Research Article

A New Classification Method in Ultrasound Images of Benign and Malignant Thyroid Nodules Based on Transfer Learning and Deep Convolutional Neural Network

Weibin Chen¹, Zhiyang Gu², Zhimin Liu³, Yaoyao Fu⁴, Zhipeng Ye⁴, Xin Zhang⁴, and Lei Xiao¹

¹College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou, China

²Department of Electrical and Electronic Engineering, Wenzhou Polytechnic, Wenzhou, China

³Wenzhou Hospital of Traditional Chinese Medicine, Wenzhou, China

⁴School of Ophthalmology & Optometry, School of Biomedical Engineering, Wenzhou Medical University, Wenzhou, China

Correspondence should be addressed to Zhiyang Gu; guzy@mail.nwpu.edu.cn and Lei Xiao; xiaolei@wzu.edu.cn

Received 11 April 2021; Revised 10 August 2021; Accepted 21 August 2021; Published 13 September 2021

Academic Editor: Shahzad Sarfraz

Copyright © 2021 Weibin Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Thyroid nodule is a clinical disorder with a high incidence rate, with large number of cases being detected every year globally. Early analysis of a benign or malignant thyroid nodule using ultrasound imaging is of great importance in the diagnosis of thyroid cancer. Although the *b*-mode ultrasound can be used to find the presence of a nodule in the thyroid, there is no existing method for an accurate and automatic diagnosis of the ultrasound image. In this pursuit, the present study envisaged the development of an ultrasound diagnosis method for the accurate and efficient identification of thyroid nodules, based on transfer learning and deep convolutional neural network. Initially, the Total Variation- (TV-) based self-adaptive image restoration method was adopted to preprocess the thyroid ultrasound image and remove the boarder and marks. With data augmentation as a training set, transfer learning with the trained GoogLeNet convolutional neural network was performed to extract image features. Finally, joint training and secondary transfer learning were performed to improve the classification accuracy, based on the thyroid images from open source data sets and the thyroid images collected from local hospitals. The GoogLeNet model was established for the experiments on thyroid ultrasound image data sets. Compared with the network established with LeNet5, VGG16, GoogLeNet, and GoogLeNet (Improved), the results showed that using GoogLeNet (Improved) model enhanced the accuracy for the nodule classification. The joint training of different data sets and the secondary transfer learning further improved its accuracy. The results of experiments on the medical image data sets of various types of diseased and normal thyroids showed that the accuracy rate of classification and diagnosis of this method was 96.04%, with a significant clinical application value.

1. Introduction

In recent years, the incidence of thyroid cancer has continued to rise. As a malignant tumor of the head and neck, it continues to threaten people's health [1]. It is reported that, in the United States, thyroid carcinoma is expected to be the third disease threat among women, with approximately 37 cases per 100,000 people [2]. The thyroid nodule is a symptom of thyroid-related disease. The nodule may be caused by the growth of thyroid cells or thyroid cyst. The thyroid tissues around the scattered lesion of thyroid nodule

can be clearly distinguished through images [3, 4]. If the benign and malignant nodules can be judged earlier, even malignant nodules can be cured. In addition, the accurate distinguishing methods can provide an effective basis for the proper subsequent clinical treatments. Besides, the accurate diagnosis earlier can also reduce the medical risk to be suffered by patients and a large amount of health care costs caused by acupuncture detection.

Currently, there are two major methods for examining the nature of thyroid nodules: ultrasound image analysis and computer tomography imaging analysis. Between

