CNN++ [10], DMPNet [11], FOTS [12], RPN [13], and DDR [14]. However, all of these oriented object detectors must be tested on OOD.

A large-scale HOD (OOD, respectively) is a set made up of a large number of images annotated with HBR (OBR, respectively). Each image contains objects of a wide variety of scales, orientations, and shapes. These objects are divided into classes (or categories) that vary from one dataset to another.

Table 1 (Table 2, respectively) presents the number of classes, instances, images, and year of creation of the most cited HOD (OOD, respectively) in the literature. To our knowledge, DOTA is the largest public Earth vision object detection dataset [27]. It contains objects exhibiting a wide variety of scales, orientations, and shapes. Moreover, images of DOTA are manually annotated by experts in aerial image interpretation.

2. State of the Art

In order to generate a dataset for object identification, we need to gather a large number of photographs, all of which must be annotated using the same technique and organized in accordance with a set of categories that have been pre-determined. The dataset has to include many photographs that match to various instances of each object class. These images must be included for each object class. The currently available annotation techniques are amenable to being split up into two primary categories. In the first class, all methods of manual annotation are grouped together, and in the second class, all techniques of semi-automatic annotation are categorized together [28].

2.1. Manual Annotation. There are two main methods of manually annotating objects with bounding boxes, which are used in the construction of most large-scale object datasets. The first one is called the consensus method, and the second

TABLE 1: Examples of large-scale horizontal object datasets.

Name	Classes	Instances	Images	Year
DIOR [15]	20	190288	23463	2019
TUT indoor [16]	7	4595	22137	2018
MS COCO [17]	80	886266	123287	2014
PASCAL VOC [18]	17	1793658	11268	2012
ImageNet [19]	200	478807	456567	2009

TABLE 2: Examples of large-scale oriented object datasets.

Name	Classes	Instances	Images	Year
DOTA-v2.0 [20]	17	1793658	11268	2021
FGSD [21]	1	5634	2612	2020
HRSC2016 [22]	1	2976	1061	2016
UCAS-AOD [23]	2	14596	1510	2015
DLR 3k [24]	2	14235	20	2015
VEDAI [25]	3	2950	1268	2015
SZTAKI-INRIA [26]	1	665	9	2012

one is called the sequential tasks method. For each instance of an object in an image, the first method asks several annotators to draw a rectangle around the object and then defines the position of the object by the rectangle elected by the majority of annotators. To annotate an object by the second method, we need at least three annotators. The first one is asked to draw a rectangle around a single instance of the object. The task of the second annotator is to validate the drawn rectangle. The third person investigates whether additional object class instances require annotation. Determining a precise bounding rectangle takes more time and resources than validating the annotation, and thus, the consensus technique is more efficient. Annotation quality affects the development of accurate object detectors. The determination of a large-scale object dataset requires expert annotators, significant working time, and remuneration in line with the desired accuracy. As examples of common use datasets, annotated with the manual methods, we cite PASCAL VOC [18], MS COCO [17], ImageNet [19], DOTA [20], etc.

Although studies on the description of OOD are abundant, very few of them explain how to draw a rectangle enclosing an oriented object. The only annotation method explicitly described in the literature appears in the article [27] by Ding et al. This method consists in drawing a HBR around the object of interest and then adjusting the angle manually by rotating the rectangle with the mouse.

2.2. Semi-Automatic Annotation. The traditional semi-automatic approach of annotating objects with rectangles consists of four phases.

First phase: in this phase, the images of the object dataset are divided into two subsets of unequal sizes. The smallest subset is annotated with rectangles by the manual method.

Second phase: the second phase consists of choosing an object detector and training this detector with the images from the dataset already annotated.

Third phase: once the detector is ready, it will be used to annotate the images of the second object dataset with the prediction rectangles.

Fourth phase: during this phase, the annotator validates the correctly annotated objects during the third phase and manually draws the bounding rectangles of the poorly annotated objects.

As this study is limited to locating objects by bounding boxes, we cite the Faster Bounding Box (FBB) as an example of semi-automatic annotation method [16]. This method was used to generate the Tampere University of Technology (TUT) Indoor dataset. The largest subset is annotated with prediction rectangles, generated by the Faster R-CNN object detector [29], trained on the smallest annotated subset.

2.3. Performance Measure of Detectors Using Rectangles. The evaluation of the effectiveness of the object detector is then performed on the second subset of the annotated dataset. Since we are only dealing here with annotations using bounding boxes, the most used criterion to compare two rectangles is the Jaccard's similarity index, also known as the Intersection over Union (IoU). For each image of the testing dataset, the IoU measures the percentage of overlap between the prediction bounding rectangle $\widehat{\mathbb{A}}$ generated by the detector, and the ground truth bounding box $\mathbb{A}$, as follows:

$$IoU(\widehat{\mathbb{A}}, \mathbb{A}) = \frac{\text{Area}(\widehat{\mathbb{A}} \cap \mathbb{A})}{\text{Area}(\widehat{\mathbb{A}} \cup \mathbb{A})}. \qquad (1)$$

We note that the IoU score lies in the interval $[0, 1]$. The closer this index gets to 1, the better the detection of the object. We say that an object is well detected (or true positive) if the IoU score is greater than or equal to a threshold discussed by experts. In most studies, the threshold value is set at 0.5 [30]. The IoU score causes a problem when its value is zero. This value is embarrassing since it does not explain how far is the prediction bounding rectangle from the ground truth bounding box. To work around this problem, Rezatofighi et al. [30] suggest replacing the IoU with the Generalized Intersection over Union (GIoU) index. The GIoU between two rectangles $\widehat{\mathbb{A}}$ and $\mathbb{A}$ is defined as follows:

$$\text{GIoU}(\widehat{\mathbb{A}}, \mathbb{A}) = IoU(\widehat{\mathbb{A}}, \mathbb{A}) - \frac{\text{Area}(\mathbb{C} \setminus (\widehat{\mathbb{A}} \cup \mathbb{A}))}{\text{Area}(\mathbb{C})}, \qquad (2)$$

where $\mathbb{C}$ is the smallest convex set enclosing both $\widehat{\mathbb{A}}$ and $\mathbb{A}$.

The exact IoU computation between two HBR is simple. Much software integrates functions to calculate this index. On the other hand, the exact IoU computation between two OBR is not as simple as that between two HBR. Liu et al. (Yao et al., respectively) proposed in Ref. [31] (in [32]) heuristic to estimate the IoU between two OBR. Recently, Zaidi has shown in Ref. [28] that these heuristics give reliable results only when the centers of the rectangles are very close and the angle between the rectangles is small. Moreover, he has developed in Ref. [28] an algorithm that calculates the exact IoU value between two OBR, as well as an (ε, α) estimator of the IoU.

The annotation method has a direct bearing on the performance measurement of an object detector. The OOD reserved for the test phase must be carefully annotated so that the IoU computation between the ground truth and prediction rectangles is not biased.

3. Motivation for the Study

The manual or semi-automatic annotation methods described in Section 2 have made it possible to build many large-scale object datasets. These datasets have given rise to very powerful object detectors. However, we cannot ignore the following concerns:

(1) According to paragraph A, accurate manual annotation is expensive, time-consuming, and requires annotation experts.

(2) In many situations, the annotator can be in front of an object with an ambiguous orientation as shown in Figure 1. In this case, he will take a considerable time to decide whether the object will be annotated with a horizontal rectangle or an oriented rectangle. Moreover, in both cases, the final decision is not based on any objective argument and depends solely on the dexterity of the annotator.

(3) The semi-automatic annotation defined in paragraph B does not guarantee the optimality of the annotation rectangle, in the sense that it must have a minimum area and cover the whole of the object.

(4) According to paragraph C, the use of any approximation methods of $IoU(\mathbb{A}, \mathbb{B})$, whether to measure the accuracy of an oriented object detector or to compare the performance of two oriented object detectors, could lead to biased results. Indeed, the function $g(x) = x/a + b - x$ is strictly increasing over the interval $[0, a + b[$, where a and b denote the areas of $\mathbb{A}$ and $\mathbb{B}$, respectively. Therefore, a large or a small ground truth bounding rectangle directly induces a bias in the computation of the IoU.

The contribution of this study is the scarcity of annotation methods motivated us to develop a robust semi-automatic annotation method of OOD. Our method is semi-automatic because it consists of a manual step followed by a computer-assisted step. In addition, this method is robust because the bounding rectangle generated by our algorithm is insensitive to the dexterity of the annotator. More precisely, the steps are as follows:

(i) We develop an algorithm called RANGE-MBR, which determines from a set $\mathcal{M}_n$ of the n most relevant (in the sense given by Definition 1) points picked on the object outline, a rectangle enclosing $\mathcal{M}_n$ and having a quasi-minimal area, in $\mathcal{O}(n)$ time.

(ii) We propose a new approach to simultaneously build HOD and OOD from a large-scale image

(a)   (b)
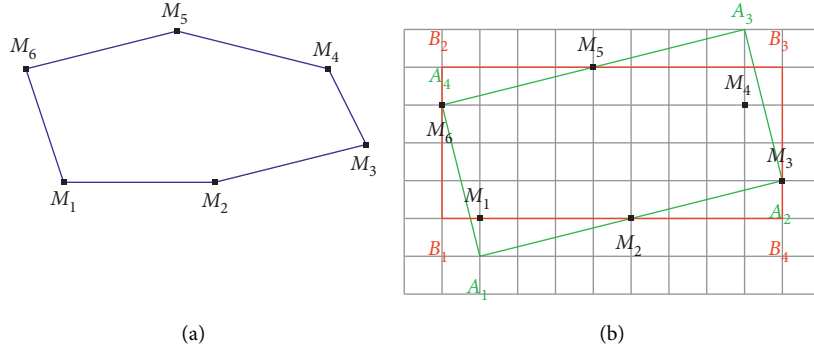
FIGURE 1: (b) Intuitively, it seems that the horizontal rectangle $\mathbb{B} = [B_1, \ldots, B_4]$ is best suited to annotate the colored object in blue (Figure (a)). The method we propose takes as input the set $\mathscr{M}_6 = \{M_1, \ldots, M_6\}$ and returns the oriented rectangle $\mathbb{A} = [A_1, \ldots, A_4]$. Since $\mathbb{A}$ covers the whole object and $\text{Area}(\mathbb{A}) < \text{Area}(\mathbb{B})$, then it is natural to annotate the object by the rectangle $\mathbb{A}$.

bank, based on both the RANGE-MBR algorithm and threshold angles.

(iii) We conduct a large experimental study to quantify the performance of the RANGE-MBR algorithm.

(iv) We compare the performance of RANGE-MBR to that of the benchmark algorithm, RC-MBR, which determines the minimum rectangle enclosing $\mathscr{M}_n$ in $\mathcal{O}(n \ln(n))$ time.

This study seems essential to us because it responds to a real need in the development of oriented object detectors. In addition, it allows researchers to build their own oriented object datasets in a rigorous manner.

Definition 1. (relevant point of an object) Any point located on the contour of the object is said to be relevant if it is a local maximum, a local minimum, the most right point, the most left point, a cusp of the first type, a cusp of the second type, an inflection point, and so on.

The remainder of this manuscript is structured as follows. Section 4 deals with the parametrization of rectangles. Section 5 is devoted to the development of the RANGE-MBR algorithm. We explain in Section 6 how to use the RANGE-MBR (or RC-MBR) algorithm to simultaneously generate horizontal and oriented datasets from a large-scale image bank. In Section 7, we perform various numerical experiments to evaluate the performance of the RANGE-MBR algorithm.

4. Parametrization of Bounding Rectangles

First, we give an overview of the different types of parameterization encountered in the literature. Then, we justify the choice of the parameterization associated with the annotation method that we propose in this study. Our choice of parameterization was discussed in a previous study on the accurate computation of the IoU [28]. However, it seems useful to us to recall this choice to facilitate the reading of the manuscript.

4.1. Parameterization of Horizontal Rectangles. The minimum parameterization of a horizontal rectangle $\mathbb{A} = \{A_1, A_2, A_3, A_4\}$ requires four parameters. The object dataset PASCAL VOC [18] annotates the rectangle $\mathbb{A}$ with (x_4, y_4, x_2, y_2), where (x_i, y_i) are the coordinates of the vertex A_i, $i = 1, \ldots, 4$ (cf. Figure 2). The object dataset Ms COCO [17] annotates the rectangle $\mathbb{A}$ with (x_4, y_4, w, h), where w and h denote the lengths of the line segments $[A_1, A_2]$ and $[A_1, A_4]$, respectively, and (x_4, y_4) are the coordinates of the vertex A_4 (cf. Figure 2). The object dataset ImageNet [19] annotates the rectangle $\mathbb{A}$ with (x_0, y_0, w, h), where w and h denote the lengths of the line segments $[A_1, A_2]$ and $[A_1, A_4]$, respectively, and (x_0, y_0) are the coordinates of the center A_0 of $\mathbb{A}$ (cf. Figure 2).

