Hindawi
Scientific Programming
Volume 2023, Article ID 2426601, 11 pages
https://doi.org/10.1155/2023/2426601

Research Article

Local Binary Convolutional Neural Networks' Long Short-Term Memory Model for Human Embryos' Anomaly Detection

Sajad Einy,[1] Esra Sen,[2] Hasan Saygin,[1] Hemrah Hivehchi,[1] and Yahya Dorostkar Navaei [3]

[1]*Istanbul Aydin University, Department of Application and Research Center for Advanced Studies, Istanbul, Turkey*
[2]*Istanbul Aydin University, Faculty of Medicine, Department of Histology and Embryology, Istanbul, Turkey*
[3]*Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran*

Correspondence should be addressed to Yahya Dorostkar Navaei; y.dorostkar@qiau.ac.ir

Received 14 February 2022; Revised 20 October 2022; Accepted 18 March 2023; Published 17 April 2023

Academic Editor: Sadiq Hussain

Accurate selection of embryos with the maximum implementation condition is a necessary step to increase the effectiveness of fertility treatment in *in vitro fertilization* (IVF). The deep learning algorithms presented high potential for monitoring and visualizing embryo features such as cell numbers and their morphological development in time series manner. Due to the ability of the computer vision and deep learning algorithms, this paper aimed to present a novel deep learning approach to distinguish simultaneous abnormality of embryos in time-lapse systems for detecting live and non-live births in IVF. The approach is composed of local binary convolutional neural network (LBCNN) and long short-term memory (LSTM). The LBCNN improved accuracy of classification by employing deep and local feature sets with lower number of learnable parameters in comparison with a standard convolutional layer. Moreover, LSTM network is employed to analyze temporal information of time-lapse embryos. The results indicate that the proposed approach achieves significant results in ROC analysis (0.98) in 5 days of monitoring compared to state-of-the-art models. In addition, the approach showed compatible results in early diagnosis of abnormality detection (72 hours) with 82.8% accuracy of classification compared to the pretrained well-known convolutional neural network (CNN) models and baseline CNN.

1. Introduction

Computer-aided detection (CAD) is designed to reduce experimental mistakes of physicians interpreting healthcare data. Recently, CAD systems with the help of machine learning approaches are utilized for different types of disease treatment and detection [1]. These kinds of approaches could be categorized into two types of models such as deep [2] and shallow-based models [3]. The shallow-based models consist of optimization algorithms (genetic algorithm, particle swarm optimization, etc.) [4] and shallow machine learning algorithms (support vector machine, linear regression, etc.) [5]. For instance, improved machine learning-based system based on improved adaptive particle swarm optimization algorithm and artificial immune recognition system were designed for wart disease treatment [6]. The main advantages of these approaches are remarkable results with less training data with fewer processing cost. Commonly, the deep learning approach-based frameworks for CAD systems are utilized as pretrained well-known models in case of deep feature learning and extractions [7]. The main reason for utilizing transfer learning is reduction of processing cost and enhanced efficiency during training new models. For example, the authors in [8] proposed deep and local convolutional neural network for brain anomaly detection. Furthermore, CAD systems with the help of deep learning algorithms have grown significantly because of high accuracy rate [9]. In this study, anomaly detection based on deep learning algorithms is utilized for CAD systems. Anomaly detection (also referred to as outlier detection)

commonly implies the detection of rare issues, events, or observations that deviate considerably from most data and do not adapt to a well-defined concept of normal performance. In the context of CAD systems, anomaly detection is such that the process can be used to alert physicians of abnormal physiological data that could indicate health complications [10, 11].

Due to the effects of CAD systems on healthcare, we presented anomaly detection with the help of deep learning approach for in vitro fertilization (IVF) [12]. Nowadays, the world's population suffers from many infertility problems that prevent normal reproduction. One of the most widely used technologies for infertility treatment is IVF by the collecting multiple follicles for the fertilization and in vitro culture [13, 14]. This technique is a useful and non-invasive method for reproduction since it allows to evaluate the fetus without harm. Moreover, the time-lapse videos and photos of fetuses are employed for fetal growth at different time lapse using brief time intervals, and they can be used as an advanced technology to record the growth of the embryos in IVF. They are also widely applied in various cases, for example, the supervision of reproduction by medical centers during the embryonic development [15]. The embryo quality (cleavage embryo scoring) is categorized into four main stages. Stage one embryos have four blastomeres on day two and eight blastomeres on day three. Their blastomeres are of equal size, large, round, and with clear cytoplasm and no fragments. Although the amount and shape of blastomeres are like stage one embryo, stage two embryos might have 10% fragments and have irregular blastomeres. In stage three embryos, the amount of blastomeres is fewer than normal, and the fragmentation ratio is 20% or above. The amount of uneven blastomeres improves in stage one and two embryos. In stage four embryos, the structure and volume of the blastomeres are distinct from each other, and the disintegration ratio is above 50%, as presented in Table 1 [16].

Gardner's classification criteria were used to score the embryos. In Gardner's classification, each embryo is scored based on blastocoel size, internal cell mass, and trophectoderm structure as presented in Table 2 [17].

For these reasons, the main objective of this study is to present anomaly detection with novel deep learning approach based on time-lapse device video data. In addition, the accuracy of implantation potential has been detected by morphological analysis of human embryos in the early stages of development (stage 1). In addition, the anomaly detection in this scenario means that blastocyst structure grade is not in *A*, *B*, and *C* grades based on Gardner's classification [18].

In the past, the common strategy for selecting quality embryos was mainly by examining the number of cells, the degree of fragmentation, and the number of nuclei in the incision stage, while poor quality embryos (based on morphology) were discarded [19]. Today, with the improvement of laboratory culture conditions and further development of physiological culture media, long cultures up to the blastocyst stage are performed. In humans, blastocyst formation begins about 5 days after fertilization, when a fluid-filled cavity occurs in the morula, which is the early embryonic stage of a 16-cell embryo. It is essential to select high-quality embryos for IVF [20]. Scoring systems for morphological evaluation of embryos have been developed to increase the birth rate. However, these methods are not sufficient to predict live birth, as there is no obvious link between morphology and chromosomal aneuploidy [21].