them, ultrasound imaging is cheap and common in hospital. This is why the ultrasound image analysis is more common. However, in the ultrasound image, the malignant thyroid nodule with prominent histopathological components and blurred boundaries usually adhere to other tissues, difficult to distinguish the morphology. This requires an efficient image classification method to improve accuracy and reduce the misdiagnosis rate. In the past studies, radiographers have summarized thyroid nodules ultrasonographic features according to their characteristics, which function as signs of cancer. However, thyroid nodules diagnosis relying on these characteristics is time-consuming and poorly robust. To this end, the accurate computer aided diagnosis system based on ultrasound images is still to be judged by doctors. A fully automatic computer aided diagnosis system consists of image preprocessing, such as denoising, ROI extraction, and classification. Nowadays, most researches mainly focus on image denoising and ROI extraction. At present, it is still difficult to judge by ultrasonic images alone. The low quality and noise pollution of the ultrasound image makes it extremely challenging to classify it. Tsantis et al. [5] proposed an SVM classifier to divide thyroid nodules into high-risk and low-risk malignant tumors. Ma et al. [6] present a noninvasive and automatic approach for differentiating benign and malignant thyroid nodules based on support vector machines (SVM). Acharya et al. [7] proposed a wavelet transform filter to classify it. Shukla et al. [8] utilize artificial neural network dealing with thyroid disease. Prochazka et al. [9] proposed a dual threshold binary decomposition method to classify it.

Rapid progress in the automatic classification of medical image data is also made by this method. Thandiackal et al. [10] identified skin lesions through some pretrained classical classification networks. Convolutional Neural Network (CNN) models are a type of deep learning architecture introduced to achieve the correct classification of breast cancer [11]. It proposed an in-depth model that uses limited chest CT data to distinguish malignant nodules and benign nodules [12–15]. It proposed a classification algorithm for thyroid nodule ultrasound images based on DCNN [16]. Nevertheless, these methods are defective in the following aspects at present:

Needless to say, transfer learning has played an important role in ultrasound imaging diagnosis of thyroid cancer. However, few-shot learning is the challenging problem of making predictions based on a limited number of samples [17–20]. Also, data labelling is a task that requires a lot of manual work [21–23]. Finally, the inappropriate model and imbalanced training data are difficult to get better classification accuracy [24–30].

Therefore, in view of the above problems, this paper discusses and does the following work:

In response to the abovementioned few-shot learning, in this paper, a TV model is introduced for the automated preprocessing of original data collected by various institutions. Some image marks made by doctors also need to be removed. The original image is then expanded by data

augmentation for the purpose of supplementing inadequate training samples. Also, in response to what is mentioned above to select suitable learning transfer model, the GoogLeNet model was established for the experiments on thyroid ultrasound image data sets. The results showed that the model enhanced the accuracy for the nodule classification. Finally, in response to the above imbalanced training data problem, this paper puts forward the secondary transfer learning conducted on public thyroid database and the actual data sets collected by hospitals, which improves the classification accuracy.

The structure of this paper is presented as follows. In Section 1, the writing motivation is given, and the relevant literature has been examined. Section 2 provides the traditional CNN structure and describes the Tv-Based Image Restoration. Section 3 describes the network structure of the proposed methods based on GoogLeNet. Section 4 shows the experimental results, including the application of the proposed method for diagnosis of thyroid cancer. The research results are summarized in Section 5.

2. Related Work

2.1. Abbreviations and Acronyms. This kind of network structure is usually called CNN, local connection, weight sharing, and other characteristics of feed-forward neural network [31]. It then inputs the extracted features into the fully connected network; thereby, the parameters are to be optimized. In their research, Moon et al. used ultrasound images for cancer diagnosis. The difference from the previous method is that they use a variety of data sets and combine different CNN algorithms for fusion diagnosis. It was found that the accuracy rates of different data sets were 91.1% and 94.62%, respectively [32]. Kim et al. used deep learning methods for intelligent diagnosis of breast ultrasound images. By calculating different performance standards, the AUC value was 89% [33]. There are also some methods that use a three-dimensional convolutional neural network structure. Through experiments, different performance standards have been found, and the accuracy rate can reach 96.7%.

It is an effective method to extract image features. The image input to the convolutional layer. In this layer, it can perform feature extraction tasks. Each feature can be extracted from each feature map through the convolutional layer. The weight is updated through continuous backward propagation during the training. The computing formula of the convolution layer is

$$X_j = f\left(\sum_i x_i k_{ij} + b\right), \quad (1)$$

where X_j is the output neuron cell, x_i is the input signal of each network cell, f is the activation function, K_{ij} is the convolution kernel, and b is the offset.

