

Deepfake Detection Based on Deep Learning and Quantum Particle Swarm Optimization

Ebeid A. Ebeid¹, Ahmed Sobhi^{2,a}, H. A. El Shenbary³

^{1,2,3}Department of Mathematics, Faculty of Science, Al-Azhar University, Nasr-City, Cairo, Egypt.

^aCybersecurity Department, Engineering and Information Technology College, Buraydah Private Colleges, Buraydah, 51418, Saudi Arabia

Article Info

Article history:

Received Aug 28, 2025

Revised Apr 2, 2026

Accepted Apr 16, 2026

Keywords:

Deepfake

deep learning

ResNet

Celeb-DF

Real and Fake Face Detection and Quantum Particle Swarm Optimization (QPSO).

ABSTRACT

Recently, many fake images and videos can be created and manipulated easily by employing Artificial Intelligence technologies, mainly based on Generative Adversarial Network (GAN). These synthesized media (images or videos) can be used later in unethical ways to create panic among individuals. This type of technology, which synthesizes human images and videos, is called deepfake. This paper proposes a deepfake detection methodology relying on Deep Learning (DL) approach and quantum particle swarm optimization (QPSO). The input image undergoes preprocessing by choosing the Region of Interest (ROI), followed by the application of an image downsampling algorithm to minimize the dimension of the new sub-image and thus decrease processing time costs. After that, the image has been normalized. The optimal subset of dataset has been determined by adapting QPSO during training phase. To perform classification of images and videos, ResNet50 model accomplishes this task perfectly. Experiments indicate that the suggested algorithm demonstrates a significant level of accuracy on "Real and Fake Face Detection" and "Celeb-DF" datasets. The proposed approach achieves accuracy 99.6 % on Celeb-DF and 99.8 % on Real and Fake Face Detection dataset.

Copyright © 2026 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Hassan Ahmed El Shenbary

Department of Mathematics, Faculty of Science, Al-Azhar University.

Nasr-City, Cairo, Egypt.

Email: h.a.elshenbary@azhar.edu.eg

1. INTRODUCTION

Deepfake is a phenomenon that is rapidly growing in recent years. This technology has the ability to generate automatically altered or swapped facial images and videos of any person. Deepfake can convincingly depict individuals, particularly public figures, stating or performing actions they did not undertake. It can be utilized to produce false information, influence elections, and create divisive content. Additionally, it can be used to bypass biometric security measures like facial recognition enabling unauthorized access to accounts, services and fraudulent financial transactions. Based on artificial intelligence technology (specifically DL), system can learn patterns from the data and then use it in many purposes such as classification and segmentation as in [1], [2] or generates new multimedia contents that cannot be easily revealed as real or fake by human eye. GANs are the most DL networks utilized to generate high quality deepfake images [3]. Initially, the images of deepfake were distinguished by the human eye because of the variation in pixels that creates an artificial visual inconsistency in the facial shape of image. Step by step, deepfake technology has been improved and carried out on images, videos and audio. These advances in deepfake versions have become almost indistinguishable from original photos [4]. On the other side, the begining of twentyfirst century has watched a widespread of