Recently, various applications with deep learning and computer vision methods have been proposed to enable the automation of embryo assessments for IVF analysis. For instance, in [22], a deep learning approach was presented based on extracting hierarchical features from input data instead of rule-based image processing programming. In another similar study [23], Google's Inception model was applied for time-lapse images of blastocyst selection in *in vitro* fertilization. The study of [14] proposed the multitask deep learning with dynamic programming approach in the classification of the development of embryos based on time-lapse images. Since stereoscopic cells have potential to overlap at different sizes in the time-lapse imaging, even for an experienced embryologist, it is difficult to count the number of cells in a single image. Therefore, some studies focused on early periods of embryo development. For instance, a set of zygotes was analyzed and significant results in blastocyst stage classification were obtained on the second day after fertilization [24]. In case of recognition of ploidy status, two-stream inflated 3D ConvNet [25] proposed for classifying time-lapse videos. Another important field of study in deep learning-based embryo monitoring systems was the automatic grading of blastocysts. In the study of [26], a convolutional neural network was proposed to detect inner cell mass and trophectoderm grades from each image. The recurrent neural network was utilized on the top of the network to classify blastocyst temporal information from video. Another time series-based embryo classification study [27] suggested the two-classifier vote-based method using a convolutional neural network (CNN). In this study, the number of cells of embryo was detected and classified. Another similar study [19] presented non-invasive classification of embryos with the help of attention branch network (ABN). Another study examined the automatic framework, namely, Cell-Net [12]. This CNN-based approach includes residual incremental atrous pyramid for counting and centroid localization of embryonic cells. Another study, a computer-automated based time-lapse with image analysis approach presented in [28]. In the case of a smaller number of training data and improving the accuracy of classification, [29] presented a novel approach, namely HEMIGEN. This approach employed a Generative Adversarial Network (GAN) for the production of one-, two-, and four-cell time images and classification with a deep neural network (DNN).

According to these studies, it is clear that most of the research in the field of embryo monitoring systems has been faced with two main challenges: detection abnormalities and qualitative classification. The first challenge is to extract robust and advanced features for identifying one-, two-, and four-cell embryonic stages for training deep learning models. Another challenge is the time series training model with recurrent neural networks to evaluate the cellular phase divided at specific times of embryo development. Therefore,

Table 1: Blastocyst stage and descriptions [16].

Blastocyst stage	Stage	Description
Early blastocyst	1	Blastocoel that is less than half of the volume of the embryo
Blastocyst	2	Blastocoel that is half of or greater than half of the volume of the embryo
Full blastocyst	3	Blastocoel completely filling the embryo
Expanded blastocyst	4	Blastocoel volume larger than that of the early embryo, with a thinning zona pellucida
Hatching blastocyst	5	Trophectoderm starting to herniate though the zona pellucida
Hatched blastocyst	6	Blastocyst has completely escaped from the zona pellucida

Table 2: Blastocyst grading system and descriptions [17].

Blastocyst structure	Grade	Description
Inner cell mass	*A*	Tightly packed, many cells
	B	Loosely grouped, several cells
	C	Very few cells
Trophectoderm	*A*	Many cells forming a cohesive epithelium
	B	Few cells forming a loose epithelium
	C	Very few large cells

we aimed to conduct a novel deep learning approach for time series abnormality embryo detection. This approach included local and deep extracted features from each input frames of time-lapse embryoscope device. Due to fusion of deep and local features in presented deep learning recurrent neural network approach, the accuracy of classification improved along with detection of abnormality in early stages. To sum up, the main contributions of this paper are as follows:

(1) This study proposed a novel deep learning framework consisting of local and deep features in time series manner for abnormality detection of embryo time lapse, namely, LBCN-LSTM.

(2) The proposed LBCN-LSTM demonstrated significant results in the accuracy rate and receiver operating characteristic (ROC) curve analysis compared to experimented baseline CNN-LSTM and pretrained models.

(3) This proposed approach achieved compatible results in early diagnosis of anomalies in embryo monitoring. Furthermore, the present technique is unique and has much more advantages compared to the traditional methods in terms of training cost.

(4) We presented a public access embryo monitoring database.

The rest of this paper is categorized as follows. Section 2 presents methodology. Section 3 presents experimental results, and finally the last section gives the conclusion.

2. Methodology

Figure 1 illustrates the summary of the approach which is the automated system for detecting abnormalities in embryo monitoring. The time-lapse image sequence in time series manner is fed into the proposed deep learning framework. For each input image, deep and local features are considered as extracted contribution of the LBCNN [30]. Furthermore, last layer of this CNN approach contains 2D global average pooling layer connected to the fully connected layer as input of LSTM network to leverage temporal information. Finally, the last layer of the LSTM network contains SoftMax function with one node hidden layer to detect anomaly in embryo time-lapse monitoring with $Y_1, Y_2, \ldots, Y_n$ outputs. In the following paragraphs, the input images, the LBCNN, and the LSTM are described in more detail.

2.1. LBCNN. In this study, we applied the local binary convolutional neural network (LBCNN) in the deep learning approach to decrease the computational complexity of CNNs with better classification accuracy. The LBCNN is inspired by open-source codes available in the following link: https://github.com/whoisraibolt/LBCNN. The LBC layer consists of fixed sparse binary filters, an activation function, and a set of trainable weights which reduces with optimizing algorithm. Local binary convolutional neural network (LBCNN) consists of LBC layers. Each local binary pattern (LBP) [31] extracted 8 resulting sums of all the bit maps which achieves the same results of eight 3×3 convolutional filters followed by simple binarization. The predefined weight vector v includes $[2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^8]$. In this case, LBP can be reformulated as follows:

$$y = \sum_{i=1}^{8} \sigma(b_i \times X).v_i. \quad (1)$$

In this formula, $x \in R^d$ is defined as vectorization of the input images, b_i represents sparse convolutional filters, σ is the non-linear binarization (Heaviside step function) operator, and y is the result of LBP image. In this case, LBCNN includes m fixed convolutional filters and input image is filtered by LBC, which produces m variance feature maps (bit maps). The Heaviside step function is employed for backpropagation in the LBC layer with sigmoid or ReLU