After the features are extracted through the convolution layer, the output feature map reaches the pooling layer for feature selection and information screening. The pooling formula is shown in the following equation:

$$X_j = f(\beta_j L(x_j) + b), \quad (2)$$

where β_j is the weight coefficient and L is the sampling function [34–36].

After pooling, the data is input to the fully connected layer that is equivalent to the traditional forward neural propagation. The connected end of convolutional neural only transmits signals to other fully connected layers. The traditional CNN Structure is shown in Figure 1.

In the traditional CNN structure, the forward propagation is adopted to build the network structure, and the backward propagation is adopted to train the network parameters. The loss function, learning rate, and moving average are used for network optimization. Regularization and cross entropy are the loss functions in CNN. The cross entropy formula is shown as

$$H = - \sum y * \log y', \quad (3)$$

where y is the standard answer and y' is the predicted value. The exponential decay learning rate is adopted, i.e., the magnitude of each parameter update. The formula of parameter update is given by

$$w_{n+1} = w_n - r * f', \quad (4)$$

where r is the learning rate and f' is the gradient of the loss function.

2.2. TV-Based Image Restoration. The data sets collected for this experiment were few and needed to be augmented. In the present study, the data set was augmented only by rotation and translation.

The current data contained manual marks, as shown in Figures 2(a) and 2(c). Manual marks mainly refer to the marks made by the professionals on the lesion area in the ultrasound image, which destroy the part of the texture and affect the accuracy and integrity of the image of the area to be analyzed. This also impacts the subsequent training. Therefore, restoring the image was essential. In 2002, Shen et al. [37] extended the TV model to image inpainting and proposed an image inpainting method based on the TV model. The Total Variation- (TV-) based self-adaptive image restoration was adopted for images to estimate the value after pixel restoration:

$$G_O^{(n)} = \sum_{p \in A} H_{Op}^{(n-1)} G_p^{(n-1)} + H_{OO}^{(n-1)} G_O^0, \quad (5)$$

where G_O represents the pixel of the current point O to be restored, G_p represents the pixel of neighboring points of the current point O at four directions, H_{Op} ($p \in A$, $A = \{N, S, W, E\}$), and H_{OO} is the weight coefficient, which was mainly determined by W_p . It is defined in

$$W_p = \frac{1}{\sqrt{a^2 + |\nabla g_p|^2}}, \quad p \in A, \quad (6)$$

$$H_{Op} = \frac{W_p}{\sum W_p + \lambda(O)},$$

$$H_{OO} = \frac{\lambda(O)}{\sum W_p + \lambda(O)},$$

where ∇g_p is the divergence;

$$\nabla g_e \approx \frac{1}{h} \sqrt{(G_W - G_O)^2 + \left[\frac{(G_{NW} - G_{SW})}{2} \right]^2}, \quad (7)$$

where G_W , G_{NW} , and G_{SW} are the pixels of the left neighboring point, the upper left neighboring point, and the lower left neighboring point of the current pixel; $\lambda(O)$ is the parameter of λ at point O .

Finally, as shown in Figure 2, the image was well restored to an extent that its texture was similar to the surrounding texture. The same method was applied to restore the pixel of Figures 2(a) and 2(c), from which Figures 2(b) and 2(d) were obtained.

3. Proposed Methods

3.1. Proposed Network Structure. The CNN model of GoogLeNet was established to realize the diagnosis of thyroid classification. The process is shown in Figure 3. Initially, the TV-based preprocessing was performed for the thyroid nodule image. Subsequently, the training of CNN model was conducted to extract the features of images of various sizes. Thereafter, the transfer learning was implemented based on the open source database and the database actually collected. The features were integrated, and the dual-softmax assisted forward propagation was conducted. In the end, a softmax classifier was adopted to classify features. The diagnosis of thyroid classification was thus completed.

3.2. GoogLeNet CNN Structure. GoogLeNet adopts the structure of the inception proposed in the Going Deeper with Convolutions [38]. Generally, a CNN structure just simply augments the network, with two disadvantages, namely, overfitting and increase in the computation amount. Generally, the network depth and width can be increased by reducing the parameters, while the reduction of parameters turns the full connection to a sparse connection. For the dense matrix optimization mode, the computation amount does not have a qualitative improvement with this kind of change. The inception structure has a sparse structure and high computing performance. The inception structure is shown in Figure 4.