4.2. Parameterization of Oriented Rectangles. Five parameters—b_1, b_2, w, h, and θ—are necessary for the minimal parameterization of an orientated rectangle denoted by the formula $\mathbb{B} = \{B_1, B_2, B_3, B_4\}$; these are the only parameters that are required. The coordinates for the center B_0 of $\mathbb{B}$ are (b_1, b_2), which correspond to the coupling b_1 and b_2. The lengths of the line segments $[B_1, B_2]$ and $[B_1, B_4]$ are indicated by the parameters w and h, respectively (cf. Figure 2). The parameter $\theta \in [0, \pi/2]$ is used to determine the acute angle that exists between the lines Δ and (B_1, B_2), where Δ denotes the horizontal line that goes through B_1. These two object datasets, HRSC2016 [22] and UCAS-AOD [23], are annotated with the help of this parametrization.

The above θ-based parametrization is not suitable for annotating satellite image datasets. These OOD have the particularity of containing a large number of instances per image, at different scales, and whose parts overlap [20]. A simple solution that overcomes the drawbacks of the θ-based annotation consists in defining a rectangle $\mathbb{B} = \{B_1, B_2, B_3, B_4\}$ using the coordinates $(x_i, y_i)_{1 \le 4}$ of the vertices $B_1, \ldots, B_4$, respectively. To solve the indeterminacy problem linked to the permutations of the vertices, we can sort the vertices clockwise and fix the first point according to the following rules [27]:
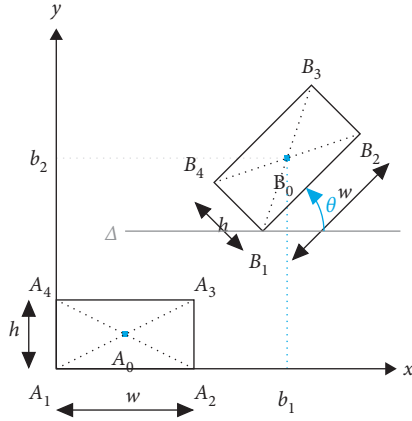
FIGURE 2: Parameterization of horizontal and oriented rectangles.

(i) For objects with a distinguished head and a tail (e.g., vehicles and helicopters), the annotator carefully selects B_1 to indicate the left corner of the instance head.

(ii) For other objects (e.g., tennis courts and bridges), B_1 is the point at the top left of the instance.

This type of parametrization was used for the annotation of the DOTA object dataset [20].

4.3. Our Choice of Parametrization. We use the notation $\mathbb{R} = [R_1, R_2, R_3, R_4]$ to denote an oriented rectangle whose vertices are R_1, R_2, R_3, and R_4. We assume that these vertices are sorted in counterclockwise order, such that R_1 is the left vertex with the smallest vertical coordinate.

To sort the vertices set $\{B_1, B_2, B_3, B_4\}$ of the rectangle $\mathbb{B}$ as described above, we first consider the barycenter G of the rectangle $\mathbb{B}$. Then, we calculate the polar angle of each vertex relative to the barycenter G [28]. We denote by $B_{(1)}, \ldots, B_{(4)}$, the sorted vertices in ascending order of their polar angle. If $B_{(3)}(2) \le B_{(4)}(2)$, then $\mathbb{B} = [B_{(3)}, B_{(4)}, B_{(1)}, B_{(2)}]$, else $\mathbb{B} = [B_{(4)}, B_{(1)}, B_{(2)}, B_{(3)}]$ (cf. Figure 2). For more details on the implementation of this sorting method, we refer the reader to our article [28].

In this investigation, we describe a bounding rectangle $\mathbb{B}$ by using the eight-tuple $(x_1, y_1, \ldots, x_4, y_4)$, where $(x_i, y_i)^t$ are the coordinates of the i-th vertex of the sorted rectangle $[B_1, \ldots, B_4]$.

This choice of parameterization is ideally suited to the annotation approach that we suggest since it eliminates the need for extra computations to draw rectangles or to locate the images of rectangles by applying affine transformations. This is one of the reasons why the method is so efficient. In addition, we demonstrate in Section B how to format the outputs of the annotation technique that we propose in accordance with the annotation rule of the DOTA dataset. This is done by referring to the annotation rule.

5. Determination of the Bounding Rectangle

It is natural to annotate the object with the minimum rectangle enclosing the set of relevant points. This problem is formulated as follows:

Problem 1. Given a set of points $\mathcal{M}_n = \{M_i = (x_i, y_i)^t \in \mathbb{R}^2; i = 1, \ldots, n\}$, we find a rectangle with the smallest area, enclosing $\mathcal{M}_n$. Such a rectangle is called a minimum bounding rectangle of $\mathcal{M}_n$ and denoted by MBR $(\mathcal{M}_n)$.

The solution of problem 5.1 is not unique as shown in Figure 3. So, MBR $(\mathcal{M}_n)$ denotes any solution of Problem 1.

Freeman and Shapira proved in 1975 that one edge of MBR $(\mathcal{M}_n)$ must be collinear with an edge of the convex hull (CH $(\mathcal{M}_n)$) of $\mathcal{M}_n$. They proposed in Ref. [33] a natural algorithm to find MBR $(\mathcal{M}_n)$ in $\mathcal{O}(n^2)$ time, based on sweeping all minimum rectangles, enclosing $\mathcal{M}_n$, and having an edge collinear with an edge of CH $(\mathcal{M}_n)$. In 1978, Shamos proposed in Ref. [34] the famous rotating calipers algorithm, which return all pairs of antipodal vertices of a n-sided convex polygon in $\mathcal{O}(n)$ time. In 1983, Toussaint used the rotating calipers technique and developed the algorithm RC-MBR to find MBR $(\mathcal{M}_n)$ in $\mathcal{O}(n \ln(n))$ time [35]. In 2006, Dimitrov et al. proposed in Ref. [36, 37] the algorithm PCA-MBR, to approximate MBR $(\mathcal{M}_n)$, in $\mathcal{O}(n \ln(n))$ time, by the minimum bounding rectangle, which is aligned with the eigenvectors of the covariance matrix of CH $(\mathcal{M}_n)$. They also proved that the relative error between PCA-MBR $(\mathcal{M}_n)$ and MBR $(\mathcal{M}_n)$ is bounded from above by $\sqrt{8\sqrt{2}} - 1 \approx 2.36$. The main drawback of the algorithm PCA-MBR is that it admits an infinite number of solutions if the covariance matrix of CH $(\mathcal{M}_n)$ has a double eigenvalue. To overcome the problem of nonuniqueness inherent in algorithms RC-MBR and PCA-MBR, we propose the method RANGE-MBR to approximate MBR $(\mathcal{M}_n)$ in $\mathcal{O}(n)$ time.

5.1. The Range-MBR Algorithm. Let $\mathcal{B} = (O, \mathbf{i}, \mathbf{j})$ be a Cartesian frame of the two-dimensional affine plane. All angles we refer to here are measured counterclockwise from the positive x-axis. Let Δ be the line passing through O and making an angle θ with the axis $(O, \mathbf{i})$. Let $\Delta^\perp$ be perpendicular to Δ and passing through O. Let P_1 and Q_1 (P_2 and Q_2, respectively) be the extreme points of the orthogonal projection of $\mathcal{M}_n$ on Δ ($\Delta^\perp$, respectively). Let $\mathbb{A} = [A_1, A_2, A_3, A_4]$ be the minimum bounding rectangle having an edge collinear with Δ. Then, $A_1, \ldots, A_4$ are the intersection points of the two parallel lines to $\Delta^\perp$ and passing through P_1 and Q_1, respectively, with the two parallel lines to Δ and passing through P_2 and Q_2 respectively. The area of $\mathbb{A}$ is given by the product of lengths of the line segments $[P_1, Q_1]$ and $[P_2, Q_2]$. From now on, the MBR $\mathbb{A}$ enclosing $\mathcal{M}_n$ and having an edge collinear with the direction making an angle θ from the positive x-axis will be denoted by $MBR(\mathcal{M}_n, \theta)$.

Let $\mathbf{u} = (\cos(\theta), \sin(\theta))^t$, and $\mathbf{v} = (-\sin(\theta), \cos(\theta))^t$) be two unit direction vector of Δ and $\Delta^\perp$, respectively. Then, $\mathcal{B} = (O, \mathbf{i}, \mathbf{j})$ and $\mathcal{B}' = (O, \mathbf{u}, \mathbf{v})$ are two Cartesian frames of the two-dimensional affine plane. For all $i = 1, \ldots, n$, we denote by $(x_i, y_i)^t$ ($(s_i, t_i)^t$, respectively) the coordinates of M_i with respect to the frame $\mathcal{B}$ ($\mathcal{B}'$, respectively). Let R_θ be the rotation matrix of angle θ, and then $(x_i, y_i)^t = R_\theta(s_i, t_i)^t$, where $\mathbf{u}^t$ is the transpose of $\mathbf{u}$.

Let X be a random variable with mean μ and standard deviation σ. Let $\mathbf{x}_{1: n} = (x_1, \ldots, x_n)$ be n observations on the
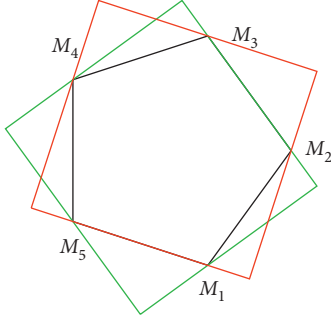
Figure 3: $\mathcal{M}_5 = \{M_1, \ldots, M_5\}$. The green and red rectangles are two distinct solutions of Problem 1.

variable X. Let $x_{(1)} \leq \ldots \leq x_{(n)}$ be the order statistics of $\mathbf{x}_{1:\,n}$. We define the range of $\mathbf{x}_{1:\,n}$ by $\mathrm{Range}(\mathbf{x}_{1:\,n}) = x_{(n)} - x_{(1)}$. Therefore, the lengths of $[P_1, Q_1]$ and $[P_2, Q_2]$ are given by:

$$P_1 Q_1 = \mathrm{Range}(\{s_1, \ldots, s_n\}) = s_{(n)} - s_{(1)},$$
$$P_2 Q_2 = \mathrm{Range}(\{t_1, \ldots, t_n\}) = t_{(n)} - t_{(1)}. \tag{3}$$

We denote by $\hat{\sigma}_n$ any estimation of σ obtained from $\mathbf{x}_{1:\,n}$. On the basis of numerous works carried out on the estimation of the standard deviation from the sample range ([38–43]), we can assume that there is a real constant $\alpha(X, n)$ such that

$$\hat{\sigma}_n = \frac{\mathrm{Range}(\mathbf{x}_{1:\,n})}{\alpha(X, n)}. \tag{4}$$

Besides, it is well known that

$$\mathrm{Var}(\mathbf{x}_{1:\,n}) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x}), \tag{5}$$

which is an estimate of σ^2, where $\overline{x} = 1/n \sum_{i=1}^{n} x_i$. It follows that there is a real constant $\alpha(X, n)$ such that:

$$\mathrm{Range}(\mathbf{x}_{1:\,n})^2 \approx \alpha(X, n)^2 \mathrm{Var}(\mathbf{x}_{1:\,n}). \tag{6}$$

We assume that the (x, y) coordinates of the set of points $\mathcal{M}_n$ are n observations of two random variables X and Y. Since the $\mathrm{Area}(\mathbb{A}) = P_1 Q_1 \cdot P_2 Q_2$, then combining equations (6) and (3) gives an approximation $\widehat{A}_n^2(\theta)$ of $\mathrm{Area}(\mathbb{A})$, satisfying the following relation:

$$\widehat{A}_n^2(\theta) = \alpha(S, n)^2 \alpha(T, n)^2 \mathrm{Var}(\mathbf{s}_{1:\,n}) \mathrm{Var}(\mathbf{t}_{1:\,n}), \tag{7}$$

where $(S, T)^t = R_\theta^t (X, Y)^t$. To alleviate notions, we will also use $\mathbb{V}_\mathbf{x}$ and $\mathbb{C}_\mathbf{x,y}$ to designate $\mathrm{Var}(\mathbf{x}_{1:\,n})$ and $\mathrm{Cov}(\mathbf{x}_{1:\,n}, \mathbf{y}_{1:\,n})$. Since $(x_i, y_i)^t = R_\theta (s_i, t_i)^t$, we obtain

$$\mathbb{V}_s = \cos^2(\theta) \mathbb{V}_x + \sin^2(\theta) \mathbb{V}_y + \sin(2\theta) \mathbb{C}_{x,y},$$
$$\mathbb{V}_t = \sin^2(\theta) \mathbb{V}_x + \cos^2(\theta) \mathbb{V}_y - \sin(2\theta) \mathbb{C}_{x,y}. \tag{8}$$

Using (8) and some trigonometric identities, we prove that

$$8 \mathbb{V}_\mathbf{s} \mathbb{V}_\mathbf{t} = K - f(\theta), \tag{9}$$

where $K = \mathbb{V}_\mathbf{x}^2 + \mathbb{V}_\mathbf{y}^2 + 6 \mathbb{V}_\mathbf{x} \mathbb{V}_\mathbf{y} - 4 \mathbb{C}_{\mathbf{x,y}}^2$ and

$$f(\theta) = \left(\mathbb{V}_\mathbf{x}^2 + \mathbb{V}_\mathbf{y}^2 - 2 \mathbb{V}_\mathbf{x} \mathbb{V}_\mathbf{y} - 4 \mathbb{C}_{\mathbf{x,y}}^2 \right) \cos(4\theta)$$
$$+ 4 \mathbb{C}_{\mathbf{x,y}} \left(\mathbb{V}_\mathbf{x} - \mathbb{V}_\mathbf{y} \right) \sin(4\theta). \tag{10}$$