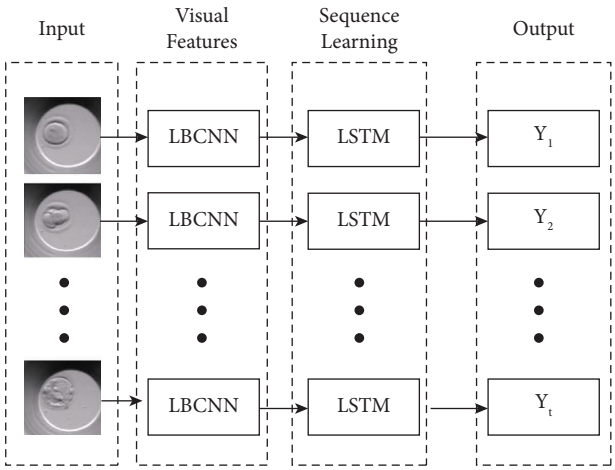

Figure 1: LBCNN-LSTM model for time-lapse embryo anomaly detection.

differentiable activation function. Each LBC layer feeds feature map as input (x_l) to the next layer, and generalized multichannel input is presented as follows:

$$x_{l+1}^{t} = \sum_{i=1}^{m} \sigma\left(\sum_{s} b_i^s \times x_l^s \right).v_{l,i}^{t}, \tag{2}$$

where t and s are output and input channels, respectively. The last step in calculating the total weight of the activations can be implemented through a convolution operation with (1×1) size filters. Consequently, each LBC layer contains two blocks of convolutional layers with fixed and non-learnable weights continually. The architecture of LBCN layer is presented in Figure 2.

As presented in Figure 3, the proposed model includes residual LBCN blocks with different hidden layer sizes. The residual LBCN block consists of two LBCNs with batch normalization and three Conv2D layers. Furthermore, the first LBCN layer is fused to last Conv2D layer for each block.

Moreover, we utilized different hidden unit sizes for the LBCN blocks. The efficient and high accuracy model is selected with four layers of LBCN with hidden unit size [512, 256, 128, 64] architecture, named as LBCNN-4L, according to the experimental results. The architecture of the LBCNN-4L is presented in Table 3.

2.2. LSTM Networks. The recurrent neural network (RNN) [32] model has been widely employed for sequential data analysis in machine learning. Nevertheless, RNN has limitations in terms of learning long-term dependencies due to the vanishing of gradients during several backpropagation processes. The LSTM [33] network has been developed to reduce the weaknesses of RNN models for long periods of time. The LSTM is an RNN-enhanced version that can process long-term consecutive data with a low gradient vanishing rate compared to other algorithms. The LSTM algorithm with long-term memory can predict multivariate time series data with high accuracy. The LSTM block structure can model time series predictions such as long-term dependencies. Therefore, this paper employed LSTM networks in case of predicting the data in time series.

3. Experimental Results

Table 4 indicates the model configuration for the proposed approach. Moreover, experimental analysis was conducted based on default input image size of well-known pretrained networks such as VGG16 [34], Resnet-50 [35], Inception V3 [36], MobileNet V2 [37], and Xception [38] with training configuration. In addition, for fair comparison of these models, we set up optimization algorithm (Adam), activation function (ReLU), momentum (0.9), weight decay $(1e-6)$, mini-batch size (16), and epochs (1000). Furthermore, we utilized Adam optimization algorithm due to first-order gradient of stochastic objective functions because of adaptive values of lower-order moments. In this case, Adam has advantage for problems that are in conditions of large parameters [39]. Furthermore, the input window size of LBCNN and baseline CNN is set up (1024, 1024). In case of learning rate due to fine-tuning of the pretrained models, the rate is lower than that of LBCNN and baseline CNN, respectively, with 0.0001 and 0.01.

3.1. Embryo Database. The analysis data were obtained from Vitrolife Embryoscope device in Istanbul Aydin University. This database contains eight non-healthy embryos and twelve healthy embryos with 102 hours of monitoring. One example of healthy and non-healthy embryo is presented in Figure 4. In this figure, healthy and non-healthy videos in specific times such as 2, 20, 50, and 90 hours are presented. These embryo data were labeled by Dr. Esra Sen. The experimented datasets used to support the findings of this study are available from the corresponding author upon request.

3.2. Architecture Analysis Details. To test the effect of the input image window size on the classification accuracy for the LBCNN model, we utilized different window sizes including (128×128), (256×256), (512×512), and (1024×1024) and compared the model results. As presented in Figure 5, the test images of each video of the blastocyst are extracted and the model is trained with SoftMax classification layer with a single node for the anomaly detection.

The blastocyst image dataset is split into three sets: train, validation, and test sets, with 70%, 20%, and 10% ratio. The results showed that the input size of image with 1024×1024 in the LBCNN achieved a highest accuracy among the other input image sizes for baseline CNN such as (512×512) and (256×256). In addition, the result of Table 5 shows that the LBP features with convolutional layers increased the accuracy of classification compared to baseline convolutional layers. The architecture of LBCNN and baseline CNN is based on four layers with 3×3 kernel size convolutional layer with nodes 512, 256, 128, and 64.

After the testing the input size of the image, we analyzed the effects of the deep network on the accuracy of the abnormality detection. In this experiment, we designed three different models with four different node sizes for each LBCN block. The three experimented LBCNN models are named as LBCNN-4L, LBCNN-5L, and LBCNN-6L with

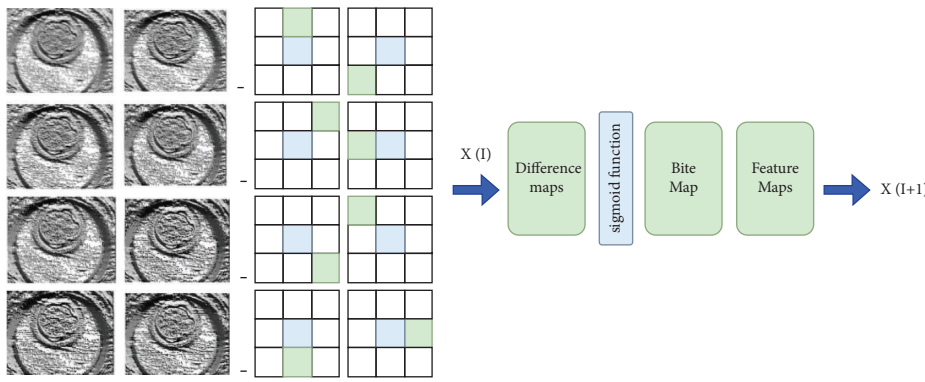

Figure 2: LBCN layer architecture.