FIGURE 1: Traditional CNN structure.

FIGURE 2: TV-based image restoration.

FIGURE 3: Proposed network model.

The use of various scale convolution kernels can get various sizes receptive fields. The final stitching refers to the integration of various scales. Different kernel sizes were set for alignment, such as 1×1 , 3×3 . Also, the convolution stride = 1 and the pad = 0, 1, 2, respectively, which was directly stitched together later. However, as the use of 5×5 convolution kernel still generated a large amount of computation, hence, the 1×1 convolution kernel was utilized to reduce the dimension. The specifically improved inception structure is shown in Figure 5.

3.3. Improved GoogLeNet Structure. The GoogLeNet network model is stacked based on the Inception module. Being a network with a relatively large given depth, there is a problem with the backward propagation of effective communication gradient through all layers. For this task, the performance of the shallower network shows that the features generated by the intermediate layer of the network should be very discernible. The discriminative ability of classifiers at low stages can be expected to add complex classifiers. It is considered as a method that overcomes the problem of vanishing gradient. It can adopt the forms of small CNN that are placed above the output of the inception module. These auxiliary networks are discarded in

case of inference. The subsequent control experiment results show that the influence of the complex networks is almost the same. One of them is adequate to achieve the same effect.

Dropout determined what percentage of fully connected nodes was shut off for a training cycle. Dropout improved the model generalizability by preventing nodes from overlearning the training data. The average pooling was finally adopted for the network to replace the fully connected layer. Furthermore, in order to prevent the gradient from vanishing, the network was provided with two additional softmax for the forward propagation gradient. The structure of inception is shown in Figure 1. The computation was performed after the number of channels was reduced through the 1×1 convolution to aggregate the information, effectively making use of computing power. The integration of multidimensional features by combining the convolution and pooling of different scales also contributed to a better effect in terms of recognition and classification. By changing the computing power from being deep to being wide, it avoided the problem of dispersion of the training gradient. The global average pooling adopted by GoogleNet solved the typical problem of the complicated and weakly generalized parameters of traditional CNN network.

FIGURE 4: Inception module.

FIGURE 5: Improved inception module.

4. Experiments

4.1. Selected Data Sets and Evaluation Indicators. We verified through a lot of experiments that the accuracy for predicting the morphological classification of candidate star clusters depended on the following characteristics of the training sample:

- (i) Origin of classifications: Primary classification or the mode of two classifiers is shown in Table 1;
- (ii) Size of images used for training: $240p \times 240p$;
- (iii) Using a random selection of 80 percent of the samples described in Table 1 separately, and the remaining 20 percent was reserved for validation.

The thyroid nodule ultrasound image data used was obtained from the hospitals. After the data augmentation, there were 2,763 images of malignant cases and 541 images of benign cases, with a total of 3,304 images. All images were cropped into a size of 240×240 . The images were extracted from the thyroid ultrasound video sequence by the ultrasonic apparatus, at a frequency of 12 MHz. The TI-RADS score was given by a professional physician after the image diagnosis. 3,123 images of cases were used for the training of improved models. 541 images, as a test data set, were then randomly divided into 5 groups to test the above three models. Each of the benign and malignant samples is divided into a verification set, test set, and training set. The specific classification scheme is shown in Table 1. The overall condition is shown in Table 1 below.

4.2. Comparative Analysis of Experimental Results. The comparison of accuracy is important for different models. Table 2 shows that our mean (Improved Inception) was

TABLE 1: Distribution of samples in training and validation of database 1.

	Malignant	Benign	Normal	Total samples
Training	2210	433	500	3143
Validating	553	108	500	1161
Testing	50	20	500	570
Total samples	2813	561	1500	4874

improved in terms of accuracy than the common GoogLeNet model, and it exhibited the highest accuracy rate in determining whether a thyroid nodule changed pathologically.

The confusion matrix and performance standards obtained in LeNet5, VGG16, GoogLeNet, and GoogLeNet (improved) models are shown in Figure 6.

