Investigating the effectiveness of deep learning approaches for deep fake detection

Article *in* Bulletin of Electrical Engineering and Informatics · December 2023

DOI: 10.11591/eei.v12i6.6221

Berrahal Mohammed
Université Mohammed Premier
10 PUBLICATIONS **85** CITATIONS
SEE PROFILE

Mohammed Boukabous
Université Mohammed Premier
17 PUBLICATIONS **211** CITATIONS
SEE PROFILE

Mimoun Yandouzi
Université Mohammed Premier
9 PUBLICATIONS **49** CITATIONS
SEE PROFILE

Mounir Grari
Université Mohammed Premier
11 PUBLICATIONS **62** CITATIONS
SEE PROFILE

Bulletin of Electrical Engineering and Informatics
Vol. 12, No. 6, December 2023, pp. 3853~3860
ISSN: 2302-9285, DOI: 10.11591/eei.v12i6.6221

Investigating the effectiveness of deep learning approaches for deep fake detection

Mohammed Berrahal[1], Mohammed Boukabous[1], Mimoun Yandouzi[2], Mounir Grari[1], Idriss Idrissi[1]
[1]Mathematics, Signal and Image Processing, and Computing Research Laboratory (MATSI), Higher School of Technology (ESTO), Mohammed First University, Oujda, Morocco
[2]Engineering Sciences Laboratory (LSI), National School of Applied Sciences (ENSAO), Mohammed First University, Oujda, Morocco

Article Info

Article history:

Received Mar 14, 2023
Revised May 25, 2023
Accepted Jun 5, 2023

Keywords:

Convolutional neural network
Deep learning
Fake face detection
Generative adversarial net
Transfer learning

ABSTRACT

As a result of notable progress in image processing and machine learning algorithms, generating, modifying, and manufacturing superior quality images has become less complicated. Nonetheless, malevolent individuals can exploit these tools to generate counterfeit images that seem genuine. Such fake images can be used to harm others, evade image detection algorithms, or deceive recognition classifiers. In this paper, we propose the implementation of the best-performing convolutional neural network (CNN) based classifier to distinguish between generated fake face images and real images. This paper aims to provide an in-depth discussion about the challenge of generated fake face image detection. We explain the different datasets and the various proposed deep learning models for fake face image detection. The models used were trained on a large dataset of real data from CelebA-HQ and fake data from a trained generative adversarial network (GAN) based generator. All testing models achieved high accuracy in detecting the fake images, especially residual neural network (ResNet50) which performed the best among with an accuracy of 99.43%.

Corresponding Author:

Mohammed Berrahal
Mathematics, Signal and Image Processing, and Computing Research Laboratory (MATSI)
Higher School of Technology (ESTO), Mohammed First University
Oujda, Morocco
Email: m.berrahal@ump.ac.ma

1. INTRODUCTION

Fake face detection is a critical challenge, given the potential harm it can cause in various domains such as social media, journalism, politics, and law enforcement [1], [2]. The use of counterfeit faces can result in identity theft, fraud, and other malicious activities that have the potential to cause harm to both individuals and organizations. Additionally, the proliferation of fabricated facial imagery can result in the dissemination of false news and misinformation, complicating the task of discerning genuine from counterfeit information. Researchers are exploring the efficacy of deep learning methods to identify fraudulent facial features in order to address this problem [3], [4].

By training on a large dataset of real and fake faces, deep learning techniques, including convolutional neural networks (CNNs) and generative adversarial networks (GANs) are capable of identifying subtle differences between the two types. These approaches can identify patterns and features that are difficult to be detected by the human eye, making it a promising solution for detecting fake faces as shown in Figure 1, we present samples of human faces from the CelebA-HQ dataset in Figure 1(a), and in Figure 1(b) we show generated images from a GAN-based model. It is evident that we face difficulty in distinguishing between the sets of images. However, there are several challenges in developing effective deep learning models for fake face detection. One of the major challenges is the lack of a large and diverse dataset of fake faces, which is essential for training and testing deep learning models [5]. Furthermore, the creation of fake faces is becoming increasingly sophisticated, developers are creating new methods to produce human faces that are highly realistic and fake, which can evade detection.

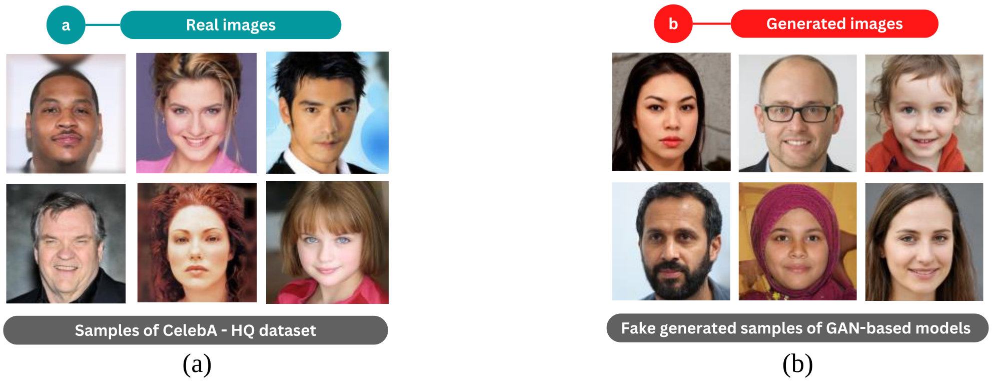

a
Real images
Samples of CelebA - HQ dataset
(a)
b
Generated images
Fake generated samples of GAN-based models
(b)

Figure 1. Samples of images used in our work: (a) sample from CelebA-HQ dataset and (b) generated sample using GAN-based trained model

Therefore, this research aims to address these challenges and investigate the effectiveness of different deep learning approaches for detecting fake faces. The study will involve exploring various deep learning architectures, training, testing datasets, and evaluating the performance of these models on different metrics. The aim is to discover the most efficient methodology for detecting counterfeit faces and to contribute to the advancement of dependable techniques that identify and prevent the dissemination of fraudulent images and videos. This study holds the potential to have a significant impact on impeding the proliferation of false faces and safeguarding the credibility of digital media.