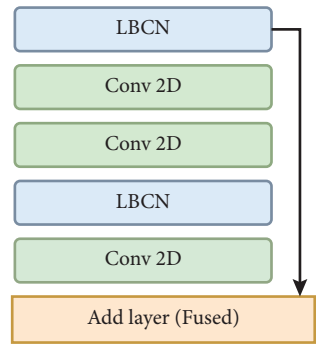

Figure 3: The architecture of LBCN block.

Table 3: Details of the architecture of LBCNN-4L network.

Layer (type)	Output shape	Param #	Connected to
Input (input layer)	(1024, 1024, 3)	0	
Lbc_1 (LBC)	(1024, 1024, 3)	13824	Input
conv2d_2 (Conv2D)	(1024, 1024, 3)	8192	lbc_1
lbc_2 (LBC)	(512, 512, 512)	294912	conv2d_2
conv2d_3 (Conv2D)	(512, 512, 512)	262144	lbc_2
conv2d_4 (Conv2D)	(512, 512, 512)	262656	conv2d_3
lbc_3 (LBC)	(512, 512, 512)	4718592	conv2d_4
conv2d_5 (Conv2D)	(512, 512, 512)	262144	lbc_3
add_1 (Add)	(512, 512, 512)	0	conv2d_3 and conv2d_5
lbc_4 (LBC)	(256, 256, 512)	4718592	add_1
conv2d_6 (Conv2D)	(256, 256, 256)	131072	lbc_4
conv2d_7 (Conv2D)	(256, 256, 256)	65792	conv2d_6
lbc_5 (LBC)	(256, 256, 512)	2359296	conv2d_7
conv2d_8 (Conv2D)	(256, 256, 256)	131072	lbc_5
add_2 (Add)	(256, 256, 256)	0	conv2d_6 and conv2d_8
lbc_6 (LBC)	(128, 128, 512)	2359296	add_2
conv2d_9 (Conv2D)	(128, 128, 128)	65536	lbc_6
conv2d_10 (Conv2D)	(128, 128, 128)	16512	conv2d_9
lbc_7 (LBC)	(128, 128, 512)	1179648	conv2d_10
conv2d_11 (Conv2D)	(128, 128, 128)	65536	lbc_7
add_3 (Add)	(128, 128, 128)	0	conv2d_9 and conv2d_11
lbc_8 (LBC)	(64, 64, 512)	1179648	add_3
conv2d_12 (Conv2D)	(64, 64, 64)	32768	lbc_8
conv2d_13 (Conv2D)	(64, 64, 64)	4160	conv2d_12
lbc_9 (LBC)	(64, 64, 512)	589824	conv2d_13
conv2d_14 (Conv2D)	(64, 64, 64)	32768	lbc_9
add_4 (Add)	(64, 64, 64)	0	conv2d_12 and conv2d_14
Average pool (GlobalAveragePooling2)	(64)	0	add_4

Table 4: Values of the parameters in the proposed approach used in this study.

CNN model	Image size	Optimization	Activation function	Momentum	Decay	Mini-batch	Learning rate	Epoch
Pretrained models	Default size	Adam	ReLU	0.9	$1e-6$	16	0.0001	1000
LBCNN	(1024, 1024)	Adam	ReLU	0.9	$1e-6$	16	0.01	1000
Baseline CNN	(1024, 1024)	Adam	ReLU	0.9	$1e-6$	16	0.01	1000

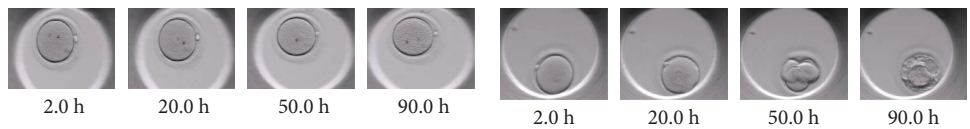

Figure 4: Monitoring the embryo time lapse with Vitrolife system. (a) Non-healthy embryo. (b) Healthy embryo.

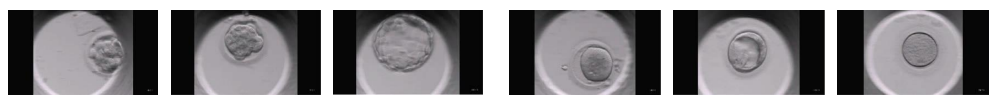

Figure 5: Blastocyst images extracted from time-lapse device. (a) Healthy embryo. (b) Non-healthy embryo.

Table 5: Different windows sizes for input image for LBCNN.

Input image	(128×128)	(256×256)	(512×512)	(1024×1024)
LBCNN	0.78	0.82	0.96	0.98
Baseline CNN	0.62	0.79	0.93	0.95

Table 6: Different architectures for LBCNN model.

Model	LBCNN-4L	LBCNN-5L	LBCNN-6L
Accuracy of classification	0.98	0.98	0.96

Conv layer (3×3 window size). The input block for all experimental models is (1024×1024) with the input block containing two layers of LBC with batch normalization and Conv 2D continually. Besides, we utilized hidden unit sizes [512, 256, 128, 64], [512, 256, 128, 64, 32], and [512, 256, 128, 64, 32, 16] continually for LBCN blocks of LBCNN-4L, LBCNN-5L, and LBCNN-6L models, respectively.

Contrary to the highest accuracy in the LBCNN-4L and LBCNN-5L, the lowest accuracy was obtained in the LBCNN-6L. The main reason for the inefficiency of the LBCNN-6L model is the lack of training data. According to the experimental results of the studied articles [1–10], it is clear that a large number of training data have great impact on the performance of large structure of convolutional neural networks. Besides, the results showed that similar results were obtained from the LBCNN-4L and LBCNN-5L. To reduce the processing cost, we selected the LBCNN-4L. Therefore, throughout the work, the architecture of the LBCNN is selected as 4 layers with (1024×1024). The architecture of the LBCNN-4L is presented in Figure 4, and the details are presented in Table 6.