The LeNet5 architecture correctly predicted 860 out of 1000 images and incorrectly predicted 140. The VGG16 architecture correctly predicted 920 out of 1000 images and incorrectly predicted 80. Although the GoogLeNet architecture correctly predicted 960 of the 1,000 pictures, it incorrectly predicted 40 of them. The most successful class of the GoogLeNet (improved) architecture is the ordinary class. The GoogLeNet (improved) architecture correctly predicted 970 out of 1000 images and incorrectly predicted 30.

True Positive Rate (TPR) is shown on the vertical axis of Figure 6, and False Positive Rate (FPR) is shown on the horizontal axis of Figure 6. The entire graph is also called the ROC curve. Figure 6 shows the result of classification accuracy percentage of our proposed algorithm as 96.65%, 97.81%, 97.32%, 95.97%, and 0.97%, respectively.

The loss values of different CNN models are shown in Table 3, and the GoogLeNet (Improved Inception) model was relatively minimal. The change in the trend in

TABLE 2: Classification of testing samples by different models.

Sample no	LeNet5	VGG16	GoogLeNet	GoogLeNet (improved inception)
1	0.5427	0.3326	0.2907	0.2876
2	0.8965	0.5024	0.5041	0.4992
3	0.7129	0.5598	0.3716	0.3218
4	0.6872	0.4901	0.3524	0.3412
5	0.7428	0.4983	0.4033	0.3916
Average	0.7164	0.4766	0.3844	0.3683

FIGURE 6: Confusion matrices obtained in the Softmax classifier (1 is benign, 2 is malignant, and 3 is normal).

TABLE 3: Loss values of different CNN models.

Sample no	LeNet5	VGG16	GoogLeNet	GoogLeNet (improved inception)
1	87.29	93.02	96.31	96.65
2	84.68	91.31	95.83	97.81
3	85.03	90.98	96.47	97.32
4	86.39	92.17	95.96	95.97
5	85.42	91.54	95.63	97.12
Average	85.76	91.80	96.04	96.97

continuous iteration is shown in Figure 7. Table 4 shows the results of the time consumed by the different models to diagnose the same test image.

As shown in Table 5, the LeNet5 model exhibited a shortest time to diagnose the thyroid ultrasound image, and the GoogLeNet model exhibited the second shortest.

4.3. Joint Training and Secondary Transfer Learning. In GoogLeNet, the transfer is from the MINIST data set to the thyroid image. Generally, it is believed that the transfer effect is worse than that of the two similar data sets, when the two data sets have a great difference. MINIST, as a natural image, greatly differs from the medical image. Therefore, the joint data training was conducted herein, based on the public

database, and it was provided by the cooperative organization of this paper. Because of the lack of samples, the joint database was deemed as a whole in the training, which further expanded the overall database.

In transfer learning, the database of small samples was used as the aiming field, and a great quantity of marked database was used as the source domain. In the previous experiments, 2,210 images of malignant and 553 images of benign cases, a total of 3,374 images, were collected from hospitals.

4.4. Analysis of Experimental Results. Table 4 shows the difference in the performance between the secondary transfer learning and the primary transfer, the data joint training, and the VGG16-based system. The results showed

FIGURE 7: TPR/FPR curves.

TABLE 4: Comparison of evaluation index obtained by different methods.

Method	F1 score	Precision	Recall	Accuracy
Twice transfer learning based on LeNet5	0.9748	0.9678	0.9671	0.9649
Once transfer learning based on LeNet5	0.9584	0.9788	0.9685	0.9635
Twice transfer learning based on VGG16	0.9657	0.9718	0.9781	0.9742
Once transfer learning based on VGG16	0.9604	0.9548	0.9685	0.9692
Twice transfer learning based on inception V3 (our method)	0.9874	0.9842	0.9684	0.9885
Once transfer learning based on inception V3 (our method)	0.9674	0.9747	0.9774	0.9745

TABLE 5: Time consumed by different CNN models to diagnose thyroid ultrasound images.