Combining equations (7), (9), and (10) gives

$$\widehat{A}_n^2(\theta) = \frac{\alpha(S, n)^2 \alpha(T, n)^2}{8} (K - f(\theta)). \tag{11}$$

Thus, the area of MBR $(\mathcal{M}_n, \theta)$ is approximately equal to

$$\frac{\alpha(S, n)^2 \alpha(T, n)^2}{8} (K - f(\theta)). \tag{12}$$

Therefore, we propose to approximate MBR $(\mathcal{M}_n)$ by MBR $(\mathcal{M}_n, \theta^*)$ such that:

$$\theta^* = \underset{\theta \in [-(\pi/2), \pi/2]}{\arg\min} \widehat{A}_n^2(\theta) = \underset{\theta \in [-(\pi/2), \pi/2]}{\arg\max} f(\theta). \tag{13}$$

The first derivative of f with respect to θ is given by:

$$\dot{f}(4\theta) = -4\lambda \sin(4\theta) + 4\delta \cos(4\theta). \tag{14}$$

where

$$\lambda = \left(\mathbb{V}_\mathbf{x}^2 - \mathbb{V}_\mathbf{y} \right)^2 - 4 \mathbb{C}_{\mathbf{x,y}}^2,$$
$$\delta = 4 \mathbb{C}_{\mathbf{x,y}} \left(\mathbb{V}_\mathbf{x} - \mathbb{V}_\mathbf{y} \right). \tag{15}$$

The function f has a unique critical point θ^* such that

$$\tan(4\theta^*) = \frac{\delta}{\lambda} = \frac{4 \mathbb{C}_{\mathbf{x,y}} \left(\mathbb{V}_\mathbf{x} - \mathbb{V}_\mathbf{y} \right)}{\left(\mathbb{V}_\mathbf{x}^2 - \mathbb{V}_\mathbf{y} \right)^2 - 4 \mathbb{C}_{\mathbf{x,y}}^2}. \tag{16}$$

Lemma 1. *The second derivative of f at the critical point θ^* has the same sign as $-\lambda$.*

Proof. The second derivative of f with respect to θ is given by:

$$\ddot{f}(\theta) = -16 (\lambda \cos(4\theta) + \delta \sin(4\theta)). \tag{17}$$

Multiplying both sides of the relation (17) by $\cos(4\theta)$ yields

$$\cos(4\theta) \ddot{f}(\theta) = -16 \left(\lambda \cos^2(4\theta) + \delta \sin(4\theta) \cos(4\theta) \right), \tag{18}$$

Multiplying both sides of the relation (14) by $\sin(4\theta)$ yields

$$\sin(4\theta) \dot{f}(\theta) = -4\lambda \sin^2(4\theta) + 4\delta \sin(4\theta) \cos(4\theta). \tag{19}$$

Using relations (19) and (18) with $\theta = \theta^*$ (i.e., $\dot{f}(\theta^*) = 0$), we obtain:

$$4\delta \cos(4\theta^*) \sin(4\theta^*) = 4\lambda \sin^2(4\theta^*),$$
$$\cos(4\theta^*) \ddot{f}(\theta^*) = -16\lambda. \tag{20}$$

Therefore, $\ddot{f}(\theta^*)$ has the same sign as $-\lambda \cos(4\theta^*)$. Since $4\theta^* = \arctan(\delta/\lambda) \in \,] - (\pi/2), \pi/2\,[$, then $\cos(4\theta^*) > 0$ whatever the set $\mathcal{M}_n$ may be. Thus, $\ddot{f}(\theta^*)$ has the same sign as $-\lambda$.

Therefore, the determination of θ^* goes through six cases:

Case 1: if $\lambda > 0$, then $\theta^* = 1/4\arctan(\delta/\lambda)$.

Case 2: if $\lambda < 0$ and $\mathbb{V}_{\mathbf{x}} \neq \mathbb{V}_{\mathbf{y}}$, then θ^* is determined through the eigen decomposition of the empirical covariance matrix of $\mathcal{M}_n$. This case is discussed below.

Case 3: if $\lambda < 0$ and $\mathbb{V}_{\mathbf{x}} = \mathbb{V}_{\mathbf{y}}$, then $\theta^* = \pi/4$.

Case 4: if $\lambda = 0$ and $\mathbb{C}_{\mathbf{x}}(\mathbb{V}_{\mathbf{x}} - \mathbb{V}_{\mathbf{y}}) > 0$, then $\theta^* = \pi/8$.

Case 5: if $\lambda = 0$ and $\mathbb{C}_{\mathbf{x}}(\mathbb{V}_{\mathbf{x}} - \mathbb{V}_{\mathbf{y}}) > 0$, then $\theta^* = \pi/8$.

Case 6: if $\lambda = 0$ and $\mathbb{C}_{\mathbf{x}}(\mathbb{V}_{\mathbf{x}} - \mathbb{V}_{\mathbf{y}}) > 0$, then it is equivalent to $\mathbb{V}_{\mathbf{x}} = \mathbb{V}_{\mathbf{y}}$ and $\mathbb{C}_{\mathbf{x}} = 0$. This case is discussed below. □

5.1.1. Illustrative Example of Case 6. It is difficult to illustrate all the scenarios, which fall into Case 6. The special case of regular polygons is the one that the annotator may encounter when annotating regular objects such as road signs, buildings, and human faces. Moreover, we have pointed out this case to warn the user to pay attention to regularly shaped objects when using our annotation method.

Proposition 1. *If the elements of $\mathcal{M}_n$ are the vertices of a regular n-sided polygon, then $\mathbb{V}_{\mathbf{x}} = \mathbb{V}_{\mathbf{y}}$ and $\mathbb{C}_{\mathbf{x},\mathbf{y}} = 0$.*

Proof. In the general case, the vertices of a regular polygon are uniformly distributed over a circle with radius r and center G equal to the barycenter of $\mathcal{M}_n$. Thus, without loss of generality, we can assume that $r = 1$, $G = (0,0)^t$, and the kth vertex of $\mathcal{M}_n$ is

$$M_k = \left(\cos\left(\frac{2\pi(k-1)}{n} \right), \ \sin\left(\frac{2\pi(k-1)}{n} \right) \right)^t. \tag{21}$$

Consider the complex sequence $(z^k)_{0 \leq k \leq n-1}$, where $z = \exp(2\pi/n\mathrm{i})$ and i is the imaginary unit. Then,

$$\sum_{k=0}^{n-1} z^k = n(\overline{x} + \overline{y}\mathrm{i}). \tag{22}$$

Since $\sum_{k=0}^{n-1} z^k = (1 - z^n/1 - z) = 0$, then $\overline{x} = \overline{y} = 0$, and

$$\mathbb{C}_{x,y} = \frac{1}{n} \sum_{k=1}^{n} x_k y_k,$$

$$\mathbb{V}_x = \frac{1}{n} \sum_{k=1}^{n} x_k^2, \tag{23}$$

$$\mathbb{V}_y = \frac{1}{n} \sum_{k=1}^{n} y_k^2.$$

Since $\cos(2x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x)$, then

$$\mathbb{V}_x = \frac{1}{2} + \frac{1}{2n} \sum_{k=0}^{n-1} \cos\left(\frac{4\pi k}{n} \right),$$

$$\mathbb{V}_y = \frac{1}{2} - \frac{1}{2n} \sum_{k=0}^{n-1} \cos\left(\frac{4\pi k}{n} \right). \tag{24}$$

Besides, $\sum_{k=0}^{n-1}\cos(4\pi k/n)$ is the real part of

$$\sum_{k=0}^{n-1} \left(z^2 \right)^k = \frac{1 - z^{2n}}{1 - z^2} = 0. \tag{25}$$

Thus, $\mathbb{V}_{\mathbf{x}} = \mathbb{V}_{\mathbf{y}} = 1/2$. Moreover, using $2\cos(x)\sin(x) = \sin(2x)$, we deduce that

$$\mathbb{C}_{\mathbf{x},\mathbf{y}} = \frac{1}{2n} \sum_{k=0}^{n-1} \sin\left(\frac{4\pi k}{n} \right). \tag{26}$$

Since $\sum_{k=0}^{n-1}\sin(4\pi k/n)$ is the imaginary part of

$$\sum_{k=0}^{n-1} \left(z^2 \right)^k = \frac{1 - z^{2n}}{1 - z^2} = 0, \tag{27}$$

then $\mathbb{C}_{\mathbf{x},\mathbf{y}} = 0$.

Note that in the case where the elements of $\mathcal{M}_n$ are the vertices of a regular n-sided polygon, problem 5.1 has exactly n different solutions $\mathrm{MBR}_i(\mathcal{M}_n)$, $i = 1, \ldots, n$. For all $i = 1, \ldots, n - 1$, $\mathrm{MBR}_i(\mathcal{M}_n)$ is the image of $\mathrm{MBR}_0(\mathcal{M}_n)$ by the rotation of angle $\theta_i = (2i/n)\pi$. □

5.1.2. Handling of the Exceptional Cases 2 and 6. Cases 2 and 6 are indeterminate cases for which the minimum of the objective function $f(\theta)$ does not exist. This is why we have thought of a trick to get out of the indeterminacy of each case, by returning to the initial definition of the solution.

(i) For Case 2, the application of an affine isometry to the set of points ($\mathcal{M}_n$) makes it possible to migrate to another case for which the angle θ is well determined. We then obtain the solution of the initial problem by applying the inverse isometry to the solution of the transformed problem, since isometry preserves the areas.

(ii) For Case 6, the use of an extra point M_{n+1} belonging to the convex hull of the set $\mathcal{M}_n$, allows to migrate to another case for which the angle θ is well determined. Another alternative consists in asking the user to click on a point of the object's outline until we get out this state.

Case 2: $\lambda = (\mathbb{V}_{\mathbf{x}} - \mathbb{V}_{\mathbf{y}})^2 - 4\mathbb{C}_{\mathbf{x},\mathbf{y}}^2$ is negative because of $\mathbb{C}_{\mathbf{x},\mathbf{y}}^2$. To get rid of $\mathbb{C}_{\mathbf{x},\mathbf{y}}^2$, we have just to follow the same reasoning on a set of points

$$\mathcal{M}_n' = \left\{ M_i' = (x_i', y_i')^t \in \mathbb{R}^2; i = 1, \ldots, n \right\}, \tag{28}$$

whose coordinates $\mathbf{x}_{1:\,n}' = (x_1, \ldots, x_n')$ and $\mathbf{y}_{1:\,n}' = (y_1, \ldots, y_n')$ are uncorrelated. Let

$$\Gamma = \frac{1}{n} \sum_{i=1}^{n} (M_i - \overline{M})(M_i - \overline{M})^t, \tag{29}$$

be the covariance matrix of the elements of $\mathscr{M}_n$, where $\overline{M} = (\overline{x}, \overline{y})^t$. Let $\Gamma = U\Delta U^t$ be an eigen decomposition of Γ, where $U \in \mathscr{O}_2(\mathbb{R})$ (the set of 2-orthogonal matrices), and $\Delta = \mathrm{diag}(\delta_1, \delta_2)$. For all $i = 1, \ldots, n$, set $M_i' = U^t M_i$ and $\mathscr{M}_n' = (M_i')_{1 \le n \le n}$. Then,t

$$\Gamma' = \frac{1}{n} \sum_{i=1}^{n} (M_i' - \overline{M'})(M_i' - \overline{M'})^t = \Delta, \tag{30}$$

where $\overline{M'} = U^t \overline{M}$. Since $U \in \mathscr{O}_2(\mathbb{R})$, then the linear map defined by $L(M) = U^t M$ preserves the norm and the dot product. Thus, $\mathrm{Area}(\mathrm{MBR}(\mathscr{M}_n)) = \mathrm{Area}(\mathrm{MBR}(\mathscr{M}_n'))$ and MBR $(\mathscr{M}_n)$ is the image of MBR $(\mathscr{M}_n')$ under the map L.

Proposition 2. *If the vectors* $\mathbf{x}_{1:n}$ *and* $\mathbf{y}_{1:n}$ *are uncorrelated* ($\mathbb{C}_{\mathbf{x},\mathbf{y}} = 0$) *and have different variances* ($\mathbb{V}_{\mathbf{x}} \ne \mathbb{V}_{\mathbf{y}}$), *then* $\widehat{A}_n^2(\theta)$ *has a unique minimum at the angle* $\theta = 0$.

Proof. Since $\mathbf{x}_{1:n}$ and $\mathbf{y}_{1:n}$ are uncorrelated, then $\mathbb{C}_{\mathbf{x},\mathbf{y}} = 0$, and $\lambda = (\mathbb{V}_{\mathbf{x}} - \mathbb{V}_{\mathbf{y}})^2$ and $\delta = 0$. According to equation (16), the function f has a unique critical point $\theta^* = 0$. Using Lemma 1, $f(\theta)$ is the maximum of $f(\theta)$.