3.3. Embryo Anomaly Detection. After experimental results by the blastocyst anomaly detection, we examined time series-based embryo anomaly detection with the LSTM neural network. To study the effects of LBP features in proposed approach, we compared LBCNN + LSTM with baseline CNN + LSTM considering the training cost and accuracy. For the appropriate comparison between two models, we applied the same input image windows size, layer, and size of hidden nodes for the baseline CNN. Normal convolutional layer with same size of nodes is employed rather than the baseline model of the LBCN. The architecture of baseline CNN is presented in Table 7. Furthermore, we applied the same configuration of LSTM for both systems for the appropriate comparison. This configuration is obtained with the help of the grid search methodology by the highest accuracy. In this paper, the LSTM sequential model is designed to analyze the time series anomaly detection of the developing time-lapse embryo monitoring. A linear stack of layers is utilized in two layers with return sequences. The first layer of the LSTM layer contains 600 memory units and it returns sequences. The second layer includes 600 memory units. After each LSTM layer, the dropout layer is applied. Finally, the last layer has a fully connected layer with a SoftMax activation function with one node for detecting anomaly.

ROC curve analysis is used to show the connection between two possible approaches between sensitivity and specificity. The ROC curve results showed that LBCNN has better area under the ROC curve (AUC) than baseline CNN.

Table 7: Baseline CNN architecture details.

Layer (type)	Output shape	Connected to
Input (input layer)	(1024, 1024, 3)	
conv2d_1 (Conv2D)	(1024, 1024, 3)	Input
conv2d_2 (Conv2D)	(1024, 1024, 3)	conv2d_1
conv2d_3 (Conv2D)	(512, 512, 512)	conv2d_2
conv2d_4 (Conv2D)	(512, 512, 512)	conv2d_3
conv2d_5 (Conv2D)	(512, 512, 512)	conv2d_4
conv2d_6 (Conv2D)	(512, 512, 512)	conv2d_5
conv2d_7 (Conv2D)	(512, 512, 512)	conv2d_6
add_1 (Add)	(512, 512, 512)	conv2d_4 and conv2d_7
conv2d_8 (Conv2D)	(256, 256, 512)	add_1
conv2d_9 (Conv2D)	(256, 256, 256)	conv2d_8
conv2d_10 (Conv2D)	(256, 256, 256)	conv2d_9
conv2d_11 (Conv2D)	(256, 256, 512)	conv2d_10
conv2d_12 (Conv2D)	(256, 256, 256)	conv2d_11
add_2 (Add)	(256, 256, 256)	conv2d_8 and conv2d_12
conv2d_13 (Conv2D)	(128, 128, 512)	add_2
conv2d_14 (Conv2D)	(128, 128, 128)	conv2d_13
conv2d_15 (Conv2D)	(128, 128, 128)	conv2d_14
conv2d_16 (Conv2D)	(128, 128, 512)	conv2d_15
conv2d_17 (Conv2D)	(128, 128, 128)	conv2d_16
add_3 (Add)	(128, 128, 128)	conv2d_13 and conv2d_17
conv2d_18 (Conv2D)	(64, 64, 512)	add_3
conv2d_19 (Conv2D)	(64, 64, 64)	conv2d_18
conv2d_20 (Conv2D)	(64, 64, 64)	conv2d_19
conv2d_21 (Conv2D)	(64, 64, 512)	conv2d_20
conv2d_22 (Conv2D)	(64, 64, 64)	conv2d_21
add_4 (Add)	(64, 64, 64)	conv2d_18 and conv2d_22
Average pool (GlobalAveragePooling2)	(64)	add_4

As shown in Figure 6, the LBCNN and CNN succeeded at 0.985 and 0.989, respectively.

To examine the training cost, we conducted the test between LBCNN and CNN in Figure 7. Both systems are analyzed in specific number of epochs and compared with accuracy of classification measurement. LBCNN and baseline CNN are trained with different values of learning epochs, and their accuracies of training and testing for each size of epochs are presented in Table 3. In this case, we selected different number of epochs such as 10, 25, 50, 75, and 100 for training of LBCNN and baseline CNN. The lowest accuracy rate in baseline CNN is achieved in 10 epochs with 32%; in contrast, the LBCNN model attained 37% of accuracy rates. Furthermore, both experimented models achieved same accuracy rates with 55%. In addition, in 100 epochs, LBCNN compared with CNN model improved accuracy rates with 8%. In conclusion, these results show that the proposed method ensured better accuracy in a smaller number of training epochs because of few number of trainable parameters compared with the baseline CNN.

Other reason for this advantage can be explained by extracting robust and enhanced features from entire video for classification. For better visualization of extracted features, we employed t-distributed stochastic neighbor embedding (t-SNE). Figure 8 shows that LBCNN extracted more enhanced features and the anomaly embryo t-SNE features were clearly separated from normal ones compared to baseline CNN. In this figure, the healthy and non-healthy features are presented in blue and red points (Figure 8).

3.4. Comparison of Time Series-Based Extracted Deep and Local Features. In the second step, we examine different well-known pretrained models as a deep feature extractor with the LSTM for the video classification in Table 8. In addition, these results are compared with the LBCNN as input for time series classification model (LSTM). The embryo video dataset is split into three sets: train, validation, and test sets, with the ratio of 70%, 20%, and 10%. The results indicated that the best experimental result is obtained by Resnet-50 + LSTM with the values 0.97, 0.98, 0.93, and 0.98. In addition, the second highest accuracy is obtained by LBCNN + LSTM with the values 0.96, 0.95, 0.96, and 0.96, respectively, for accuracy, precision, recall, and $F1$-score, respectively. Furthermore, the lowest accuracy was achieved by the LeNet 5 + LSTM with 0.76. Based on these experimental results, it can be concluded that the LBCNN can extract robust features and achieve significant accuracy with few number of learnable parameters compared to VGG16, Inception V3, and Xception models.

3.5. Early Diagnosis of Embryo Anomaly. To test the early diagnosis of embryo anomaly, we conducted a test by the different time sections, and we selected five different time-lapse videos, namely, 12 h, 24 h, 48 h, 60 h, and 72 h, as presented in Figure 9 by the standard morphology of embryo. This figure separately presented each class of the data (train and test). For instance, the class of 12 h contains the video frames from 0.0 h up to 12.0 h. Similarly, the 24 h class contains the video frames from 0.0 h up to 24.00 h.