social networking development. Therefore, deepfake content has spread faster than ever before [5]. Relying on deepfake technology, people have the ability to select their clothing more efficiently, which aids the fashion and online retail sectors. In addition, this technology can help in some other business by offering synthetic voices for performers who are unable to dub promptly. In the medical field, deepfake technology may help potentially Alzheimer's patients to retain their memories by communicating with earlier version of them. In addition, GAN can be used to detect anomalies in X-ray images [6]. Several methods have been introduced to detect deepfake images and videos including the traditional classifiers of machine learning (e.g. Support Vector Machine(SVM)) and deep neural networks include but not limited to convolutional neural networks (CNNs) (e.g. DenseNet, VGG and ResNet), recurrent neural networks (RNN) and Long Short-Term Memory (LSTM). The subtle and local tiny variations of fake faces make it more similar to fine-grained visual classification problem [7]. A lot of methods have been proposed to enhance the accuracy and beat the challenges in deepfake images or videos. These approaches include but limited to: Guarnera et al. [8] examined bogus videos of human faces to propose a new discernment approach that has the ability to detect forensics trail buried in photos. Xin et al. [9] presented a system for detecting AI-generated forgeries in face images or videos. They compared head locations computed by by employing visual cues for those evaluated using solely the central region. Falko et al. [10] proposed an approach to collect characteristics to identify produced faces, Face2Face and deepfake images in the teeth, eyes, and facial contours. Nadimpalli et al. [11] proposed a method that chooses the top-k augmentations when test each sample. Because of the performance may degrade significantly during the evaluation across datasets, an reinforcement learning (RL) agent is used in an image-specific manner. Guera et al. [12] have employed recognition pipeline to detect deepfakes automatically. The proposed method composed of two steps. Firstly, after preprocessing important features are selected at frame level using CNN. RNN captures the introduced erratic frames for face swapping process. Ciftci et al. [13] have proposed a method to extract important features from the facial regions in the fake and real video pairs. SVM and CNN have been trained to find probabilities of authenticity. Fran et al. [14] provide an examination of inception modules in CNNs. Inception modules are used as an intermediate step in the midst of regular convolution and the depthwise separable convolution operations. A pre-trained ResNet-50 to generate images using computer graphics has been initiated by Rezende et al. [15]. The uppermost layer was substituted with a fully connected layer then SVM was used as a classifier for classification purposes. Dong et al. [16] developed a robust method of two-path deep semisupervised learning. First path was utilized for the supervised learning and the second was used for the unsupervised learning case. The purpose of supervised learning path was to learn in case of small amount of data having labels. An unsupervised path was applied to learn in case of a huge amount of data and the data are unlabeled. Authors in [17] developed a new system by using a hybrid CNN composed of Xception and InceptionResnet v2 for extracting features from frames. İlhan et al. [18] in their study proposed a feature extraction and classification approach. The introduced method has been experimented on the dataset with NASNetLarge CNN model. Nguyen et al. [19] have detected deepfakes based on CNN combined with LSTM model. The major constraints and limitations in most previous studies are the usage of the whole image despite of the little information embedded on the boundaries of image. Furthermore, many works do not have a high degree of accuracy that enables them to efficiently distinguish between original and fake images or videos. In addition, the authors implemented their models on the full dataset to identify the model performance. One factor that can affect the model performance is selecting an optimal dataset for training. When dataset is divided into many subsets, some of these subsets can obtain higher performance comparing with the others. Finding the subset of data points that maintain a similar level of quality as the full dataset has been proposed in few literatures which handle this issue as in study [20], [21]. Using a subset of data can speedup the training process with only a very small accuracy degradation. By employing appropriate optimization algorithms we can select an optimal subset of data and execute a smaller number of epochs to get good approximation of the boundary of the decision target model even faster. The primary contributions of this research are: 1) selecting the ROI in the facial image to ameliorate the performance. 2) Reducing the image size by down sampling technique to reduce the computation time. 3) Adapting QPSO to select the optimal subset of training data to enhance time cost of model and reduce the required storage. 4) Adapting ResNet50 as a deep learning model to enhance feature extraction problem and classification task accuracy. The remainder structure of this article is prepared as follows: In section two, the details of ResNet50 model, QPSO methods have been explained. The contributed proposed methodology is illustrated in Section 3. Section 4 introduces the main results and discussion. In section 5, the paper's conclusion is presented.

2. METHODS

The proposed method is based on ResNet-50 and QPSO components. Resnet-50 model is used basically to extract important features from the input data (images or videos), then their features are used to complete the classification task. Resnet-50 is considered as a convolutional neural network (CNN) mainly consists of 50 layers. On the other hand, QPSO is a powerful and effective computational method that enhances a problem solution by progressively refining a potential solution. It explores a wider search space comparing with PSO and increases the ability to find a global optimum. The basic details have been explained as follows:

2.1. ResNet

ResNet is a DL architecture introduced by He Kaiming et al. in 2015 [22]. It has been utilized in many advanced works to handle many issues [23–29]. It has the ability to mitigate the issue of vanishing gradient that can occur in most deep networks. It can accomplish this task by skipping connections (residual connections) that enable the network to “understand” residual mappings. The variance between the expected result and the present result at every layer can be learned by these residual connections. This procedure helps the network to hold the important information from early layer during training. One of the significant variants of ResNet is ResNet50. It integrates of 50 layers organized into 5 blocks, and each block containing a collection of residual blocks. ResNet50 is an effective and robust deep neural network, offering benefits compared to other networks, making it appropriate for different tasks. The three types of network layers are described in next subsections:

2.1.1. Convolutional Layers

The initial layer of the network is a convolutional layer. It has many filters to execute convolution on the input data. The convolutional layer is then followed by a max-pooling layer with stride 2 to reduce the size of the output of such layer. The obtained output of the pooling layer is then given to the next residual blocks layers.

2.1.2. Residual Blocks

Every residual block is made up of two convolutional layers, which are then succeeded by a batch normalization layer and a rectified linear unit (ReLU) activation function. The output of these layers are considered as the input of the residual block and given to ReLU function. The last output of the residual block is then sent to the next block.

2.1.3. Fully Connected Layer

The third type of layers is the fully connected layer which represents the final layer of the network. It takes the results of the final residual block and convert it to the output categories.