Case 6. To get out of the indeterminacy of Case 6, we propose to add to the set $\mathscr{M}_n$ an artificial point M_{n+1} located in the convex hull of $\mathscr{M}_n$ so that the empirical covariance matrix of the new set $\mathscr{M}_{n+1} = \mathscr{M}_n \cup M_{n+1}$ is different from a scalar matrix. We assume that $(x_{n+1}, y_{n+1})^t$ are the Cartesian coordinates of M_{n+1} with respect to $\mathscr{B}_G = (G, \mathbf{i}, \mathbf{j})$, where $G = (\overline{x}, \overline{y})^t$ is the barycenter of $\mathscr{M}_n$. Without any loss of generality, we assume that $\mathbf{x}_{1:n+1}$ ($\mathbf{y}_{1:n+1}$, respectively) is the vector of the x-coordinates (y-coordinates, respectively) of the points of $\mathscr{M}_{n+1}$ with respect to $\mathscr{B}_G = (G, \mathbf{i}, \mathbf{j})$. We denote by $\widetilde{\mathbb{V}}_{\mathbf{x}} = \mathrm{Var}(\mathbf{x}_{1:n+1})$ and $\widetilde{\mathbb{V}}_{\mathbf{y}} = \mathrm{Var}(\mathbf{y}_{1:n+1})$. Lemma 2 gives the relation between $\widetilde{\mathbb{V}}_{\mathbf{x}}, \mathbb{V}_{\mathbf{y}}, \mathbb{C}_{\mathbf{x},\mathbf{y}}$ and $\widetilde{\mathbb{V}}_{\mathbf{x}}, \widetilde{\mathbb{V}}_{\mathbf{y}}, \widetilde{\mathbb{C}}_{\mathbf{x},\mathbf{y}}$, respectively. $\square$

Lemma 2. *If* $\mathbb{V}_{\mathbf{x}} = \mathbb{V}_{\mathbf{y}}$ *and* $\overline{x} = \overline{y} = \mathbb{C}_{\mathbf{x},\mathbf{y}} = 0$, *then*

$$\widetilde{\mathbb{V}}_{\mathbf{x}} = \frac{n}{n+1}\mathbb{V}_{\mathbf{x}} + \frac{n}{(n+1)^2}x_{n+1}^2,$$

$$\widetilde{\mathbb{V}}_{\mathbf{y}} = \frac{n}{n+1}\mathbb{V}_{\mathbf{y}} + \frac{n}{(n+1)^2}y_{n+1}^2, \tag{31}$$

$$\widetilde{\mathbb{C}}_{\mathbf{x},\mathbf{y}} = \frac{n}{(n+1)^2}x_{n+1}y_{n+1}.$$

Proof. We set $\widetilde{x} = (1/n+1)\sum_{i=1}^{n+1} x_i$ and $\widetilde{x} = (1/n+1)\sum_{i=1}^{n+1} y_i$. Since $\overline{x} = \overline{y} = 0$, then $\widetilde{x} = x_{n+1}/n+1$ and $\widetilde{y} = y_{n+1}/n+1$. On the one hand, since $\overline{x} = 0$, we have

$$\sum_{i=1}^{n+1} (x_i - \widetilde{x})^2 = \sum_{i=1}^{n} x_i^2 + x_{n+1}^2 - (n+1)\widetilde{x}^2 \tag{32}$$

$$= n\mathbb{V}_{\mathbf{x}} + \frac{n}{n+1}x_{n+1}^2.$$

On the other hand, since $\sum_{i=1}^{n+1} x_i y_i = 0$, we have:

$$\sum_{i=1}^{n+1} (x_i - \widetilde{x})(y_i - \widetilde{y}) = \sum_{i=1}^{n+1} x_i y_i - (n+1)\widetilde{x}\widetilde{y} \tag{33}$$

$$= \left(\frac{n}{n+1}\right)x_{n+1}y_{n+1}.$$

Let $\widetilde{f}(\theta)$ be the function defined from $\mathscr{M}_{n+1}$ in the same way that $f(\theta)$ was defined from $\mathscr{M}_n$ in (10). Then,

$$\widetilde{f}(\theta) = \widetilde{\lambda}\cos(\theta) + \widetilde{\delta}\sin(\theta), \tag{34}$$

where

$$\widetilde{\lambda} = \left(\widetilde{\mathbb{V}}_{\mathbf{x}}^2 + \widetilde{\mathbb{V}}_{\mathbf{y}}^2 - 2\widetilde{\mathbb{V}}_{\mathbf{x}}\widetilde{\mathbb{V}}_{\mathbf{y}} - 4\widetilde{\mathbb{C}}_{\mathbf{x},\mathbf{y}}^2\right),$$

$$\widetilde{\delta} = 4\widetilde{\mathbb{C}}_{\mathbf{x},\mathbf{y}}(\widetilde{\mathbb{V}}_{\mathbf{x}} - \widetilde{\mathbb{V}}_{\mathbf{y}})\sin(4\theta). \tag{35}$$

Using the identities (31), we have

$$\widetilde{\lambda} = \frac{n^2}{(n+1)^4}\left(x_{n+1}^4 + x_{n+1}^4 - 6x_{n+1}^2 y_{n+1}^2\right), \tag{36}$$

It follows that

(1) $\widetilde{\lambda} \ne 0$ i.f.f $M_{n+1} \notin D_1 \cup D_2 \cup D_3 \cup D_4$, where $D_1, \ldots, D_4$ are the lines whose Cartesian equations are given by

$$D_1: x = (1 + \sqrt{2})y D_2: x = (1 - \sqrt{2})y,$$
$$D_3: x = (\sqrt{2} - 1)y D_4: x = (-1 - \sqrt{2})y. \tag{37}$$

(2) $\widetilde{\delta} \ne 0$ i.f.f $M_{n+1} \notin D_5 \cup D_6$, where $D_5: x = y$ and $D_6: x = -y$.

In summary, any point $M_{n+1} = (\overline{x} + x_{n+1}, \overline{y} + y_{n+1})^t$ defined as a convex combination of $M_1, \ldots, M_n$, such that $x_{n+1}y_{n+1} \ne 0$ overcomes the indeterminacy problem posed by Case 6. However, the resulting RANGE-MBR $(\mathscr{M}_n)$ depends on the values of x_{n+1} and y_{n+1}. We tested this technique on a set $\mathscr{M}_n$ composed of vertices of a regular n-sided polygon, and we observed that the choice

$$M_{n+1} = \frac{1}{2}\left(M_j + M_{\mathrm{nod}(j+1,4)+1}\right). \tag{38}$$

We allow our algorithm to reach MBR $(\mathscr{M}_n)$ whatever the value of n and $j = 1, \ldots, n$. The generalization of the technique applied to the vertices of a regular polygon requires the calculation of CH $(\mathscr{M}_n)$, and the determination of the midpoint, M_{n+1}, of any two consecutive vertices of CH $(\mathscr{M}_n)$, such that the product of the coordinates of M_{n+1} with respect to $\mathscr{B}_G = (G, \mathbf{i}, \mathbf{j})$, is different from 0. $\square$

5.2. Annotate the DOTA Dataset Using Range-MBR. To adapt our annotation method to the output format required by the DOTA dataset, it suffices to ask the annotator that the first click points to the object's head, and to add a binary variable T equals to 1 if we can differentiate the head of the object from its tail, and zero otherwise. When we pass the set $\mathcal{M}_n$ to the RANGE-MBR algorithm, we save the coordinates of the first point M_1. Next, we determine the rectangle $\mathbb{R} = [R_1, R_2, R_3, R_4] = \text{RANGE} - MBR(\mathcal{M}_n)$.

(1) If the binary variable T equals one, then we determine the vertex R_j of $\mathbb{R}$ that is closest to M_1, and we permute the vertices of $\mathbb{R}$ clockwise so that R_j occupies the first position in this permutation.

(2) If the binary variable T equals zero, then we determine the top left vertex R_j of $\mathbb{R}$, and we permute the vertices of $\mathbb{R}$ clockwise so that R_j occupies the first position in this permutation.

6. Robust Semi-Automatic Annotation

In this section, we provide a new approach for building object detection datasets based on both MBR algorithm and threshold angles. This method is semi-automatic because it consists of a manual step followed by a computer-assisted step. In addition, this method is robust because the bounding rectangle generated by our algorithm is insensitive to the dexterity of the annotator.

(i) The problem of MBR was dealt with in Section 5. If the user needs an optimal annotation in the sense given by Definition 2, then he calls the RC-MBR algorithm. Otherwise, he calls the RANGE-MBR algorithm for quasi-optimal annotation. However, optimality and complexity are inversely proportional.

(ii) By default, the threshold angle is equal to zero. It can also be adjusted by experts in object detection, or it can be defined experimentally as the largest angle between the ground truth bounding box and the positive x-axis that gives no significant difference between the performance of horizontal and oriented detectors when tested on oriented objects. The experimental determination of the threshold angle requires horizontal and oriented object detectors, as well as the already existing oriented object datasets.

6.1. Properties of the Robust Semi-Automatic Annotation Method. By construction, the robust semi-automatic annotation method ensures the following properties:

(1) The bounding rectangle generated by our approach is quasi-optimal in the sense that it covers the whole object, and its area is close to the area of the MBR enclosing the object.

(2) The angle of the rectangle is determined by an algorithm based on some relevant points collected on the contour of the object.

(3) The bounding rectangle is insensitive to annotators provided all relevant points on the object have been selected.

(4) The determination of the bounding rectangle requires $\mathcal{O}(n)$ elementary operations, where n is the number of relevant points.

(5) It allows the user to build simultaneously, from an image bank, two databases: one for horizontal objects and another for oriented objects.

In summary, the robust semi-automatic annotation provides a simple solution to all the drawbacks mentioned in paragraph III, which are inherent in the old annotation methods.

Let:

(i) θ_0 be a threshold angle fixed by the user (by default, $\theta_0 = 0$),

(ii) $\mathcal{M}_n$ be the set of relevant points selected on the contour of the object,

(iii) H-MBR $(\mathcal{M}_n)$ be the horizontal minimum bounding rectangle enclosing $\mathcal{M}_n$,

(iv) O-MBR $(\mathcal{M}_n)$ be the minimum bounding rectangle enclosing $\mathcal{M}_n$,

(v) θ be the angle of O-MBR $(\mathcal{M}_n)$,

then, the Algorithm 1, and the flowchart shown in Figure 4 summarize the stages of construction of horizontal and oriented datasets, from a large-scale image bank.

7. Experimental Study

This experimental study was carried out exclusively with the Matlab R2007b software. We implemented the RC-MBR, PCA-MBR, and RANGE-MBR algorithms in MATLAB language, and we wrote a script (see the Appendix), which

(1) reads the image then displays it

(2) ask the annotator to click on the most relevant points of the object ($\mathcal{M}_n$ is a $2 \times n$ matrix)

(3) determine the rectangle $\mathbb{A}$ corresponding to RC-MBR $(\mathcal{M}_n)$

(4) determine the rectangle $\mathbb{B}$ corresponding to RANGE-MBR $(\mathcal{M}_n)$

(5) draw the two rectangles with the colors red and green, respectively

The images used in experiments B,...,D are free of rights and collected on the net. Moreover, the optimality criterion of an annotation is given in Definition 2.

Definition 2. (optimal annotation) A rectangle $\mathbb{A} = [A_1, \ldots, A_4]$ enclosing an object is said to be optimal, if it fulfills the conditions (i) and (ii):

(i) the rectangle $\mathbb{A}$ covers the whole object,

(ii) the area of $\mathbb{A}$ is minimal.
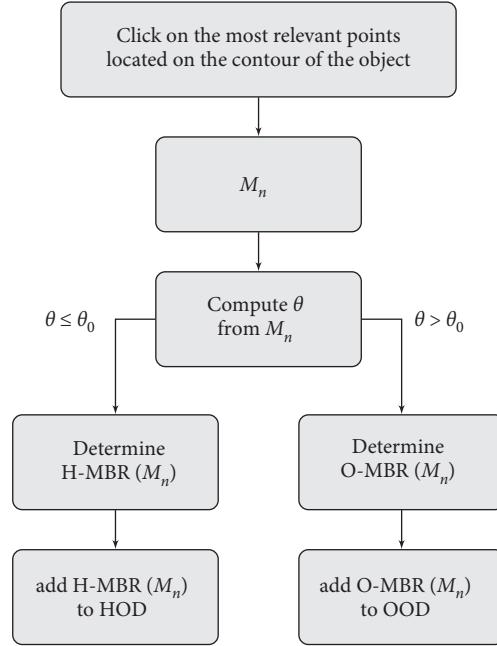
FIGURE 4: Our approach consists in choosing between O-MBR ($\mathcal{M}_n$) and H-MBR ($\mathcal{M}_n$)). If the angle of O-MBR ($\mathcal{M}_n$) is less than or equal to a threshold angle, then we annotate the object using H-MBR ($\mathcal{M}_n$) instead of O-MBR ($\mathcal{M}_n$).