Sample no	LeNet5	VGG16	GoogLeNet	Our method (improved inception)
1	14.65	19.23	15.69	16.70
2	13.97	20.56	16.53	16.87
3	15.56	18.59	17.59	18.47
4	15.55	21.95	18.81	18.91
5	15.13	22.37	17.03	17.05
Average	14.57	20.50	17.09	17.36

FIGURE 8: Performance comparison of system transfer and nontransfer learning based on (a) LeNet5, (b) VGG16, and (c) inception V3.

that, for small medical data sets, the secondary transfer significantly improved the system performance. Figure 8 shows the comparison of system transfer and non-transfer learning based on LeNet5, VGG16, Inception V3.

With $\alpha = 0.05$, the p value of the VGG16 model is less than 0.001, and the t value = -28.71. Because the p value is

less than α , there is enough evidence to reject these invalid hypotheses. The p value of the LeNet model is 0.05 and the t value = -1.66, which shows that there is not enough evidence to reject the null hypothesis $\alpha = 0.05$. However, between LeNet and VGG16, the average value of VGG16 is higher, and the average value of the other two groups of GoogleNet

and GoogLeNet (improved) is higher than that of LeNet and VGG16.

In case of data sets with similar category, the data joint training also showed a close agreement to the experimental result of secondary transfer. The data joint training and secondary transfer were effective in further improving the system performance, while introducing the transfer learning. This provided a reference for the classification of small data sets and medical image data sets.

5. Conclusions

In the present study, the thyroid ultrasound image was preprocessed by the TV-based self-adaptive image restoration method. Subsequently, the CNN model was established using the corresponding loss function, learning rate, moving average, and optimization algorithm set for optimization. Three improved models, namely, the LeNet5 model, VGG16 model, and GoogLeNet model, were trained to diagnose the benign and malignant thyroid nodules. Thereafter, the accuracy rate of each model in terms of diagnosing results was obtained through the tests.

Although all of the three trained models completed the recognition, to verify the best CNN model for diagnosing such ultrasound images, we collected a large amount of image data for training and testing. In the comparison studies, it was found that the GoogLeNet (Improved) exhibited the relatively highest accuracy rate in determining whether a thyroid nodule changed pathologically. The average accuracy rate of the GoogLeNet model was up to 96.04%; furthermore, GoogLeNet (Improved) achieves classification accuracy of 97%, with a loss value of 0.3844. It explains that the GoogLeNet model can diagnose whether the patient's thyroid is in diseased state or is normal. In the end, the data joint training and secondary transfer learning were performed for the open source data sets, and the thyroid ultrasound image data was collected from the hospitals, which further improved the classification accuracy.

In the experiment in this paper, deep learning was applied to the auxiliary medical diagnosis. Our next step is to gradually optimize the model and study the improvement of model, so as to ensure a high accuracy rate of the results. The image classification and diagnosis method based on deep learning will provide a reference to the doctors to diagnose such diseases, help them improve diagnosis efficiency and accuracy, immensely save manpower, and provide new concepts for the ultrasound diagnosis of the thyroid nodules in future.

Data Availability

The data used to support the findings of this study are available from Weibin Chen via e-mail: sun@wzu.edu.cn.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

WB.CHEN conceptualized the study; WB.CHEN and ZY.GU developed the methodology; WB.CHEN and ZY.GU

worked on software; ZM.LIU, YY.FU, and ZP.YE validated the study; XIN.ZHANG carried out formal analysis; L.XIAO investigated the study; L.XIAO helped with the resources; WB.CHEN and L.XIAO wrote, reviewed, and edited the manuscript; XIN.ZHANG visualized the study. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This work was financially supported by Zhejiang Provincial Natural Science Foundation of China under Grant nos. LY21F020001 and LY19F030006, Wenzhou Science and Technology Bureau of China (Wenzhou major scientific and technological innovation project, under Grant nos. ZG2020026 and ZY2019019), and Science and Technology of Wenzhou (Y20180232).