2.2. Quantum Particle Swarm Optimization

Particle Swarm Optimization (PSO) is the method which mimics the action of various types of animal societies such as: Shoals of fishes and groups of birds as a stochastic optimization [30]. The QPSO algorithm is the counterpart of traditional PSO that mimics the action of particles in the realm of quantum mechanics[31]. The basic characteristic of QPSO against PSO is the smaller number of parameters needed through adaptation process [32]. In quantum space, the square of the wave function $|\psi|^2$ signifies the likelihood density of particles existing at a certain location. The probability density function (PDF) attains scaling requirements as follows:

$$\int_{-\infty}^{+\infty} |\psi|^2 dx dy dz = 1 \quad (1)$$

Utilizing stochastic simulation based on Monte Carlo principles on Eq 1. The location of the particle x can be determined in the following way:

$$x = q + \frac{L}{2} \cdot \ln \left(\frac{1}{u} \right) \quad (2)$$

While every particle possesses D dimensions in quantum space), the position of any particle is depicted by the vector $x = (x_1, x_2, \dots, x_D)$. parameter u denotes a random number evenly distributed between 0 and 1 as in [31, 33]. The parameter q denotes a local attractor that possesses the elements $(q_1, q_2, \dots, q_D)$ having D dimensions for any particle. Parameter q enables PSO to convergence to the correct position. The local attractor of i th particle is performed by:

$$q_{id} = \alpha pbest_{id} + (1 - \alpha) gbest_d \quad (3)$$

Where d from 1 to D and I from 1 to n . Respectively $gbest$ and $pbest$ are the global best position and best position [?]. Parameter α is a stochastic value within the range $[0, 1]$. In Eq. 2, the variable L can be computed as the formula $L = 2\beta |mbest_{id} - x_{id}|$. The variable β refers to the "Contraction-Expansion Coefficient" [30] while $mbest$ indicates to the average value of the optimal positions of all particles which employed to prevent early convergence. Finally, each particle updates its own position according the equation:

$$x_{id}(t+1) = q_{id} + \beta |mbest_d - x_{id}(t)| \ln\left(\frac{1}{u}\right) \quad (4)$$

Relying on Eq. 4, the position of the particle in the subsequent iteration relies on the local attractor's mean value of all positions and the correspond position of particles in the present iteration. Depending on the problem at hand, QPSO uses a fitness function to choose the optimal value which is $gbest$.

3. METHODOLOGY

In this work, the input image is preprocessed to adapt and ameliorate its quality. Then, the optimal subset of training data has been determined by adapting QPSO. Finally, ResNet50 model performs binary classification of images or videos into real or fake class. Figure 1 illustretes the suggested methodology

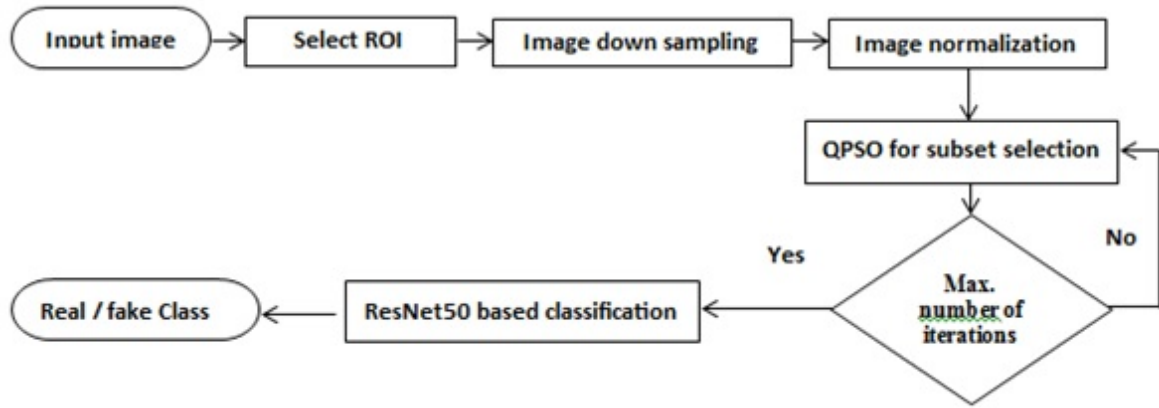


Figure 1. Proposed methodology for deepfake detection

The structure of the proposed algorithm is outlined below:

- **Image Pre-processing:** in this step the facial image (or frame) is corrupted to create Region of Interest (ROI). Selecting ROI mitigates redundant facial regions and boosts the overall efficiency of the model. The employed approach is named as Single Shot MultiBox Detector (SSD) which constructs a single profound neural network for identifying objects [34]. It presents better accuracy and faster than some of other approaches. Then, image down sampling approach is applied to enhance the computations complexity and optimize storage required for the model execution by reducing image size. The Nearest Neighbor algorithm is utilized in this work because it is one of the simplest and computationally efficient down sampling techniques. It works by selecting the value of the nearest pixel in the original image to represent the down sampled pixel [35]. After that, a normalization process is employed to make the pixels' values in the interval $[0,1]$. The utilized approach in this work is called standardization. If I refers to the input image then:

$$\text{Normalization}(I) = \frac{I - \mu(I)}{\sigma(I)} \quad (5)$$

Where σ is the standard deviation. Thus, this transformation on the original data converts all the scaled data into the range $[0, 1]$.