Input: An image; threshold angle; HOD and OOD datasets
Output: Annotated image assigned to HOD or OOD
(1) for (each object in the image) do
(2) Determine the class of the object
(3) Ask the annotator to click on the most relevant points ($\mathcal{M}_n$) of the object as defined in Definition 1
(4) Determine the RANGE-MBR $\mathbb{A} = [A_1, A_2, A_3, A_4]$ enclosing $\mathcal{M}_n$ as described in Section A
(5) Set α the angle between the positive x-axis and $\overrightarrow{A_1 A_2}$
(6) if ($\alpha >$ threshold angle) then
(7) Assign (image, object, $\mathbb{A}$, class) to the OOD dataset
(8) else
(9) Determine the horizontal MBR $\mathbb{A} = [A_1, A_2, A_3, A_4]$ enclosing the relevant points, as shown in Figure 5, by setting $\theta = 0$
(10) Assign (image, object, $\mathbb{A}$, class) to the HOD dataset
(11) end if
(12) end for

ALGORITHM 1: Robust semi-automatic annotation

7.1. Experiment 1. This experiment consists of studying the optimality condition (ii) of the RANGE-MBR algorithm. Equation (28) describes how to generate the coordinates of the vertices $\{M_n = (x_k, y_k); k = 1, \ldots, n\}$ of a random n-sided polygon where each rand is call of a generator of uniform random numbers on the interval $[0, 1]$. The parameter v_y, called the factor of dilation, controls the aspect ratio of the polygon as shown in Figure 6. The more v_y is greater than 1, the more the polygon is dilated in the direction of the y-axis. Note that such a polygon is not necessarily convex as shown in Figure 7. This is also the case for any polygon whose vertices are defined by the relevant points of an object.

Since

(1) the optimal annotation criterion that we have chosen results in a rectangle which has a minimum area and which covers the maximum of visible parts of the object,

(2) RC-MBR is the fastest algorithm that determines the smallest rectangle enclosing a set of points,

it seems natural to consider this algorithm as a reference in the comparative study that we carried out.
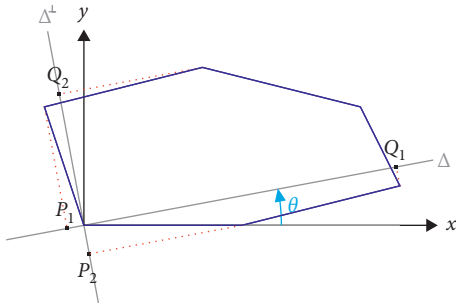
FIGURE 5: When θ is fixed, the minimum bounding rectangle enclosing the blue polygon (object) is the rectangle $\mathbb{A}$ whose vertices $A_1, \ldots, A_4$ are the intersection points of the two parallel lines to $\Delta^\perp$ and passing through P_1 and Q_1, respectively, with the two parallel lines to Δ and passing through P_2 and Q_2, respectively. $\mathbb{A}$ is the rectangle whose edges pass through the red dotted segments.

FIGURE 6: The original object is the regular blue colored heptagon. The black, red, and green polygons correspond to the dilation of the blue polygon by a factor $v_y = 2, 3, 4$, respectively, in the vertical direction.

$$
\begin{cases}
x_k = (1 + \text{rand})\cos\left(\dfrac{2\pi k}{n} + \dfrac{2\pi}{n}\text{rand}\right), \\[2mm]
y_k = \left(1 + v_y \text{rand}\right)\sin\left(\dfrac{2\pi k}{n} + \dfrac{2\pi}{n}\text{rand}\right).
\end{cases}
\tag{39}
$$

For each value of the pair (n, v_y), we generate $r = 10000$ random n-sided polygons. For each polygon $\mathcal{M}_n^{(i)}$, $i = 1, \ldots, r$, we determine the relative error e_i, between the areas of RANGE-MBR $(\mathcal{M}_n^{(i)})$ and RC-MBR $(\mathcal{M}_n^{(i)})$, as well as the CPU times $\hat{t}_i$ and t_i, used by the algorithms RANGE-MBR and RC-MBR to compute RANGE-MBR $(\mathcal{M}_n^{(i)})$ and RC-MBR $(\mathcal{M}_n^{(i)})$, respectively. Finally, we denote by $\bar{e}, \bar{\hat{t}}, \bar{t}$, and $\text{Std}(e)$, the means and the standard deviation of the sequences $(e_i)_{1 \leq i \leq r}$, $(\hat{t}_i)_{1 \leq i \leq r}$, $(t_i)_{1 \leq i \leq r}$, and $(e_i)_{1 \leq i \leq r}$ respectively.

Figure 8 and 9 represent $\bar{e}$ and $\text{Std}(e)$ versus n and v_y for $n \in \{4, \ldots, 20\}$ and $v_y \in \{1, \ldots, 4\}$. We have retained the ranges of values $[4, 20]$ for n and $[1, 4]$ for v_y, because in our estimation, they are those which correspond most to reality.

FIGURE 7: P_1 and P_2 are two random heptagons, generated according to Equation (28) with $v_y = 4$. P_1 is concave, while P_2 is convex.
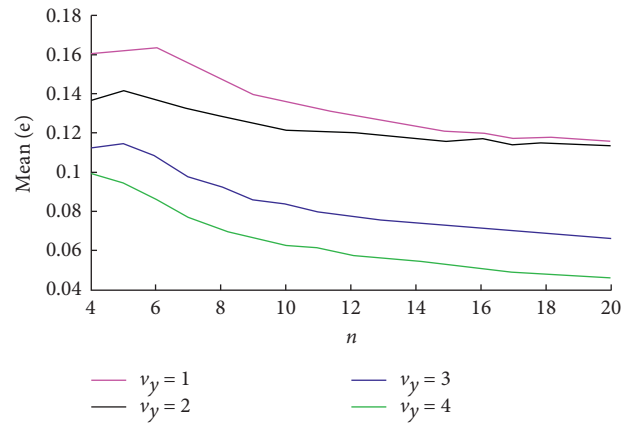
FIGURE 8: Mean of $(e_i)_{1 \leq i \leq r}$ versus n and v_y.

FIGURE 9: Standard deviation of $(e_i)_{1 \leq i \leq r}$ versus n and v_y.

In practice, rare are objects that have more than 20 relevant points or less than 4 relevant points.

We deduce from Figures 8 and 9 that:

(1) The more the polygon is dilated in one direction, the more the algorithm RANGE-MBR is accurate and precise

(2) The more vertices the polygon has, the more the algorithm RANGE-MBR is accurate and precise

We ran other simulations with larger values of n and v_y, and we got the same previous conclusion. The parameter v_y controls the dilation of the polygon in the vertical direction. In fact, we could choose any other direction, since the RANGE-MBR algorithm is not sensitive to the direction of expansion. On the other hand, it is sensitive to the number of vertices of the polygon and to its dilation.

Out of the 680000 generated polygons, we do not encounter any case for which $\mathbb{V}_x = \mathbb{V}_y$ and $\mathbb{C}_{x,y} = 0$. In addition, the algorithm RANGE-MBR is 9 times faster than algorithm RC-MBR. The mean CPU times of RANGE-MBR and CR-MBR are both independent of n and v_y. It is about 1.38×10^{-4} seconds for the algorithm RANGE-MBR, and 1.20×10^{-3} seconds for the algorithm RC-MBR. Since the complexity of the algorithm RANGE-MBR is $\mathcal{O}(n)$, and that of the algorithm RC-MBR is $\mathcal{O}(n \ln(n))$, then the difference in CPU time, for each algorithm, is only observed if n is large enough. Based on the response time per click given in [44] and considering the mean CPU time of the RANGE-MBR algorithm, we can state that determining a bounding rectangle requires $2.5 + 1.5(n-1)$ seconds, where n is the number of relevant points collected on the contour of the image.

7.2. Experiment 2.
We made a comparison between RANGE-MBR and PCA-MBR based on sets of vertices of regular n-sided polygons, with $n = 4, \ldots, 20$. We observe that the relative error of the algorithm RANGE-MBR is always equal to 0, while that of the algorithm PCA-MBR is different from 0 for the values of n reported in Table 3. Although Case 6 does not contain only regular polygons, this experiment shows that the RANGE-MBR method achieves optimality on regular polygons, but for the other shapes (which go in Case 6), it offers a better solution than that obtained by PCA-MBR.

7.3. Experiment 3.
This experiment consists in studying the effect of the dilation factor on the performance of the RANGE-MBR algorithm. Figure 10 represents a basic experiment. The relevant points $\mathcal{M}_n$ are colored in yellow. We observe that RANG-MBR $(\mathcal{M}_n)$ is close to RC-MBR $(\mathcal{M}_n)$. Although the bird on the left is smaller than the bird on the right, the relative error e equals 0.65% for the bird on the right, and 3.12% for the bird on the left. These scores are expected, since the bird on the right has a more elongated shape than the bird on the left, and RANGE-MBR is more effective on dilated objects.

7.4. Experiment 4.
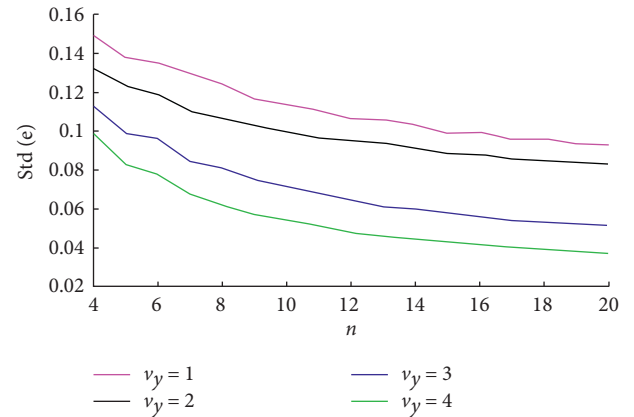This experiment consists in confirming the conclusion obtained in the experiment D. Figure 11 contains two mangoes. The one on the left is almost circular, while the one on the right is clearly elongated. The relevant points $\mathcal{M}_n$ are colored in yellow. The red rectangle corresponds to RANG-MBR $(\mathcal{M}_n)$, while the green rectangle corresponds to RC-MBR $(\mathcal{M}_n)$. The relative error e equals 6.93% for the left mango and 1.86% for the right mango. This real example agrees with the simulation results shown in Figure 8. The mango on the right is more dilated than the

Table 3: Relative error e of PCA-MBR on regular n-sided polygon.

n	4	6	8	9	12	14	15	16	20
$100e$	100	6.6	17.2	0.8	0.7	1.3	0.3	4.0	2.5

—— RC-MBR
—— RANGE-MBR

Figure 10: Illustration of the semi-automatic annotation using RC-MBR and RANGE-MBR algorithms.

Figure 11: The relevant points $\mathcal{M}_n$ are colored in yellow. The red rectangle corresponds to RANG-MBR $(\mathcal{M}_n)$, while the green rectangle corresponds to RC-MBR $(\mathcal{M}_n)$.

one on the left that is why the relative error for the circular mango is greater than relative error for the oval mango.

7.5. Experiment 5.
In order to verify the robustness of our annotation method with respect to the annotator, we asked a colleague to click on the relevant points of the two mangoes. Figure 12 illustrates the result of this experiment. The relative error e equals 8.92% for the left mango, and 2.32% for the right mango. The relative errors reported in Figure 12 are larger than those reported in Figure 11. We explain this difference by the number of points used in each experiment: the more points we use, the more the RANGE-MBR algorithm reduces the relative error. This real example agrees with the simulation results shown in Figure 8.

7.6. Experiment 6.
This experiment consists in comparing our annotation method to the FBB method introduced in Section B. For this, we have chosen a random image from the
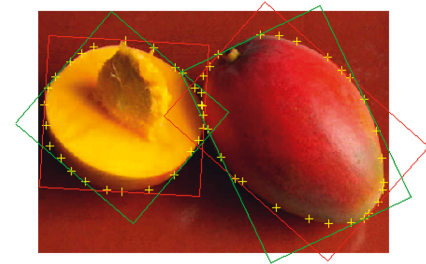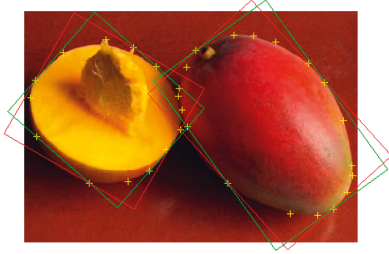
FIGURE 12: The relevant points $\mathcal{M}_n$ are colored in yellow. The red rectangle corresponds to RANG-MBR ($\mathcal{M}_n$), while the green rectangle corresponds to RC-MBR ($\mathcal{M}_n$).

TUT indoor dataset, and we have annotated the objects it contains with FBB, RC-MBR, and RANGE-MBR. Note that the FBB annotation uses blue rectangles, RC-MBR annotation uses green rectangles, RANGE-MBR annotation uses red rectangles, and the relevant points are marked in yellow.

Based on Figure 13 and Table 4, it can be seen that the annotation by the FBB method is not optimal. Indeed,

(i) For the upper extinguisher and the exit sign, condition (i) is violated

(ii) For the lower extinguisher, condition (ii) is violated

In addition, the annotation by the RANGE-MBR method satisfies condition (i) and the relative error between the area of RANGE-MBR and RC-MBR is 12% for the exit sign, 3% for the upper extinguisher, and 0.3% for the lower extinguisher. We can conclude that condition (ii) is almost satisfied by the RANGE-MBR method.

(i) The values of the relative error are in agreement with the simulation results presented in Figure 8. Indeed, in terms of dilation, the lower extinguisher is the most dilated, followed by the upper extinguisher, then the exit sign.