References

- [1] D. Chen, J. Niu, Q. Pan, Y. Li, and M. Wang, "A deep-learning based ultrasound text classifier for predicting benign and malignant thyroid nodules," in *Proceedings of the 2017 International Conference on Green Informatics (ICGI)*, Fuzhou, China, August 2017.
- [2] J. Chi, E. Walia, P. Babyn, J. Wang, G. Groot, and M. Eramian, "Thyroid nodule classification in ultrasound images by fine-tuning deep convolutional neural network," *Journal of Digital Imaging*, vol. 30, no. 3, pp. 477–486, 2017.
- [3] M. Halicek, G. Lu, J. V. Little et al., "Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging," *Journal of Biomedical Optics*, vol. 22, no. 6, Article ID 060503, 2017.
- [4] H. Li, J. Weng, Y. Shi et al., "An improved deep learning approach for detection of thyroid papillary cancer in ultrasound images," *Scientific Reports*, vol. 8, no. 1, p. 6600, 2018.
- [5] S. Tsantis, D. Cavouras, I. Kalatzis, N. Piliouras, N. Dimitropoulos, and G. Nikiforidis, "Development of a support vector machine-based image analysis system for assessing the thyroid nodule malignancy risk on ultrasound," *Ultrasound in Medicine and Biology*, vol. 31, no. 11, pp. 1451–1459, 2005.
- [6] J. Ma, S. Luo, M. Dighe, D.-J. Lim, and Y. Kim, "Differential diagnosis of thyroid nodules with ultrasound elastography based on support vector machines," in *Proceedings of the 2010 IEEE International Ultrasonics Symposium*, pp. 1372–1375, San Diego, CA, USA, October 2010.
- [7] U. R. Acharya, S. V. Sree, G. Swapna et al., "Effect of complex wavelet transform filter on thyroid tumor classification in three-dimensional ultrasound," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 227, no. 3, pp. 284–292, 2013.
- [8] A. Shukla, R. Tiwari, P. Kaur, and R. R. Janghel, "Diagnosis of thyroid disorders using artificial neural networks," in *Proceedings of the IEEE International Advance Computing Conference*, Patiala, India, March 2009.
- [9] A. Prochazka, S. Gulati, S. Holinka, and D. au, "Patch-based classification of thyroid nodules in ultrasound images using direction independent features extracted by two-threshold binary decomposition," *Computerized Medical Imaging and Graphics*, vol. 71, pp. 9–18, 2019.
- [10] R. Thandiackal, C. Brandle, D. Leach, A. Jafari, and F. Iida, "Exploiting passive dynamics for robot throwing task," in

- Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vilamoura-Algarve, Portugal, October 2012.
- [11] N. Gessert, M. Bengs, L. Wittig et al., "Deep transfer learning methods for colon cancer classification in confocal laser microscopy images," *International Journal of Computer Assisted Radiology and Surgery*, vol. 14, no. 11, pp. 1837–1845, 2019.
 - [12] T. Liu, S. Xie, J. Yu, L. Niu, and W. Sun, "Classification of thyroid nodules in ultrasound images using deep model based transfer learning and hybrid features," in *Proceedings of the IEEE International Conference on Acoustics*, New Orleans, LA, USA, March 2017.
 - [13] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
 - [14] X. Zhang, H. Gu, L. Zhou, L. Chen, Q. Zheng, and W. Chen, "Improved dual-domain filtering and threshold function denoising method for ultrasound images based on non-subsampled contourlet transform," *Journal of Medical Imaging and Health Informatics*, vol. 7, no. 7, pp. 1624–1628, 2017.
 - [15] W.-B. Chen, L. Zhou, H.-B. Gu, and X. Zhang, "Ratio weighted and condition weighted medical image fusion based on contourlet transformation coefficients," *International Journal of Knowledge-Based and Intelligent Engineering Systems*, vol. 19, no. 3, pp. 191–198, 2015.
 - [16] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems*, Red Hook, NY, USA, December 2012.
 - [17] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. De Vries, M. J. N. L. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1252–1261, 2016.
 - [18] W.-B. Chen, M. Hu, L. Zhou, H. Gu, and X. Zhang, "Fusion algorithm of multi-focus images with weighted ratios and weighted gradient based on wavelet transform," *Journal of Intelligent Systems*, vol. 8, no. 19, pp. 0077–0078, 2017.
 - [19] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowledge-Based Systems*, vol. 108, no. 15, pp. 42–49, 2016.
 - [20] X. Zhang, M. Fan, D. Wang, P. Zhou, and D. Tao, "Top- k feature selection framework using robust 0-1 integer programming," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 7, pp. 3005–3019, 2021.
 - [21] X. Zhang, D. Wang, Z. Zhou, and Y. Ma, "Robust low-rank tensor recovery with rectification and alignment," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 238–255, 2021.
 - [22] T. Williams and R. Li, "Advanced image classification using wavelets and convolutional neural networks," in *Proceedings of the 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Anaheim, CA, USA, February 2017.
 - [23] Y. Tokuoka, S. Suzuki, and Y. Sugawara, "An inductive transfer learning approach using cycle-consistent adversarial domain adaptation with application to brain tumor segmentation," 2020, <https://arxiv.org/abs/2005.04906>.
 - [24] V. Jayaram, M. Alamgir, Y. Altun, B. Scholkopf, M. Grosse-Wentrup, and Moritz, "Transfer learning in brain-computer interfaces abstractuFFFDThe performance of brain-computer interfaces (BCIs) improves with the amount of avail," *IEEE Computational Intelligence Magazine*, vol. 11, no. 1, pp. 20–31, 2016.
 - [25] X. Zhang, T. Wang, J. Wang, G. Tang, and L. Zhao, "Pyramid channel-based feature attention network for image dehazing," *Computer Vision and Image Understanding*, vol. 197–198, Article ID 103003, 2020.
 - [26] X. Zhang, J. Zhang, W. Chen, M. Hu, and X. Zhang, "Double fidelity terms unidirectional variation model for single-image rain removal," *Jouanal of Electronic Imaging*, vol. 28, no. 2, pp. 15–23, 2019.
 - [27] M. Mehdipour Ghazi, B. Yanikoglu, and E. Aptoula, "Plant identification using deep neural networks via optimization of transfer learning parameters," *Neurocomputing*, vol. 235, no. 26, pp. 228–235, 2017.
 - [28] B. Zoph, D. Yuret, J. May, and K. K. Jonathan, "Transfer learning for low-resource neural machine translation," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, TX, USA, November 2016.
 - [29] P. Kulkarni, J. Zepeda, F. Jurie, P. Perez, and L. Chevallier, "Hybrid multi-layer deep CNN/aggregator feature for image classification," 2015, <https://arxiv.org/abs/1503.04065>.
 - [30] D. Mahapatra, P. K. Roy, S. Sedai, and R. Garnavi, "Retinal image quality classification using saliency maps and CNNs," in *Proceedings of the International Workshop on Machine Learning in Medical Imaging*, pp. 172–179, Athens, Greece, October 2016.
 - [31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, <https://arxiv.org/abs/1409.1556>.
 - [32] W. K. Moon, Y.-W. Lee, H.-H. Ke, S. H. Lee, C.-S. Huang, and R.-F. Chang, "Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks," *Computer Methods and Programs in Biomedicine*, vol. 190, Article ID 105361, 2020.
 - [33] S. Y. Kim, Y. Choi, E. K. Kim et al., "Deep learning-based computer-aided diagnosis in screening breast ultrasound to reduce false-positive diagnoses," *Scientific Reports*, vol. 11, no. 1, pp. 1–11, 2021.
 - [34] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: a brief review," *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 7068349, 13 pages, 2018.
 - [35] T. O Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications & Networking*, vol. 3, no. 4, pp. 563–575, 2017.
 - [36] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and R. S. Tan, "A deep convolutional neural network model to classify heartbeats," *Computers in Biology and Medicine*, vol. 89, 2017.
 - [37] J. Shen, S. H. Kong, and T. F. Chan, "Euler's elastica and curvature-based inpainting," *SIAM Journal on Applied Mathematics*, vol. 63, no. 2, pp. 564–592, 2003.
 - [38] C. Szegedy, W. Wei Liu, Y. Yangqing Jia et al., "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, June 2015.