- Data selection: During training, incoming data are being shuffled before feeding into the model. QPSO algorithm is employed to shuffle and select the optimal data based on the model performance which determined by the fitness function. In this work, training accuracy represents the fitness function. QPSO improves the model's efficiency by determining the optimal size of the subset that is representative of the overall dataset. The elected optimal subset is smaller than the complete data which accelerate the model training. The the selection of QPSO parameters' values such as population size, maximum iteration and β coefficient is determined by values in previous study [36] and then adapted experimentally based on the result of training accuracy of our model. The overall process of QPSO steps are as follows:

1. Initialize QPSO by selecting a subset of data randomly. Each subset of training images represents a particle in the QPSO algorithm.
2. Calculate the fitness function for every particle.
3. Global and local bests are updated.
4. Data points are also updated according to Eq. 4.
5. Continue repeating Steps 2 through 4 until the stopping criteria are met.

The evaluation process can be speeded up by using one of many methods [37–39].

- 50 based classification: The process of classification is accomplished by using ResNet50 as a binary classifier. The early layers extract features from image to produce feature map then fully connected layer produce the final output by classifying the input image into "real" or "fake" class.

4. RESULT AND DISCUSSION

In this work we have implemented our model on two public datasets: "Real and Fake Face Detection" and "Celeb-DF" datasets. "Real and Fake Face Detection" has been collected by Department of Computer Science, Yonsei University [40]. It composed of 140,000 images divided into three directories; train, test and valid. Each directory contains 256×256 real/fake face photos, respectively. The other dataset is Celeb-DF" dataset [41] which includes 590 authentic videos gathered from YouTube showcasing people of different eras, ethnicities, and genders, alongside 5639 related deepfake videos

Frames have different dimensions (around 500×500). Our experiments were executed on a GPU using Google Colab notebook. During training, data is shuffled before being fed into the model. In this process, the model can produce different results according to the order of the training data. QPSO is used to shuffle data and select the best subset of training data to obtain the best accuracy.

To enhance the efficiency of our model, we adapt the parameters as follows: Population size = number of particles (data batches in our model), Maximum iterations = 10 iterations, and $\beta = 0.75$. The size of the selected subset is around 70% of the total dataset.

The fitness function is determined by computing the accuracy of training. The basic parameters of ResNet50 have been adapted as follows: learning rate $\in [0.00001, 0.001]$, dropout $\in [0.4, 0.6]$, weight decay = 0.00001, optimizer = Adam, activation function = ReLU, pooling = average pooling, and loss function = cross entropy. Table 1 shows QPSO Convergence and Fitness Optimization on "Real and Fake Detection" and "Celeb-DF" datasets.

Table 1. QPSO Convergence and Fitness Optimization on "Real and Fake Detection" and "Celeb-DF" datasets

Iteration (t)	1	2	3	4	5	6	7	8	9	10
Fitness1 (f_{min})	0.266	0.190	0.095	0.07	0.033	0.006	0.004	0.002	0.002	0.002
Fitness2 (f_{min})	0.284	0.156	0.082	0.045	0.036	0.023	0.007	0.005	0.004	0.004

Figure 2 shows the convergence of QPSO on both datasets where QPSO was configured with a maximum of 10 iterations. The stability of the proposed model has been achieved in the 8th and 9th iteration with no further fluctuations observed in subsequent iterations. This confirms that the search space was sufficiently explored to achieve the optimal feature subset.

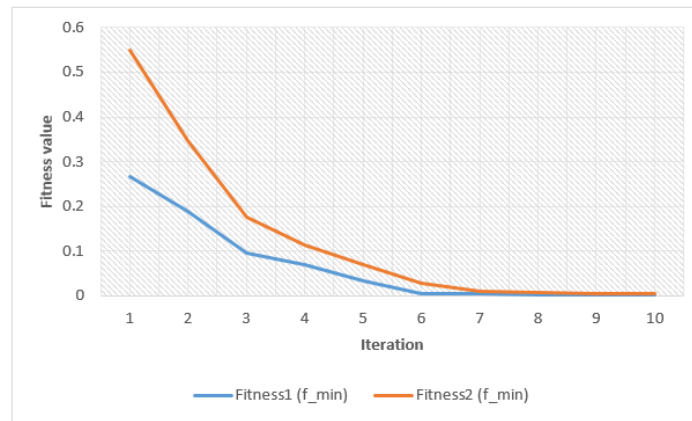


Figure 2. Convergence curve of QPSO on "Real and Fake Detection" as Fitness1 and "Celeb-DF" dataset as Fitness2 through 10 iterations

The next tables 2 and 3 show the results before and after applying QPSO on the datasets. Table 2 shows different image dimensions and the accuracy obtained on the "Real and Fake Detection" dataset. The best accuracy is achieved at an image size of 175×175 . After the fifth epoch, the best accuracy obtained is 99.8% after applying QPSO on the dataset with 1.9 % improvement. At an image size of 64×64 , the model achieves the lowest accuracy, which equals 97%.