(ii) The exit sign and the upper extinguisher have regular shapes, while the lower extinguisher has an irregular shape. We have already underlined in Section 1 and that the RANGE-MBR method is sensitive to regular forms. This example then illustrates the situation of Case 6.

For all these reasons, the annotation of the lower extinguisher by the RANGE-MBR method is the best.

7.7. Experiment 7.

This experiment highlights the importance of the annotation method on the calculation of the IoU. Figure 14 corresponds to image P_{1128} from the DOTA dataset. The objects of interest in these images are the airplanes. We use the BBAVectors detector [6] to generate the red prediction bounding boxes. Ground truth bounding boxes used in [6] are colored green. The blue rectangles correspond to the annotations of the airplanes by the RANGE-MBR method. For all $i = 1, 2, 3$, we denote by IoU_i^g (IoU_i^b, respectively) the Intersection over Union between the green (blue, respectively) rectangle enclosing the airplane A_i
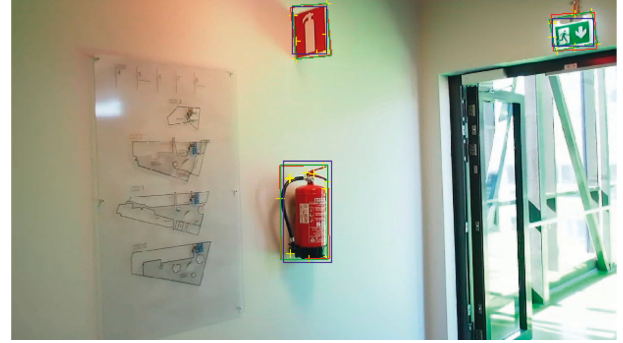
FIGURE 13: Annotation of an image with FBB, RANGE-MBR, and RC-MBR.

TABLE 4: Areas of bounding rectangles in Figure 13.

Object	Method		
	FBB	RANGE-MBR	RC-MBR
Upper fire extinguisher	7227	8295	8046
Lower extinguisher	21726	19917	19844
Exit sign	4902	6927	6146

FIGURE 14: RANGE-MBR allows to build quasi-minimum bounding rectangles.

TABLE 5: IoU computation for the objects in Figure 14.

Airplane	1	2	3
IoU_i^g	0.5230	0.5114	0.7252
IoU_i^b	0.6222	0.6186	0.7087

and the corresponding red rectangle. The results of this experiment are reported in Table 5. The use of minimum ground truth rectangles best reflects the performance of an object detector. In view of the results of Table V, it is reasonable to rely on the results of the second row rather than those of the first row.

8. Conclusion

In this piece, we provide a novel approach that is both resilient and semi-automatic in nature for the annotation of objects. According to the results of the experimental research, we are able to assert that

(1) Robust semi-automatic annotation is quasi-optimal in the sense that Definition 2 describes, and that its optimality increases with dilated objects, which is the case for the majority of the objects that appear on the images that were collected by the use of aerial photography.

(2) Robust semi-automatic annotation is quick and reliable in the sense that the bounding rectangle does not have any gaps, and is oblivious to the skill of the person doing the annotation.

(3) Robust semi-automatic annotation is simple to put into action and might be easily incorporated into platforms for annotating text.

(4) Robust semi-automatic annotation is sensitive to the annotation of items displaying symmetry, with respect to one or more directions. Given this scenario, it is not appropriate for the relevant points to follow the same symmetry in order for the created rectangle to be somewhat close to the ideal rectangle.

Thqe critical angle, as described in Section 6, would look somewhat like this: the emphasis on experimental research using large-scale collected data on horizontal and oriented object datasets, as well as detectors of state-of-the-art oriented objects. Once a good estimate of the threshold angle has been made, the robust semi-automatic annotation technique is the one that allows the simultaneous construction, from the same image library, of two different data sets: one for the elements horizontal and another for oriented objects.

Appendix

A. [Script. 1: Collect $\mathcal{M}_n$, draw RC-MBR ($\mathcal{M}_n$) and RANGE-MBR ($\mathcal{M}_n$)] Here, we assume that we have already implemented the RANGE-MBR and RC-MBR functions, in the MATLAB language. The following scripts

(1) Read the image and then display it

(2) Ask the annotator to click on the most relevant points of the object ($\mathcal{M}_n$ is a $2 \times n$ matrix)

(3) Determine the rectangle $\mathbb{A}$ corresponding to RC-MBR ($\mathcal{M}_n$)

(4) Determine the rectangle $\mathbb{B}$ corresponding to RANGE-MBR ($\mathcal{M}_n$)

(5) Draw the two rectangles with the colors red and green, respectively

```
1  % read the image
2  photo=imread('mango.jpg');
3  % display the image
4  image(photo)
5  button=1;
6  X=[];
7  Y=[];
8  % read ginputs until a mouse right-button occurs
9  while(sum(button)<=1)
10     [x,y,button]=ginput(1);
11     X=[X x];
12     Y=[Y y];
13 end
14 n=length(X);
15 Mn=[X;Y];
16 % A is a (2,4)-matrix containing the coordinates of
       RC-MBR(Mn)
17 A=RC_MBR(Mn);
18 A=[B B(:,1)];
19 % B is a (2,4)-matrix containing the coordinates of
       RANGE-MBR(Mn)
20 B=RANGE_MBR(Mn);
21 B=[B B(:,1)];
22 hold on
23 % draw the points of Mn in yellow
24 plot(M(1,:), M(2,:), '*y')
25 % draw rectangle A in green
26 plot(A(1,:), A(2,:), '-g')
27 % draw rectangle B in red
28 plot(B(1,:), B(2,:), '-r')
29 axis equal
```

Data Availability

The data used to support the findings of this study are included within the article.

Disclosure

There is a basic version [45] of this study accessible at https://www.researchsquare.com/article/rs-860574/v1. The authors confirm that this study has not been published elsewhere.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] K. He, "Mask r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision*, IEEE, Venice, Italy, October 2017.

[2] Y. Bai, "Coke: localized contrastive learning for robust keypoint detection," 2020, https://arxiv.org/abs/2009.14115.

[3] P. Hurtik, "Poly-yolo: higher speed, more precise detection and instance segmentation for yolov3," 2020, https://arxiv.org/abs/2005.13243.

[4] A. Ullah, "Mathematical problems in engineering. experimental and numerical research of paved microcrack using histogram equalization for detection and segmentation," *Mathematical Problems in Engineering*, vol. 2022, Article ID 2684983, 13 pages, 2022.

[5] J. Iqbal, M. A. Munir, A. Mahmood, A. R. Ali, and M. Ali, "Leveraging orientation for weakly supervised object detection with application to firearm localization," *Neurocomputing*, vol. 440, pp. 310–320, 2021.

[6] J. Yi, "Oriented object detection in aerial images with box boundary-aware vectors," 2020, https://arxiv.org/abs/2008.07043.

[7] J. Ding, "Learning roi transformer for detecting oriented objects in aerial images," in *Proceedings of the The IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Long Beach, CA, USA, June 2019.

[8] S. Azimi, "Towards multi-class object detection in unconstrained remote sensing imagery," in *Proceedings of the Asian Conference on Computer Vision*, pp. 150–165, Perth, Australia, 2018.

[9] J. Ma, W. Shao, H. Ye et al., "Arbitrary-oriented scene text detection via rotation proposals," *IEEE Transactions on Multimedia*, vol. 20, no. 11, pp. 3111–3122, 2018.

[10] X. Yang, "R2cnn++: multi-dimensional attention based rotation invariant detector with robust anchor strategy," 2018, https://arxiv.org/pdf/1811.07126.

[11] Y. Liu, "Deep matching prior network: toward tighter multi-oriented text detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1962–1969, IEEE, Honolulu, HI, USA, July 2017.

[12] X. Liu, Q. Yu, and J. Yan, "Fots: fast oriented text spotting with a unified network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5676–5685, IEEE, Salt Lake City, UT, USA, June 2018.

[13] L. Li, Z. Zhou, B. Wang, L. Miao, and H. Zong, "A novel cnn-based method for accurate ship detection in hr optical remote sensing images via rotated bounding box," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 1, pp. 686–699, 2021.

[14] W. He, Y. Fei, and L. Cheng-Lin, "Deep direct regression for multi-oriented scene text detection," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 745–753, IEEE, Venice, Italy, October 2017.

[15] K. Li, G. Wan, G. Cheng, L. Meng, and J. Han, "Object detection in optical remote sensing images: a survey and a new benchmark," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 296–307, 2020.

[16] B. Adhikari, P. Jussi, and H. Heikki, "Faster bounding box annotation for object detection in indoor scenes," in *Proceedings of the European Workshop on Visual Information Processing*, pp. 1–6, London, UK, 2018.

[17] T. Lin, M. Michael, B. Serge et al., "Microsoft coco: common objects in context," *European Conference on Computer Vision*, Springer, Cham, New York, NY, USA, 2014.

[18] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: a retrospective," *International Journal of Computer Vision*, vol. 111, no. 1, pp. 98–136, 2015.

[19] O. Russakovsky, J. Deng, H. Su et al., "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[20] G. Xia, D. Xiang, Z. Jian, and B. Zhen, "Dota: a large-scale dataset for object detection in aerial images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Salt Lake City, UT, USA, June 2018.

[21] K. Chen and M. Wu, "Fgsd: a dataset for fine-grained ship detection in high resolution satellite images," 2020, https://arxiv.org/abs/2003.06832.

[22] Z. Liu, H. Wang, L. Weng, and Y. Yang, "Ship rotated bounding box space for ship extraction from high-resolution optical satellite images with complex backgrounds," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 8, pp. 1074–1078, 2016.

[23] H. Zhu, X. Chen, W. Dai, K. Fu, and Q. Ye, "Orientation robust object detection in aerial images using deep convolutional neural network," *IEEE International Conference on Image Processing*, vol. 13, pp. 3735–3739, 2015.

[24] K. Liu and G. Mattyus, "Fast multiclass vehicle detection on aerial images," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1938–1942, 2015.

[25] S. Razakarivony and F. Jurie, "Vehicle detection in aerial imagery: a small target detection benchmark," *Journal of Visual Communication and Image Representation*, vol. 34, pp. 187–203, 2016.

[26] C. Benedek, X. Descombes, and J. Zerubia, "Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 33–50, 2012.

[27] J. Ding, B. Xiang, Y. Wen et al., "Object detection in aerial images: a large-scale benchmark and challenges," 2021, https://arxiv.org/abs/2102.12219.

[28] A. Zaïdi, "Accurate IoU computation for rotated bounding boxes in $${\mathbb {R}}2$$ and $${\mathbb {R}}3$$," *Machine Vision and Applications*, vol. 32, no. 6, 2021.

[29] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: towards realtime object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[30] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, and I. Reid, "Generalized intersection over union: a metric and a loss for bounding box regression," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 658–666, IEEE, Long Beach, CA, USA, June 2019.

[31] L. Liu, Z. Pan, and B. Lei, "Learning a rotation invariant detector with rotatable bounding box," 2017, https://arxiv.org/abs/1711.09405.

[32] C. Yao, B. Xiang, and L. Wenyu, "Detecting texts of arbitrary orientations in natural images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1083–1090, IEEE, Providence, RI, USA, June 2012.

[33] H. Freeman and R. Shapira, "Determining the minimum-area encasing rectangle for an arbitrary closed curve," *Communications of the ACM*, vol. 18, no. 7, pp. 409–413, 1975.

[34] M. Shamos, *Computational geometry*, Ph.D. thesis, Yale University, New Haven, CT, USA, 1978.

[35] H. Toussaint, "Solving geometric problems with the rotating calipers," in *Proceedings of the IEEE MELECON*, Canada, CA, USA, May 1983.

[36] D. Dimitrov, C. Knauer, K. Kriegel, and G. Rote, "On the bounding boxes obtained by principal component analysis," in *Proceedings of the European Workshop on Computational Geometry*, pp. 193–196, March 2006.

[37] D. Dimitrov, C. Knauer, K. Kriegel, and G. Rote, "Bounds on the quality of the pca bounding boxes," *Computational Geometry*, vol. 42, no. 8, pp. 772–789, 2009.

[38] A. Ramirez and C. Cox, "Improving on the range rule of thumb," *Rose Hulman Undergraduate Mathematics Journal*, vol. 13, no. 2, 2012.

[39] S. P. Hozo, B. Djulbegovic, and I. Hozo, "Estimating the mean and variance from the median, range, and the size of a sample," *BMC Medical Research Methodology*, vol. 5, no. 1, 2005.

[40] Z. Yi, K. Zhao, J. Sun, L. Wang, K. Wang, and Y. Ma, "Prediction of the remaining useful life of supercapacitors," *Mathematical Problems in Engineering*, vol. 2022, Article ID 7620382, 8 pages, 2022.

[41] M. Elloumi, H. Gassara, and O. Naifar, "An overview on modelling of complex interconnected nonlinear systems,"

Mathematical Problems in Engineering, vol. 2022, Article ID 4789405, 18 pages, 2022.

[42] H. Sun, J. Sun, K. Zhao, L. Wang, and K. Wang, "Data-driven ica-bi-lstm-combined lithium battery soh estimation," *Mathematical Problems in Engineering*, vol. 2022, Article ID 9645892, 8 pages, 2022.