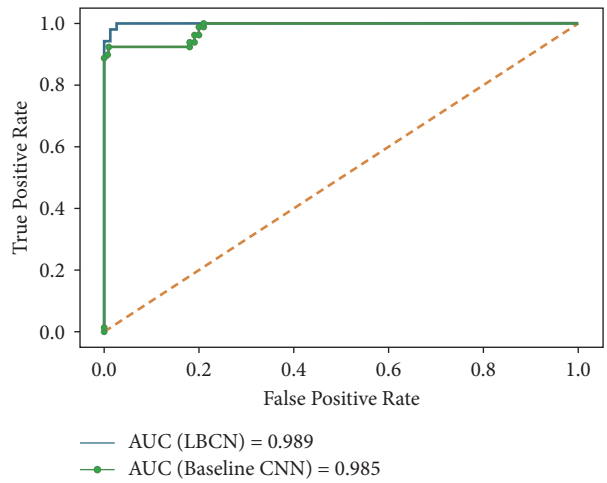

Figure 6: ROC curve analysis of embryo time-lapse anomaly detection.

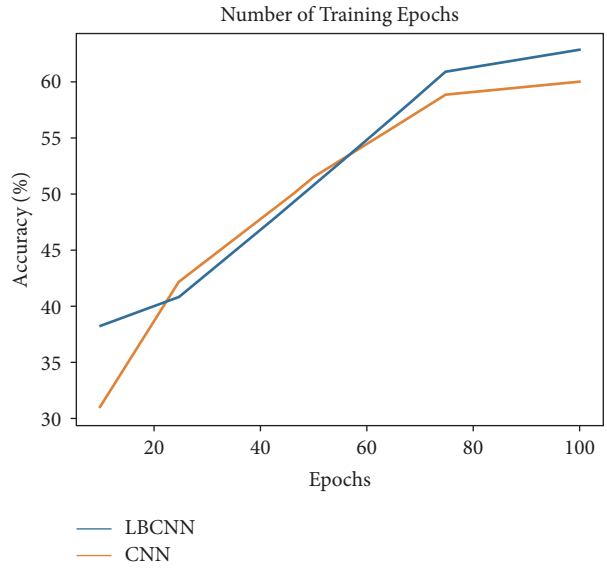

Figure 7: Number of epochs for training based on accuracy of classification.

As shown in Table 9, the LBCNN + LSTM can achieve the highest accuracy in the 72 h class compared to well-known pretrained models. In this class, Resnet-50, VGG16, and Inception V3 models with LSTM achieved 80.3, 78.9, and 79.7% accuracy of classification, respectively. The finding also indicated that the proposed method achieved successful results compared to the well-known pretrained models like VGG16, LeNet 5, Resnet-50, Inception V3, MobileNet V2, and Xception. The main reasons of improvement of embryo early detection in the proposed system can be explained by the employing deep and local combined features for the classification.

3.6. Discussion. We also evaluated the different methods and compared them with this study by the quality analysis and the abnormality detection of embryo in video or image analysis (Table 10). It can be concluded that the quality analysis with single blastocyst image by Khosravi et al. [23] had better accuracy compared to CNN-LSTM with video analysis technique by Kragh et al. [26]. Moreover, the LBCNN-LSTM approach achieved better results in abnormality detection with AUC = 0.98 compared to study of Tran et al. [22], Lee et al. [25], and Sawada et al. [19] with the results of AUC curve analysis at 0.93, 0.74, and 0.93, respectively. In addition, among the proposed approaches by the abnormality detection, the results of the LSTM with attention map [19] have still significant result. In this case, we can conclude that the blastocyst morphology analysis has more effects than time series analysis on the abnormality detection. Furthermore, Payá et al. [40] proposed a supervised contrastive learning framework for grading and anomaly detection of embryos which achieved 0.94 AUC in abnormality detection. Nevertheless, this comparison is

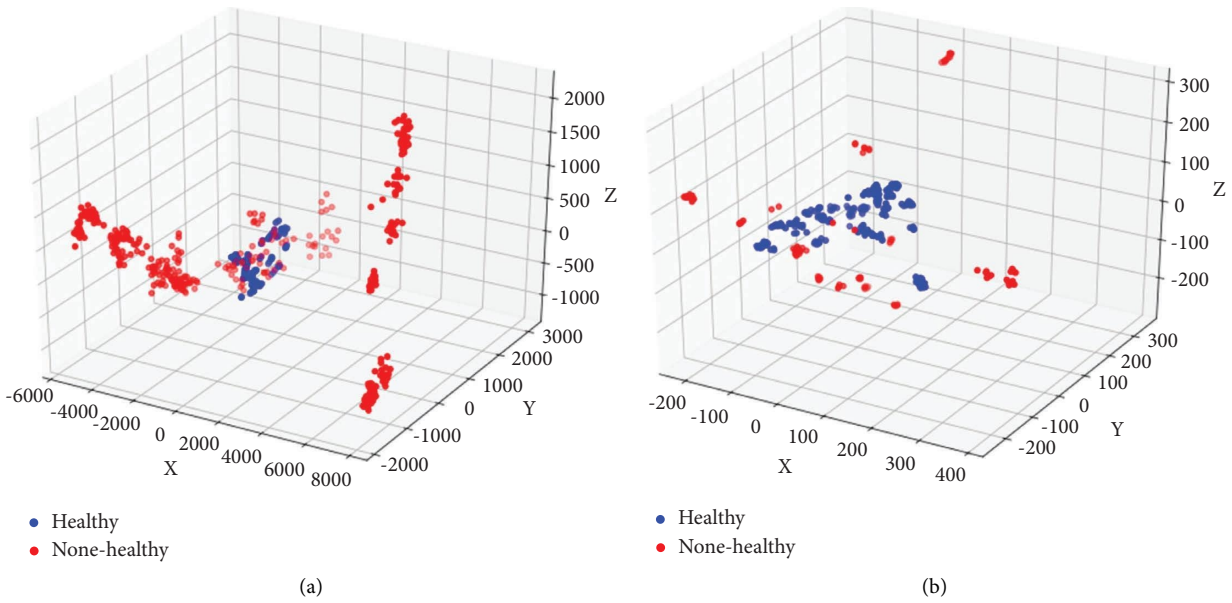

Figure 8: Features extracted from last layer of proposed deep models. (a) LBCNN. (b) Baseline CNN.

Table 8: Embryo video anomaly detection results based on different pretrained model's architectures.