Table 2. The performance at different dimensions on Real and Fake Detection dataset

Dimensions	Accuracy without QPSO	Accuracy with QPSO	Epoch
64×64	95.8	97.0	6
96×96	96.1	97.9	6
128×128	97.4	99.4	6
156×156	97.6	99.4	7
175×175	97.9	99.8	5

Table 3 shows the performance of the classifier at different dimensions on the Celeb-DF dataset. The best accuracy has been achieved at the size 150×150 at the seventh epoch after applying QPSO on the dataset with 2.1 % improvement.

Table 3. The performance at different dimensions on Celeb-DF dataset

Dimensions	Accuracy without QPSO	Accuracy with QPSO	Epoch
56×56	84.0	84.6	9
112×112	93.5	94.6	8
124×124	95.4	97.3	8
150×150	97.5	99.6	7
170×170	98.1	99.3	7

The performance at different dimensions on the utilized datasets has been presented graphically in figure 3.



Figure 3. The performance at different image dimensions on "Real and Fake Detection" and "Celeb-DF" dataset

The results which are achieved with ResNet50 classifier are better than some other algorithms such as but are not limited to VGG16, VGG19 and DenseNet169. For example, the model size of ResNet50 (when loaded from PyTorch framework) equals 97.8 MB whereas the size of VGG16 model is 528 MB, VGG19 has slightly larger file size than VGG16, as it has more layers. So it will be in the range of 549 MB. Thus, the big difference in the model size offers a big gap on the computation times for both models. The computation times of VGG16 and VGG19 are more than the double of ResNet50 computation time. DenseNet169 is 1.5 times of ResNet50 computationally. This comparison of computation times of ResNet50, VGG16, VGG19 and DenseNet169 models has been obtained by implementing them on the same machine. In the next tables, the results of some previous works are shown. Tables 4, show some of such results obtained by implementing different approaches on "Real and Fake detection dataset".

Table 4. Related works comparison in deepfake detection based on Real and Fake detection dataset

Work	Technique	Acc. (%)
[42]	VGG-16	91.97
[42]	VGG-19	92.09
[43]	IQIEA-FS	Mean normalized accuracy = 0.583
Proposed	ResNet50 + QPSO	99.8

Therefore, the accuracy and speed of ResNet50 are the major factors to make it the preferable choice when compared with VGG models. Table 5 introduces the results which are obtained by implementing different approaches on "Celeb-DF" dataset. As noted, most of the previous works utilize CNN model because it is considered the best DL model which handle the images or videos.

Table 5. Related work results comparison in deepfake detection based on Celeb-DF dataset

Work	Technique	Acc. (%)
[44]	EfficientNet-B4	64.29
[7]	Multi-attentional	67.44
[45]	InceptionResNetV2 + LSTM	91.48
[46]	CNN	98.90
[47]	CNN	93.70
[48]	CNN	95.86
[17]	Hybrid CNNs	95.75
Proposed	ResNet50 + QPSO	99.60

Finally, the proposed algorithm has succeeded in minimizing the image size in the "Real and Fake Face Detection" dataset from 256×256 to 170×170 . Whereas, the dimensions of the "Celeb-DF" images have been reduced from 500×500 to 150×150 pixels. When the input image is preprocessed by selecting the Region of Interest (ROI) and downsampling, the time cost of processing is improved effectively. Instead of employing the whole dataset in training, the proposed methodology has determined the optimal subset of data for future usage using the QPSO algorithm. By implementing the ResNet50 model for feature extraction and classification, we have achieved accuracies of 99.8% and 99.6% on both datasets, exceeding other approaches.

5. CONCLUSION AND FUTURE WORK

This paper proposed an efficient methodology for deepfake detection based on DL and quantum optimization approaches. The proposed architecture significantly enhance the process of deepfake detection by enabling both ResNet50 and QPSO techniques. The input image is preprocessed by selecting ROI then image downsampling method has been modified to decrease the dimensions of the newly created sub-image to decrease the whole computation time required. After that, the image has been normalized. The optimal subset of data has been determined by adapting QPSO during training phase. To perform feature extraction and classification, ResNet50 network has accomplished this task perfectly. Experiments have been implemented on images and videos datasets. Findings indicated that the suggested algorithm has a substantial Level of accuracy on both datasets and outperforms many other DL based approaches. In the future improvements, it is planned to merge another approaches to improve the system performance especially when deepfakes are very realistic or difficult to detect. Quantum CNN method is a promising approach that can enhance the results in such related area.