[43] X. Wan, W. Wang, J. Liu, and T. Tong, "Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range," *BMC Medical Research Methodology*, vol. 14, no. 1, 2014.

[44] D. Papadopoulos, J. R. R. Uijlings, F. Keller, and V. Ferrari, "Extreme clicking for efficient object annotation," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4940–4949, IEEE, Venice, Italy, 2017.

[45] A. Zaïdi, "Robust semi-automatic annotation of object data sets with bounding rectangles," 2021.

Retraction

Retracted: Mathematical Methods for Sensitive Information Mining Method of News Communication Platform Based on Big Data IOT Analysis

Mathematical Problems in Engineering

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] C. Jiang and Z. Yang, "Mathematical Methods for Sensitive Information Mining Method of News Communication Platform Based on Big Data IOT Analysis," *Mathematical Problems in Engineering*, vol. 2022, Article ID 1606183, 7 pages, 2022.

Research Article

Mathematical Methods for Sensitive Information Mining Method of News Communication Platform Based on Big Data IOT Analysis

Chao Jiang [iD] and ZhiXian Yang

School of Artificial Intelligence, Zhengzhou Railway Vocational & Technical College, Zhengzhou 451460, China

Correspondence should be addressed to Chao Jiang; 19402168@masu.edu.cn

It is urgent to effectively monitor the public opinion of the news communication platform. The platform designed in this paper takes microblog public opinion as the research goal, uses MongoDB to build a distributed computing platform for sensitive information of news communication platform, establishes a corpus of sensitive event topics, introduces PageRank algorithm to deal with microblog social relations, obtains the characteristics of sensitive information of news communication platform, and carries out information screening, so as to accurately screen and mine the keywords in high impact information. To ensure the practical application effect of sensitive information mining method of news communication platform based on big data analysis. Finally, the experiment proves that the sensitive information mining method of news communication platform based on big data analysis has the advantages of high timeliness and high accuracy, which fully meets the research requirements. This is fully in line with the requirements of the study.

1. Introduction

With the popularization of the Internet and the improvement of netizens' sense of social responsibility, the network public opinion broke out a huge vitality that cannot be ignored. It is the public's strong influence and tendentious views on some hot issues in real life. Generally speaking, sensitive information is mainly composed of four parts, which are sensitive words, the related words of sensitive words, the degree of correlation between them and the association rules between them [1]. At present, sensitive information mining technology mainly uses association analysis and cluster analysis to obtain sensitive information related to sensitive words. The application range of correlation analysis technology is relatively wide, and the development speed is fast. Association analysis technology mainly includes two parts: association words and association rules [2]. Clustering analysis technology is mainly to find the text information of related topics, so as to realize the

monitoring of topics and achieve the purpose of topic tracking [3]. For big data analysis, this paper aims to establish a big data platform in which IoT and smart devices can work together to collect the data. Therefore, the final objectives of this paper are to utilize the developed big data communication platform enable quick information collection and real-time feedback, to aggregate and analyze the data collected through repeaters, and to utilize structured databases in the form of big data. The application process of this technology mainly includes three steps. The first step is feature extraction, which mainly refers to filtering the information after the input of information, obtaining the feature vector of the sample, and finally obtaining a matrix; the second step is text clustering, which mainly refers to clustering the results of feature extraction, which can obtain a matrix reflecting all the features in the n-dimensional space. The final step is to select the classification threshold, which mainly refers to the determination of the threshold after the clustering spectrum is obtained, and then the

classification scheme can be directly obtained, in order to ensure the effectiveness of sensitive information mining method based on big data analysis.

2. Sensitive Information Mining Method of News Communication Platform

2.1. Information Filtering Algorithm of News Communication Platform. Network information is complex and diverse, and there are many kinds of good information to serve the public. At the same time, some reactionary, superstitious, violence, and other sensitive information pose a serious threat to social and public security [4]. Therefore, to carry out Internet public opinion information mining requires not only to identify the hot topics concerned by the public from the Internet public opinion information, but also to analyze whether the public's attitude towards an event is positive or negative. In addition, deep-seated public opinion information mining requires good control of negative information dissemination, timely discovery, and disposal of sensitive information, in order to prevent it from causing serious harm to the government, enterprises, individuals, and so on [5]. In the process of traditional topic detection, to judge the similarity between a report and a topic, we need to calculate the similarity between the report and each report in the topic cluster. When the size of topic cluster is large, the number of comparisons will increase exponentially, which will affect the processing speed [6]. To solve this problem, the center vector is used to represent a topic. At this time, we only need to calculate the similarity with the center vector, which improves the effectiveness of topic discovery.

By analyzing the public opinion information such as network news and forum posts, the existing forms and the structural characteristics of network text information, this paper puts forward the idea of "text reconstruction", that is to say, the representative information of the topic is gathered together to form a "theme block", and the remaining part forms a "content block". News headlines contain a lot of classified information, which can let the public know the basic situation of the event in the briefest prompt and evaluation, and guide the public to further read [7]. It is a high generalization of the content of the web page, and has a high accuracy when used to classify news web pages [8]. Topic is the supporting point of Title Construction, and the title of the core event under the same topic is the same or similar to that of its related follow-up reports. Title Information has significant ability to distinguish topics in topic detection, but with the continuous development and change of events, the topic center drifts, and the titles of subsequent reports also change.

The first paragraph of the news web page is a supplement to the title, which is a general description of the event, including the time, place, event, which people or units are involved and so on. It makes a great contribution to the classification [9]. "Universal" emphasizes the statistical characteristics of public opinion information, a single web page cannot be regarded as public opinion, many web pages about a topic and the participation of many Internet users can become network public opinion information [10, 11]. In this sense, network public opinion is accompanied by many news pages, BBS/forums, blog, many netizens browse or comment on a certain topic. We can say that the topic spread by multiple media and concerned by multiple netizens is a hot topic.

2.1.1. Construction of Sensitive Information Database of News Communication Platform. From the header of the file I record, you can find the first attribute address in the file *i* record body, followed by two flag words (flags), where the first bit represents the file deletion flag, and the second bit represents the directory 1 division flag (normal directory). When reading the disk information, make a suitable way to read the information according to the two flag bits. If it is a normal file or directory, continue the following reading operation [12]. Otherwise, you need to recover the file first, and then filter the recovered file to get the required file information, and then hand it to the following text information extraction module for processing.

As can be seen from the figure, a file may contain multiple attributes. A complete file needs all the attributes in the file records to be combined according to certain rules. Therefore, when reading the disk file information, all the attributes in the file records need to be read into the memory according to the order in the file records [13]. Each attribute has its specific content, that is, the operation of each attribute is different. The serial number of the file *i* recorded in the main file table NFT starts with 0, and the files *i* recorded from 0 to 16 belong to platform files, or metafiles, which are mainly used to store the metadata of the platform. These metafiles are transparent to users and are hidden files [14]. The difference between 16 files and other files and directories is that they have a unique fixed address in the MFT table, while other files and directories can be stored anywhere in the table.

The content part starts with the attribute name, and then defines whether the attribute is resident or nonresident. If it is the former, then the attribute value is the content of the attribute; conversely, if it is a nonresident attribute, then the flow of the attribute will be stored in one or more runs. For the sake of simplicity, the storage running area is continuous on the logical cluster number [15]. A run table is stored after the file attribute name, through which the run date table can access the run table belonging to the attribute. The purpose is to calculate the similarity between the text information extracted from the new web page and the existing text cluster, and taking into account the life cycle of public opinion information; its importance decreases with the loss of time [16]. That is to say, the same keyword in different time intervals is likely to represent different meanings, so we add the calculation of time interval [17, 18]. For example, after the new text information is calculated in the time interval, the smaller the similarity value is, the higher the possibility that it is a new event is, and the higher the score is. The expression is shown in formula (1):

$$\text{score}(x) = \max TT_{(t)} \left\{ \left(1 - \frac{k}{m} \right) \times \text{sim}(\vec{x}, \vec{c}_1) \right\}. \quad (1)$$

In the formula, $\overrightarrow{x}$ is the new file information, $\overrightarrow{c}_1$ is the first cluster in the time interval, i is the number of files in the time interval, and k is the number of files added between the latest file collection time in cluster $\overrightarrow{c}_1$ and the arrival time of the new file $\overrightarrow{x}$. In the case of setting the closed value, as long as the score is greater than the set value, the new file is considered to be a new topic. Based on this, the data processing steps of the news communication platform are optimized as follows in Figure 1:

The design of network public opinion monitoring platform mainly includes three modules: text preprocessing module, sensitive information analysis module, and public opinion analysis module. The text preprocessing module mainly includes two steps: Chinese word segmentation and information filtering. Among them, Chinese word segmentation mainly transforms the irregular key text obtained by the platform to form a sensitive word set, and then further processes the word set to obtain the corresponding associated word set. When using word segmentation tools for word segmentation, its speed is relatively fast, and has high efficiency. After inputting the original text information, through the process of Chinese word segmentation, filtering meaningless words, calculating word frequency, scoring feature items and so on, the feature vector of the sample is finally obtained, and the final output of this step is a matrix. Whether the feature selection is good or bad will greatly affect the later analysis. Through text clustering, we can get the distance between these sample points which can reflect the n-dimensional space. The output result of clustering algorithm operation is a clustering pedigree graph, which can generally reflect all the classification situations, or directly give a specific classification scheme, including a total of several categories, the specific sample points in each cluster, etc. After getting the cluster pedigree, we need to choose the appropriate threshold. After determining the ill value, the platform can directly see the classification scheme through the existing clustering pedigree.

2.1.2. Realization of Sensitive Information Mining in News Communication Platform.

Propensity analysis of network public opinion is essentially to distinguish the network text information and determine whether it belongs to the positive category or the negative category. The main process of classifier construction is to use the word sequence check suffix tree representation model to calculate, get the similarity calculation results in the feature space, and use support vector machine algorithm to find the optimal classification hyperplane, so as to achieve the purpose of accurate judgment of network information public opinion tendency. Different from structured data, there are polysemy and polysemy in Chinese [19, 20]. At the same time, the context understanding of sentences brings challenges to public opinion information monitoring, which needs the support of corresponding natural language processing technology.

The network public opinion monitoring based on the network information extraction and semantic analysis technology cannot understand the deeper semantics, can only stay in the stage of passive monitoring of network
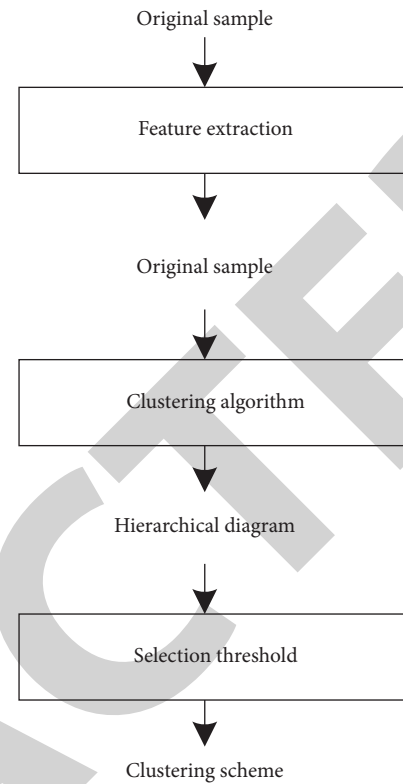

FIGURE 1: Optimization of data processing steps of news communication platform.

public opinion, and cannot realize the automatic identification of network hot spot information, or track the discovered public opinion information. With the development of natural language processing, data mining and other technologies, especially the wide application of search engine, we can efficiently organize the originally scattered information together through the analysis of its relevance [21, 22]. Build a sensitive information knowledge thesaurus, through the analysis of users' concerns about sensitive words, for the detection documents, infer the relationship of sensitive information in the knowledge base, determine the query conditions of sensitive words, submit to the search engine, build the basic analysis set, carry out sensitivity analysis, get the sensitivity evaluation of sensitive words, and issue early warning according to the analysis results. Through the analysis of Web usage records, web structure information and web content information, some quantitative indexes of public opinion are given for decision makers to use, as shown in Table 1.

The basic task of information collection layer is to collect the rich and various public opinion information from the web pages with various data formats. It provides the required data for the public opinion information mining layer, and is the premise of public opinion deep mining. It is the main task of the Internet public opinion information mining layer to carry out in-depth mining of public opinion information, find the hot issues of public concern, analyze the attitude of the public, and deal with the sensitive information that constitutes harm. By analyzing the data provided by the

TABLE 1: Evaluation index system of Internet public opinion.

Class a	Second level	Level 3
Dynamic changes of public opinion	Information variability	Total number of information items under topic
		Information growth rate under topic
	Timeliness index	Initial item information release time
		Last update time of item information
		Aging parameters
	Signature degree	Signature of information
	Publisher influence	Measurement of publisher influence
	Diffusivity	The change of flow and distribution
	Network distribution	characteristics of information in time
Public opinion information mining	Attention	Change of published papers
		Review changes
		Response change
	Content sensitivity	Over sensitive information results
	Attitude tendentiousness	Propensity analysis results

public opinion information collection layer, it can detect network topics, analyze people's attitudes, monitor network-sensitive information, evaluate public opinion situation, etc., and provide objective basis for the public opinion information service layer to serve relevant departments.