Model	Accuracy	Precision	Recall	*F*1-score
VGG16 + LSTM	0.91	0.89	0.84	0.86
LeNet 5 + LSTM	0.76	0.75	0.75	0.76
Resnet-50 + LSTM	0.97	0.98	0.93	0.98
Inception V3 + LSTM	0.96	0.95	0.94	0.96
MobileNet V2 + LSTM	0.95	0.93	0.93	0.93
Xception + LSTM	0.96	0.97	0.92	0.97
Baseline CNN + LSTM	0.95	0.94	0.94	0.94
LBCNN + LSTM	0.97	0.98	0.93	0.98

Figure 9: Time-lapse embryo development.

Table 9: Diagnosis of embryo anomaly in different periods based on accuracy.

Model	Time				
	24 h	36 h	48 h	60 h	72 h
VGG16 + LSTM	30.4	42.8	69.2	70.1	78.9
LeNet 5 + LSTM	22.8	30.8	52.1	55.6	55.6
Resnet-50 + LSTM	31.9	48.7	69.3	67.1	80.3
Inception V3 + LSTM	30.9	41.9	67.2	66.9	79.7
MobileNet V2 + LSTM	29.3	35.7	59.1	65.1	65
Xception + LSTM	30.1	47.6	61.3	66.9	73.9
Baseline CNN + LSTM	22.8	36.8	60.2	62.3	70.1
LBCNN + LSTM	**30.4**	**47.9**	**69.8**	**75.2**	**82.8**

The proposed approach (LBCNN+LSTM) experimental results declared as bold values.

Table 10: Results and discussion based on different types of approach and methodology.

Authors	Year	Device	Input image	Results	Models
Kragh et al. [26]	2019	Time-lapse incubator	Video: from days 1 to 5	Quality analysis: AUC = 0.96	Approach of CNN-LSTM
Khosravi et al. [23]	2019	Time-lapse incubator	Blastocyst	Quality analysis: AUC = 0.98	Google's Inception model
Tran et al. [22]	2019	Time-lapse incubator	Blastocyst transfer	Abnormality analysis: AUC = 0.93	Deep learning approach
Dirvanauskas et al. [27]	2019	Time-lapse incubator	Video: from days 1 to 6	Abnormality analysis: AUC = 0.98	Two-classifier vote-based CNN
Lee et al. [25]	2021	Time-lapse incubator	Video: from days 1 to 5	Abnormality analysis: AUC = 0.74	Deep learning approach
Sawada et al. [19]	2021	Time-lapse incubator	Blastocyst	Abnormality analysis: AUC = 0.93	LSTM with attention map
Payá et al. [40]	2022	Time-lapse incubator	Video: from days 1 to 4	Abnormality analysis: AUC = 0.94	A supervised contrastive learning framework
Our proposed method	2022	Time-lapse incubator	Video: from days 1 to 5	Abnormality analysis: AUC = 0.98	LBCNN-LSTM

superficial and unreliable because the researchers used different databases in terms of size and number of items.

4. Conclusion

This study presented the approach by fully automated deep learning to analyze the blastocyst morphology in case of anomaly detection from time-lapse imaging of human embryos. In this paper, we presented a novel deep learning approach, namely, LBCN-LSTM. This approach achieved significant results in case of accuracy of classification and ROC curve analysis compered to existing well-known pretrained models and state-of-the-art algorithms. The main advantage of this model is utilizing deep and local features in end-to-end manner with employing fewer number of trainable parameters compared to baseline CNN. In addition, this model can detect embryo anomaly based on Gardner's classification in early stage (stage 1) based on blastocyst stage table with higher accuracy rate compared to existing models. The results showed that the proposed LBCNN-LSTM model can be an efficient model for the real-life application regarding the accuracy of the diagnosis, process cost, and early detection of the abnormality of the human embryo in time-lapse incubator.

Data Availability

The human embryos dataset used to support the findings of this study have been deposited in the Sajad EINY repository (sajadeiny@aydin.edu.tr). This dataset is available under certain terms and conditions upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Rawat, A. Singh, H. S. Bhadauria, and J. Virmani, "Computer aided diagnostic system for detection of leukemia using microscopic images," *Procedia Computer Science*, vol. 70, pp. 748–756, 2015.

[2] A. Elaanba, M. Ridouani, and L. Hassouni, "A stacked generalization chest-X-ray-based framework for mispositioned medical tubes and catheters detection," *Biomedical Signal Processing and Control*, vol. 79, Article ID 104111, 2023.

[3] C. Yu, X. Qin, Y. Chen, J. Wang, and C. Fan, "Drowsydet: a mobile application for real-time driver drowsiness detection," in *Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pp. 425–432, Leicester, UK, August, 2019.

[4] S. Hussain, *Survey on Current Trends and Techniques of Data Mining Research London*, Journals Press, London, UK, 2017.

[5] H. Yu, N. Zhao, P. Wang, H. Chen, and C. Li, "Chaos-enhanced synchronized bat optimizer," *Applied Mathematical Modelling*, vol. 77, pp. 1201–1215, 2020.

[6] M. Abdar, V. N. Wijayaningrum, S. Hussain, R. Alizadehsani, and P. Plawiak, "A novel improved machine learning-based system for wart disease treatment," *Journal of Medical Systems*, vol. 43, 2019.

[7] H. Behzadi-khormouji, H. Rostami, S. Salehi et al., "Deep learning, reusable and problem-based architectures for detection of consolidation on chest X-ray images," *Computer Methods and Programs in Biomedicine*, vol. 185, Article ID 105162, 2020.

[8] S. Einy, H. Saygin, H. Hivehch, and Y. Dorostkar Navaei, "Local and deep features based convolutional neural network frameworks for brain MRI anomaly detection," *Complexity*, vol. 2022, Article ID 3081748, 11 pages, 2022.

[9] T. B. Chandra, K. Verma, B. K. Singh, D. Jain, and S. S. Netam, "Automatic detection of tuberculosis related abnormalities in Chest X-ray images using hierarchical feature extraction scheme," *Expert Systems with Applications*, vol. 158, Article ID 113514, 2020.

[10] W. Rasheed and T. B. Tang, "Anomaly detection of moderate traumatic brain injury using auto-regularized multi-instance one-class SVM," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 1, pp. 83–93, 2020.