ACKNOWLEDGEMENTS

We express our heartfelt appreciation to the Faculty of Science, Al-Azhar University, Cairo, Egypt, for its essential support for both scientific and resources that were crucial to the success of this work.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

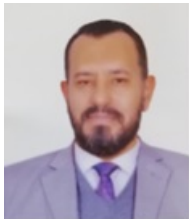
- [1] N. Atitallah, A. Abdel-Wahab, A. A. Hadi, H. Abdel-Jaber, A. W. Mohamed, M. Elserisy, and Y. Mansour, "Horsenet: a novel deep learning approach for horse health classification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 38, no. 1, pp. 555–568, 2025. [Online]. Available: <https://doi.org/10.11591/ijeecs.v38.i1.pp555-568>
- [2] E. A. Elsoud, O. Alidmat, S. Abuowaida, E. Alhenawi, N. Alshdaifat, A. Aburomman, and H. Y. Chan, "Accurate segmentation of fruit based on deep learning," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 38, no. 2, pp. 1331–1338, 2025. [Online]. Available: <https://doi.org/10.11591/ijeecs.v38.i2.pp1331-1338>
- [3] H. S. Shad, M. M. Rizvee, N. T. Roza, S. M. Hoq, M. M. Khan, A. Singh, and S. Bourouis, "Comparative analysis of deepfake image detection method using convolutional neural network," *Computational Intelligence and Neuroscience*, vol. 2021, 2021. [Online]. Available: <https://doi.org/10.1155/2021/3111676>
- [4] T. Jung, S. Kim, and K. Kim, "Deepvision: deepfakes detection using human eye blinking pattern," *IEEE Access*, vol. 8, pp. 83 144–83 154, 2020. [Online]. Available: <https://doi.org/10.1109/ACCESS.2020.2988660>
- [5] T. Nguyen, Q. Nguyen, C. M. Nguyen, D. Nguyen, D. Nguyen, and S. Nahavandi, "Deep learning for deepfakes creation and detection: a survey," *Computer Vision and Image Understanding*, pp. 1–17, 2019. [Online]. Available: <https://doi.org/10.1016/j.cviu.2022.103525>
- [6] M. Westerlund, "The emergence of deepfake technology: a review," *Technology Innovation Management Review*, vol. 9, no. 11, pp. 39–52, 2019. [Online]. Available: <https://doi.org/10.22215/timreview/14>
- [7] H. Zhao, W. Zhou, D. Chen, T. Wei, W. Zhang, and N. Yu, "Multi-attentional deepfake detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 2185–2194. [Online]. Available: <https://doi.org/10.1109/CVPR46437.2021.00222>
- [8] Y. Li and S. Lyu, "Exposing deepfake videos by detecting face warping artifacts," *arXiv preprint arXiv:1811.00656*, 2018. [Online]. Available: <https://doi.org/10.48550/arXiv.1811.00656>
- [9] F. Matern, C. Riess, and M. Stamminger, "Exploiting visual artifacts to expose deepfakes and face manipulations," in *2019 IEEE Winter Applications of Computer Vision Workshops (WACVW)*, 2019, pp. 83–92. [Online]. Available: <https://doi.org/10.1109/WACVW.2019.00020>
- [10] D. Güera and E. J. Delp, "Deepfake video detection using recurrent neural networks," in *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2018, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/AVSS.2018.8639163>
- [11] A. V. Nadimpalli and A. Rattani, "On improving cross-dataset generalization of deepfake detectors," *arXiv preprint arXiv:2204.04285*, 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2204.04285>
- [12] D. Güera and E. J. Delp, "Deepfake video detection using recurrent neural networks," in *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2018, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/AVSS.2018.8639163>