The sensitive information analysis module mainly includes three ways, namely, association analysis, cluster analysis, and feature extraction. Among them, the common algorithm of association rule mining is Apriori algorithm. The application of this algorithm must first form all frequent item sets, and then form all credible association rules from these frequent item sets. The most important feature of this algorithm is to start from the single item, and then filter it layer by layer, so as to get an effective item set, effectively avoiding the search for impossible items. The public opinion analysis module mainly provides two functions, which are the discovery of public opinion hot spots and the tracking of network public opinion topics. Among them, the application of hot spot discovery can let users know the current hot topics in time, and comprehensively grasp the current network public opinion information. In the process of hot spot discovery, the public opinion monitoring platform mainly obtains information, frequently searched words, browsed web pages, forum replies, and other relevant information based on the keywords entered by users, then monitors the hot spots, and automatically identifies the "hot spot" information in the network, thus forming hot spot alarm [23–25]. Topic tracking in public opinion monitoring platform is mainly realized by topic tracking method, which mainly forms tracking expression expressed by query vector from training set, and then uses this tracking expression to judge the newly captured web page information, and finally obtains the information related to the current topic [26, 27].

3. Experimental Analysis

If the threshold is set too low, it will lead to the separation of reports on the same topic. If the threshold is set too high, it will make the topic larger and introduce a lot of irrelevant reports. In the research of Chinese text orientation analysis, there is no open corpus at present https://www.drip.com 16000 comments were collected as experimental data, and

emotion categories of the text were manually labeled, including 8000 positive category documents and 8000 negative category documents. A total of 9804 documents are divided into 20 categories. Among them, there are no more than 100 documents in 11 categories such as literature and education, and more than 1000 documents in 6 categories such as computer, environment, agriculture, economy, politics, and sports. Because the training process of genetic algorithm needs a large number of samples, we only select 6 categories with more than 1000 documents. At the same time, because the algorithm will eventually be applied to information filtering, the project team collected 276 and 192 documents of violence and pornography, respectively. As a result, there are 7947 documents in 8 categories. The distribution of training documents is shown in Table 2:

In order to set a reasonable threshold, ten experiments were carried out. In each experiment, firstly, 10 reports were randomly selected from 8 topics in data set 1 to form the original report set; secondly, the reports corresponding to the original report set were selected from data set 2 to form the reconstructed report set; Then, each report in the original report set and the theme block of each report in the reconstructed report set are segmented, feature selected, and weighted, and each topic in the two report sets is represented as a central vector; Finally, cosine similarity is used to calculate the similarity of each report to its topic and other topics in the report set. This experiment uses eight topics to measure the performance of traditional single-pass clustering algorithm and hierarchical topic detection algorithm in topic detection. The experiments were conducted five times and the performance was evaluated by the cost of testing. Through the average measurement of five experiments, the two methods can identify the topic well under the similar threshold settings of different topics, but the detection costs are different, and the change curve is shown in Figure 2.

It can be seen from the figure that when the topic threshold is given, the detection cost of hierarchical topic detection algorithm is lower than that of traditional single-pass topic detection algorithm, which indicates that the former has better topic detection ability. With the different topic similarity threshold setting, the detection cost

TABLE 2: Sample distribution of training data.

Category	Violence	Pornographic	Computer	Environmental science	Agriculture	Economics	Politics	Sports
Number of documents	276	192	1358	1218	1022	1601	1026	1254

fluctuates up and down. When the topic similarity threshold rises from 0.24 to 0.30, the detection cost decreases. When the topic similarity threshold is greater than 0.30, the false detection rate changes to miss detection trend, and the detection rate is improved. Therefore, when the similarity threshold is 0.30, the detection cost is the least and the topic detection performance is the best. In the experiment, the topic threshold BR is set as 0.30, and the subtopic threshold value is in the range of 0.4 to 0.6. A hierarchical subject detection algorithm is used to test the subject. This method detects the topic and identifies five subtopics, and detects the same number of subtopic reports as the manual identification reports, as shown in Figure 3.

By comparing the detection results of similarity mining effect of sensitive information, we can know the discovery that the number of reports of subtopics identified by the hierarchical topic detection algorithm is roughly the same as that of manually annotated subtopics, indicating that the method can distinguish molecular topics and present the hierarchical structure of topics to a certain extent and the accuracy of the proposed method in each category of test data 1.

Through analysis, we can find that there are some similarities between the two categories with poor classification effect. For example, the political category often contains economic, environmental, agricultural, and other factors, resulting in the low accuracy. In the above experimental data in Table 3, the improved calculation method can achieve better results. However, we cannot rule out that the above experimental results are obtained on the basis of data 1, and there may be some overfitting problems. Therefore, the above second set of test data is used for further test, and the analysis data are as follows:

Among the above experimental data, in terms of accuracy in Table 4, although the computer finance and closed test have a slight decline, there is little difference, while the sports class has a big gap. After analyzing the training documents and test documents, we can find that the sports-related documents in the original training documents belong to sports theory research, while the test documents come from the network, so there is a big difference between the distance. In view of the purpose of the research which is to apply to content-based information filtering, this experiment is designed to apply the above classifier to the test experiment of network-sensitive information filtering. In the experiment, test data 1 is divided into two categories, legal documents and illegal documents. The illegal documents are composed of pornographic and violent documents in test data 1, while the legal documents are randomly selected from the other six categories. The experimental data composition and test results are as follows.

In the experimental data shown in Table 5, in the two numbers before and after each table item, the former one
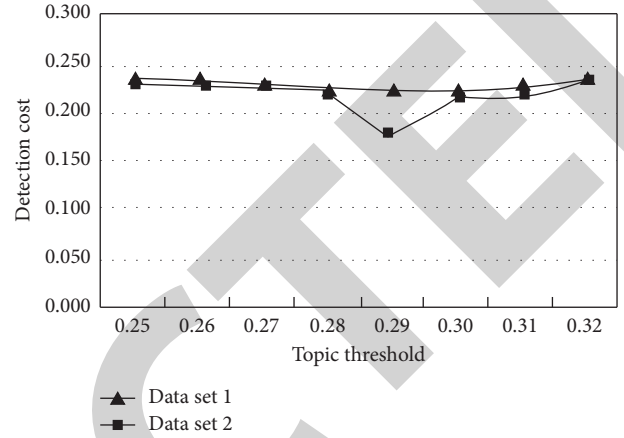

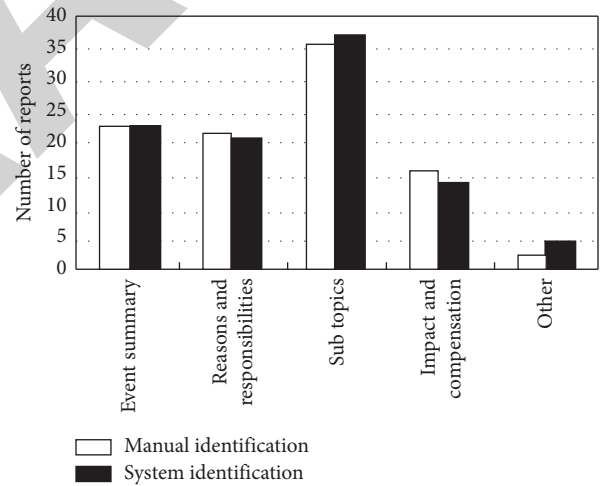FIGURE 2: Threshold detection of sensitive information similarity mining.

FIGURE 3: Contrast detection of similarity mining effect of sensitive information.

uses the template generation method based on dynamic genetic algorithm to generate the template, but does not use the weight calculation method in this paper to calculate the filtering effect of the filtered information. The latter one uses both the template generation method based on dynamic genetic algorithm and the template generation method based on this paper. The weight calculation method calculates the filtering effect of the filtered information. Simultaneously interpreting the data in the above table with traditional methods, the proposed method is obviously better than the traditional method in terms of the accuracy of illegal information. Because the test data used by traditional methods are completely consistent, the proposed method has better filtering effect.

Table 3: Accuracy of sensitive information processing.

Category	Agriculture	Politics	Sports	Violence
Accuracy	79.969/82.324	74.364/76.549	75.211/81.622	96.053/96.385
Category	Environmental science	Economics	Computer	Pornographic
Accuracy	83.345/87.936	91.585/91.762	87.468/91.763	98.446/99.045

Table 4: Comparison of data mining accuracy.

	Sports	Economics	Computer
Accuracy	46.154/56.672	90.697/91.387	84.314/88.946

Table 5: Statistics of sensitive information filtering effect test.

	Number of documents	Effective filtering	Accuracy
Illegal	300	293/295	97.67/ 98.33
Legitimate	300	261/277	87.00/ 92.33

4. Conclusions

With the increase of the number of network users, the network environment has become more complex. Therefore, the establishment of network public opinion monitoring platform is particularly important. On the basis of sensitive information mining, relevant key technologies are studied, and the design method of network public opinion monitoring platform is proposed to effectively realize the network public opinion monitoring, realize the maintenance of social stability, and promote government departments to make decisions more democratic and scientific.

Data Availability

The data used to support the findings of the study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. Zhang, Q. Ma, Y. Y. Chiang et al., "Extracting geographic features from the Internet: a geographic information mining framework," *Knowledge-Based Systems*, vol. 174, no. 15, pp. 57–72, 2019.

[2] J. Li, S. Liu, Z. Wang, and J Guo, "Neutrosophy theory based visualization report of sports news data[J]," *International Journal of Electrical Engineering Education*, vol. 7, no. 28, pp. 125–132, 2020.

[3] Y. P. Singh and P. Gautam, "Development of data mining algorithm for giving the loan in banks introduction[J]," *High Technology Letters*, vol. 27, no. 1, pp. 284–295, 2021.

[4] Z. Thomas, "Enriching metadata in an international agro-forestry research repository with funding-related information by mining grey literature - ScienceDirect[J]," *Procedia Computer Science*, vol. 146, no. 10, pp. 112–122, 2019.

[5] A. Subahi and G. . Theodorakopoulos, "Detecting IoT user behavior and sensitive information in encrypted IoT-app traffic," *Sensors*, vol. 19, no. 21, pp. 4777-4778, 2019.

[6] S. S. D. Mitchell, "'Warning! You're entering a sick zone': the construction of risk and privacy implications of disease tracking apps[J]," *Online Information Review*, vol. 43, no. 6, pp. 1046–1062, 2019.

[7] Z. Li and J. Wang, "Security storage of sensitive information in cloud computing data center[J]," *International Journal of Performability Engineering*, vol. 15, no. 3, pp. 1023–1032, 2019.

[8] X. Lu, L. Cao, and X. Du, "Dynamic control method for tenants' sensitive information flow based on virtual boundary recognition," *IEEE Access*, vol. 8, no. 10, pp. 162548–162568, 2020.

[9] M. J. Y. Suh, H. J. Yi, H. J. Kim, and S. H. Kim, "Is asymmetric hearing loss a risk factor for vestibular dysfunction? Lesson from big data analysis based on the Korean national health and nutrition survey[J]," *Otology & Neurotology*, vol. 40, no. 10, pp. 1339–1345, 2019.

[10] M. Tolani, S. R. K. Sunny, and R. K Singh, "Lifetime improvement of wireless sensor network by information sensitive aggregation method for railway condition monitoring," *Ad Hoc Networks*, vol. 87, no. 11, pp. 128–145, 2019.

[11] B. Buddhadev, P. Faruki, M. S. Gaur, S. Kharche, and A Zemmari, "FloVasion: towards detection of non-sensitive variable based evasive information-flow in android apps," *IETE Journal of Research*, vol. 35, no. 2, pp. 1–15, 2020.

[12] K. Nagaraj, S. Gs, and A. Sridhar, "Encrypting and preserving sensitive attributes in customer churn data using novel dragonfly based pseudonymizer approach," *Information*, vol. 10, no. 9, p. 274, 2019.

[13] M. Barika, S. Garg, A. Y. Zomaya, L. Wang, A. V. Moorsel, and R. Ranjan, "Orchestrating big data analysis workflows in the cloud: research challenges, survey, and future directions [J]," *ACM Computing Surveys*, vol. 52, no. 5, pp. 1–41, 2020.

[14] G. Hong, "The development and application of new media technology in news communication industry[J]," *International Journal of Electrical Engineering Education*, vol. 2, no. 24, pp. 64–80, 2021.

[15] S Jesus, C Jenny, V. V Jesus, R Ligia, S Darwin, and F Claudia, "Preservation of confidential information privacy and association rule hiding for data mining: a bibliometric review - ScienceDirect[J]," *Procedia Computer Science*, vol. 151, no. 5, pp. 1219–1224, 2019.

[16] N. Kumar, G. Akangire, B. Sullivan, K. Fairchild, and V Sampath, "Continuous vital sign analysis for predicting and preventing neonatal diseases in the twenty-first century: big data to the forefront," *Pediatric Research*, vol. 87, no. 2, pp. 210–220, 2019.

[17] M. B. Alazzam, F. Alassery, and A. Almulihi, "Diagnosis of melanoma using deep learning," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–9, Article ID 1423605, 2021.

[18] M. Gonalves Pinho and F. A. BraganaM, "Psychotic disorders hospitalizations associated with cannabis abuse or dependence: a nationwide big data analysis[J]," *International*