[11] X. K. Wang, W. Hou, H. Zhang et al., "KDE-OCSVM model using Kullback-Leibler divergence to detect anomalies in

medical claims," *Expert Systems with Applications*, vol. 200, Article ID 117056, 2022.

[12] R. Moradi Rad, P. Saeedi, J. Au, and J. Havelock, "Cell-net: embryonic cell counting and centroid localization via residual incremental atrous pyramid and progressive upsampling convolution," *IEEE Access*, vol. 7, pp. 81945–81955, 2019.

[13] Y. Liu, V. Chapple, P. Roberts, J. Ali, and P. Matson, "Time-lapse videography of human oocytes following intracytoplasmic sperm injection: events up to the first cleavage division," *Reproductive Biology*, vol. 14, no. 4, pp. 249–256, 2014.

[14] Z. Liu, B. Huang, Y. Cui et al., "Multi-task deep learning with dynamic programming for embryo early development stage classification from time-lapse videos," *IEEE Access*, vol. 7, pp. 122153–122163, 2019.

[15] E. Alexopoulou, A. Pinborg, E. Budtz-Jørgensen, and A. Zedeler, "Comparing early embryo morphokinetics with time-lapse microscopy in patients with low and normal ovarian response to ovarian stimulation," *Reproductive Biology*, vol. 19, no. 2, pp. 127–132, 2019.

[16] W. Zhao, "Clinical outcomes comparison of single fresh and frozen-thawed superior blastocyst transfer," *International Journal of Clinical and Experimental Medicine*, vol. 10, no. 8, pp. 12605–12608, 2017.

[17] A. P. Tartia, C. Q. Wu, J. Gale, D. Shmorgun, and M. C. Léveillé, "Time-lapse KIDScoreD5 for prediction of embryo pregnancy potential in fresh and vitrified-warmed single-embryo transfers," *Reproductive BioMedicine Online*, vol. 45, no. 1, pp. 46–53, 2022.

[18] S. M. Diakiw, J. M. Hall, M. VerMilyea et al., "An artificial intelligence model correlated with morphological and genetic features of blastocyst quality improves ranking of viable embryos," *Reproductive BioMedicine Online*, vol. 45, no. 6, pp. 1105–1117, 2022.

[19] Y. Sawada, T. Sato, M. Nagaya et al., "Evaluation of artificial intelligence using time- lapse images of IVF embryos to predict live birth," *Reproductive BioMedicine Online*, vol. 43, no. 5, pp. 843–852, 2021.

[20] J. Firmin and J. L. Maître, "Morphogenesis of the human preimplantation embryo: bringing mechanics to the clinics," *Seminars in Cell & Developmental Biology*, vol. 120, pp. 22–31, 2021.

[21] T. T. Huang, D. H. Huang, H. J. Ahn, C. Arnett, and C. T. Huang, "Early blastocyst expansion in euploid and aneuploid human embryos: evidence for a non-invasive and quantitative marker for embryo selection," *Reproductive BioMedicine Online*, vol. 39, no. 1, pp. 27–39, 2019.

[22] D. Tran, S. Cooke, P. J. Illingworth, and D. K. Gardner, "Deep learning as a predictive tool for fetal heart pregnancy following time-lapse incubation and blastocyst transfer," *Human Reproduction*, vol. 34, no. 6, pp. 1011–1018, 2019.

[23] P. Khosravi, E. Kazemi, Q. Zhan et al., "Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization," *Npj Digital Medicine*, vol. 2, no. 1, pp. 21–29, 2019.

[24] C. C. Wong, K. E. Loewke, N. L. Bossert et al., "Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage," *Nature Biotechnology*, vol. 28, no. 10, pp. 1115–1121, 2010.

[25] C. I. Lee, Y. R. Su, C. H. Chen et al., "End-to-end deep learning for recognition of ploidy status using time-lapse videos," *Journal of Assisted Reproduction and Genetics*, vol. 38, no. 7, pp. 1655–1663, 2021.

[26] M. F. Kragh, J. Rimestad, J. Berntsen, and H. Karstoft, "Automatic grading of human blastocysts from time-lapse imaging," *Computers in Biology and Medicine*, vol. 115, Article ID 103494, 2019.

[27] D. Dirvanauskas, R. Maskeliunas, V. Raudonis, and R. Damasevicius, "Embryo development stage prediction algorithm for automated time lapse incubators," *Computer Methods and Programs in Biomedicine*, vol. 177, pp. 161–174, 2019.

[28] J. Conaghan, A. A. Chen, S. P. Willman et al., "Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: results from a prospective multicenter trial," *Fertility and Sterility*, vol. 100, no. 2, pp. 412–419.e5, 2013.

[29] D. Dirvanauskas, R. Maskeliūnas, V. Raudonis, R. Damaševičius, and R. Scherer, "HEMIGEN: human embryo image generator based on generative adversarial networks," *Sensors*, vol. 19, pp. 3578–3616, 2019.

[30] F. Juefei-Xu, V. N. Boddeti, and M. Savvides, "Local binary convolutional neural networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4284–4293, Honolulu, HI, USA, July 2017.

[31] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657–1663, 2010.

[32] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, Article ID 132306, 2020.

[33] M. Dua, D. Makhija, P. Y. L. Manasa, and P. Mishra, "A CNN–RNN–LSTM based amalgamation for Alzheimer's disease detection," *Journal of Medical and Biological Engineering*, vol. 40, no. 5, pp. 688–706, 2020.

[34] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015*, pp. 1–14, San Diego, CA, USA, May 2015.

[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[36] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, Las Vegas, NV, USA, June 2016.

[37] A. G. Howard, "Mobilenets: efficient convolutional neural networks for mobile vision applications," 2017, http://arxiv.org/abs/1704.04861.

[38] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1800–1807, Honolulu, Hawaii, July 2017.

[39] D. P. Kingma and J. L. Ba, "ADAM: a method for stochastic optimization," *Conference Paper ICLR 2015 ADAM*, vol. 15, pp. 1–15, 2015.

[40] E. Payá, L. Bori, A. Colomer, M. Meseguer, and V. Naranjo, "Automatic characterization of human embryos at day 4 post-insemination from time-lapse imaging using supervised contrastive learning and inductive transfer learning techniques," *Computer Methods and Programs in Biomedicine*, vol. 221, Article ID 106895, 2022.