- [13] U. A. Ciftci, I. Demir, and L. Yin, "Detection of synthetic portrait videos using biological signals," *arXiv preprint arXiv:1901.02212*, 2021. [Online]. Available: <https://doi.org/10.48550/arXiv.1901.02212>
- [14] C. Fran, "Deep learning with depth wise separable convolutions," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017. [Online]. Available: <https://doi.org/10.48550/arXiv.1610.02357>
- [15] E. R. S. Rezende, G. C. S. Ruppert, and T. Carvalho, "Detecting computer generated images with deep convolutional neural networks," in *30th Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2017, pp. 71–78. [Online]. Available: <https://doi.org/10.1109/SIBGRAPI.2017.16>
- [16] X. Dong, U. Victor, and L. Qian, "Two-path deep semisupervised learning for timely fake news detection," *IEEE Transactions on Computational Social Systems*, pp. 1–13, 2020. [Online]. Available: <https://doi.org/10.1109/TCSS.2020.3027639>
- [17] S. Ikram, S. Thaseen, S. Chambial, and D. Sood, "A performance enhancement of deepfake video detection through the use of a hybrid cnn deep learning model," *International Journal of Electrical and Computer Engineering Systems*, vol. 14, no. 2, pp. 169–178, 2023. [Online]. Available: <https://doi.org/10.32985/ijeces.14.2.6>
- [18] I. İlhan, E. Balı, and M. Karaköse, "An improved deepfake detection approach with nasnetlarge cnn," in *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, 2022, pp. 598–602. [Online]. Available: <https://doi.org/10.1109/ICDABI56818.2022.10041558>
- [19] H. T. Nguyen, T. C. Dao, T. M. N. Phan, and T. T. Phan, "Fake face detection in video using shallow deep learning architectures," *International Journal of Intelligent Systems Technologies and Applications*, vol. 20, no. 6, pp. 469–486, 2022. [Online]. Available: <https://doi.org/10.1504/ijista.2022.128528>
- [20] C. Coleman, C. Yeh, S. Mussmann, B. Mirzasoleiman, P. Bailis, P. Liang, J. Leskovec, and M. Zaharia, "Selection via proxy: Efficient data selection for deep learning," *arXiv preprint arXiv:1906.11829*, 2019. [Online]. Available: <https://doi.org/10.48550/arXiv.1906.11829>
- [21] A. Mittal, D. Sivasubramanian, R. Iyer, P. Jyothi, and G. Ramakrishnan, "Partitioned gradient matching-based data subset selection for compute-efficient robust asr training," *arXiv preprint arXiv:2210.16892*, 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2210.16892>
- [22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. [Online]. Available: <https://doi.org/10.48550/arXiv.1512.03385>
- [23] M. Shafiq and Z. Gu, "Deep residual learning for image recognition: A survey," *Applied Sciences*, vol. 12, no. 18, p. 8972, 2022. [Online]. Available: <https://doi.org/10.3390/app12188972>
- [24] X. Li, L. Ding, L. Wang, and F. Cao, "Fpga accelerates deep residual learning for image recognition," in *2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2017, pp. 837–840. [Online]. Available: <https://doi.org/10.1109/ITNEC.2017.8284852>
- [25] S. Wu, S. Zhong, and Y. Liu, "Deep residual learning for image steganalysis," *Multimedia Tools and Applications*, vol. 77, pp. 10 437–10 453, 2018. [Online]. Available: <https://doi.org/10.1007/s11042-017-4440-4>
- [26] A. Jafar and L. Myungho, "Hyperparameter optimization for deep residual learning in image classification," in *2020 IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C)*, 2020, pp. 24–29. [Online]. Available: <https://doi.org/10.1109/ACSOS-C51401.2020.00024>
- [27] W. Bae, J. Yoo, and J. C. Ye, "Beyond deep residual learning for image restoration: Persistent homology-guided manifold simplification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 145–153. [Online]. Available: <https://doi.org/10.48550/arXiv.1611.06345>

- [28] Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, "Deep residual learning for image compression," in *CVPR Workshops*, 2019. [Online]. Available: <https://doi.org/10.48550/arXiv.1906.09731>
- [29] X. Cheng, Y. Zhang, Y. Chen, Y. Wu, and Y. Yue, "Pest identification via deep residual learning in complex background," *Computers and Electronics in Agriculture*, vol. 141, pp. 351–356, 2017. [Online]. Available: <https://doi.org/10.1016/j.compag.2017.08.005>
- [30] J. Sun, W. Xu, and W. Fang, "Quantum-behaved particle swarm optimization algorithm with controlled diversity," in *Computational Science – ICCS*, 2006, pp. 847–854. [Online]. Available: https://doi.org/10.1007/978-3-662-45526-5_13
- [31] Z. Huang, "Improved quantum particle swarm optimization for mangroves classification," *Journal of Sensors*, vol. 2016, p. 9264690, 2016. [Online]. Available: <https://doi.org/10.1155/2016/9264690>
- [32] S. Li, R. Wang, W. Hu, and J. Sun, "A new qpso based bp neural network for face detection," in *Fuzzy Information and Engineering*, 2007, pp. 355–363. [Online]. Available: <https://doi.org/10.1155/2016/9264690>
- [33] R. Misra and K. S. Ray, "Object tracking based on quantum particle swarm optimization," *arXiv preprint arXiv:1707.05228*, 2017.
- [34] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, and C. Berg, "Ssd: Single shot multibox detector," in *Computer Vision – ECCV 2016: 14th European Conference, Proceedings, Part I*, 2016, pp. 21–37. [Online]. Available: https://doi.org/10.1007/978-3-319-46448-0_2
- [35] "Image downsampling algorithms: A comprehensive guide," <https://saturncloud.io/blog/image-downsampling-algorithms-a-comprehensive-guide/>, accessed: 2024.
- [36] K. A. E. Dahshan, E. K. Elsayed, A. Aboshoha, and E. A. Ebeid, "Recognition of facial emotions relying on deep belief networks and quantum particle swarm optimization," *International Journal of Intelligent Engineering & Systems*, vol. 13, no. 4, 2020. [Online]. Available: <https://doi.org/10.22266/ijies2020.0831.09>
- [37] B. Wang, Y. Sun, B. Xue, and M. Zhang, "Evolving deep convolutional neural networks by variable-length particle swarm optimization for image classification," in *IEEE Congress on Evolutionary Computation (CEC)*, 2018, pp. 1–8. [Online]. Available: <https://doi.org/10.48550/arXiv.1803.06492>
- [38] T. Yamasaki, T. Honma, and K. Aizawa, "Efficient optimization of convolutional neural networks using particle swarm optimization," in *Proceedings of the Third IEEE International Conference on Multimedia Big Data (BigMM)*, 2017. [Online]. Available: <https://doi.org/10.1109/BigMM.2017.69>
- [39] Y. Wang, H. Zhang, and G. Zhang, "cpso-cnn: An efficient pso-based algorithm for fine-tuning hyper-parameters of convolutional neural networks," *Swarm and Evolutionary Computation*, 2019. [Online]. Available: <https://doi.org/10.1016/j.swevo.2019.06.002>
- [40] "Real and fake face detection dataset," <https://www.kaggle.com/datasets/ciplab/real-and-fake-face-detection>, accessed: 2024.
- [41] "Celeb-deepfakeforensics dataset," <https://github.com/yuezunli/celeb-deepfakeforensics>, accessed: 2024.
- [42] N. Nida, A. Irtaza, and N. Ilyas, "Forged face detection using ela and deep learning techniques," in *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST)*, 2021, pp. 271–275. [Online]. Available: <https://doi.org/10.1109/IBCAST51254.2021.9393234>
- [43] H. Mittal, M. Saraswat, J. C. Bansal, and A. Nagar, "Fake-face image classification using improved quantum-inspired evolutionary-based feature selection method," in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2020, pp. 989–995. [Online]. Available: <https://doi.org/10.1109/SSCI47803.2020.9308337>

- [44] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning (ICML)*, 2019, pp. 6105–6114. [Online]. Available: <https://doi.org/10.48550/arXiv.1905.11946>
- [45] P. Yadav, I. Jaswal, J. Maravi, V. Choudhary, and G. Khanna, "Deepfake detection using inceptionresnetv2 and lstm," in *International Conference on Emerging Technologies: AI, IoT, and CPS for Science Technology Applications*, 2021. [Online]. Available: <https://doi.org/10.1109/ICCMC56507.2023.10083584>
- [46] S. Agarwal, H. Farid, T. El-Gaaly, and S. Lim, "Detecting deep-fake videos from appearance and behavior," in *2020 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2020, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/WIFS49906.2020.9360904>
- [47] U. A. Ciftci, I. Demir, and L. Yin, "How do the hearts of deep fakes beat? deep fake source detection via interpreting residuals with biological signals," in *2020 IEEE International Joint Conference on Biometrics (IJCB)*, 2020, pp. 1–10. [Online]. Available: <https://doi.org/10.48550/arXiv.2008.11363>
- [48] P. Ranjan, S. Patil, and F. Kazi, "Improved generalizability of deep-fakes detection using transfer learning based cnn framework," in *2020 3rd International Conference on Information and Computer Technologies (ICICT)*, 2020, pp. 86–90. [Online]. Available: <https://doi.org/10.1109/ICICT50521.2020.00021>

BIOGRAPHIES OF AUTHORS



Ebeid Ali Ebeid     Ebeid Ali obtained his B.Sc. degree in computer science from department of mathematics, faculty of science, Al Azhar university, Cairo, Egypt in 2002 and M.Sc. degree from the same faculty in 2015. He was involved as a teaching assistant for the department of mathematics, faculty of science, Al Azhar university in 2012 then lecturer assistant in 2016. He received his PhD degree in 2020 and he is currently a lecturer of computer science at the same faculty. His research interests contain biometrics, pattern recognition, computer vision, machine learning and AI. He can be contacted at email: ebeid78@azhar.edu.eg



Ahmed Sobhi     received his B.Sc. degree in Mathematics and Computer Science from Department of Mathematics, Faculty of Science, Al Azhar University, Egypt in 2002. The M.Sc. degree in Computer Science from Faculty of Science, Minufiya University, Egypt in 2010. He obtained his PhD in Computer Science from Department of Mathematics, Faculty of Science, Al Azhar University, Egypt in 2017. He is a faculty member at Faculty of Science, Al Azhar University since 2012. His research includes IOT, Cloud Computing, Deep Learning, and Data Science. He can be contacted at email: sobhi205@azhar.edu.eg



H. A. El Shenbary     obtained his B.Sc. and M.Sc. degrees from Al-Azhar University, Cairo, Egypt. He holds a Ph.D. in Computer Science (2020) and is currently an Assistant Professor at Al-Azhar University. His research areas include Image Processing, Biometrics, Pattern Recognition, Classification, AI, Deep Learning, IoT, and Optimization Techniques. Contact: h.a.elshenbary@azhar.edu.eg