2. BACKGROUND

2.1. Computer vision

The field of artificial intelligence known as computer vision concentrates on empowering computers to comprehend and interpret visual information from the world, similar to how humans perceive and process visual information [6]. The goal of computer vision is to enable machines to automatically analyze, process, and interpret images and videos, and to extract meaningful insights and information from them. This involves developing algorithms and techniques that can perform tasks such as object recognition, segmentation, tracking, and image restoration. Computer vision has numerous real-world applications, including autonomous vehicles, medical imaging, facial recognition, security and surveillance, and robotics [7]. With its ever-evolving refinement and precision, computer vision possesses the capability to catalyze a transformative shift across a multitude of industries and disciplines.

2.2. Generative adversarial network

GANs belong to the category of deep learning generative models, capable of generating new data by learning the underlying patterns in each dataset. A GAN comprises of two neural networks, namely a generator network and a discriminator network. The generator network learns to generate synthetic data that is similar to belong to the category of deep learning generative models, capable of the training data, while the discriminator network learns to distinguish between real and fake data. During training, the generator and discriminator networks compete against each other in a game-like setup, where the generator tries to generate more realistic data, while the discriminator tries to correctly identify real and fake data [8]. The training process of a GAN involves updating the generator and discriminator networks alternatively until the generator can produce synthetic data that is indistinguishable from the real data as shown in Figure 2. GANs have demonstrated their effectiveness in various applications, such as generating images and videos, style transfer, and data augmentation. One of the major advantages of GANs is that they can generate highly realistic and diverse data, which is not possible with traditional generative models. However, GANs can be challenging to train and require careful tuning of hyperparameters to achieve good results.

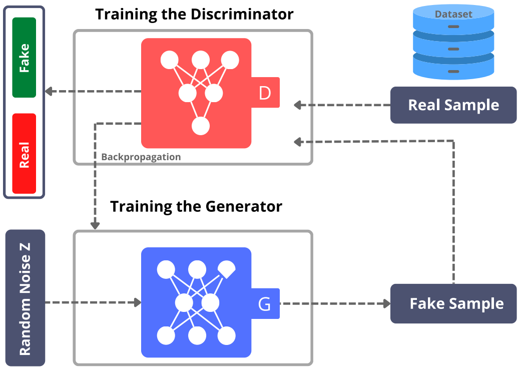


Figure 2. The architecture of GANs

2.3. Transfer learning

Transfer learning is a technique in machine learning that involves taking a model trained for one task and adapting it to another related task through fine-tuning [9]. In the context of classification, transfer learning involves using a pre-trained model as a starting point for training a new classifier. In transfer learning, the pre-trained model is typically a deep neural network that has been trained on a large dataset, such as ImageNet, which contains millions of labeled images. The pre-trained model learns a set of feature representations that can be useful for a wide range of computer vision tasks [10]. To use transfer learning for classification, the pre-trained model is first modified by removing the last layer, which is typically the output layer that predicts the class labels for the original dataset. A new output layer is then added to the model, which will be trained on the target dataset. During the fine-tuning stage, the weights of the pre-trained model are frozen, and only the weights of the new output layer are updated. This allows the model to quickly adapt to the new task, while still retaining the valuable feature representations learned during the pre-training stage. Transfer learning can significantly reduce the amount of labeled data required for training a new model and can also improve the accuracy of the classifier, especially in cases where the target dataset is small or similar to the original dataset used for pre-training. Table 1 features an exquisite compilation of transfer learning models that have been carefully selected for utilization in our paper.

Table 1. The best-performing architectures for image classification CNN-based model

Ref/year	Model name	Brief description
[11]/2015	ResNet50	Residual neural networks (ResNets) are a type of CNN that has proven to be very effective at reducing the training error of very deep networks. ResNet50 consists of 50 layers.
[12]/2014	VGG16 and VGG19	Both VGG16 and VGG19 are deep neural networks. VGG16 has 16 layers, consisting of 13 convolutional layers and 3 fully connected layers. On the other hand, VGG19 is comparable to VGG16 but has 19 layers.
[13]/2016	DenseNet	Uses dense blocks, which are blocks of layers where each layer is connected to every other layer in the block. Dense-Net has a total of 121 layers, including 100 convolutional layers.
[14]/2018	MobileNetV2	Developed specifically for use on mobile devices. MobileNetV2 has 19 layers.
[15]/2017	Xception	Is an architecture that uses depth-wise separable convolutions (a type of convolution that is very efficient computationally). It has a total of 71 layers, including 46 convolutional layers.
[16]/2017	NASNetMobile	Is an architecture that is specifically designed for use on mobile devices. It has a total of 474 layers, including 33 convolutional layers.
[17]/2015	InceptionV3	Is an architecture that is composed of a series of Inception modules. It has a total of 42 layers, including 31 convolutional layers.

2.4. Related work

There has been significant research in the field of generative fake face image detection in recent years. Many of these studies have focused on developing deep learning models that can accurately distinguish between real and fake face images. One popular approach is to use GANs [18], [19], which are generative models that can create highly realistic face images. To identify such counterfeit faces, researchers have devised techniques for training discriminative models that differentiate between authentic and fraudulent images. Zhang *et al.* [20] proposed a method for detecting GAN-generated fake face images based on the visual artifacts that are present in these images. In addition to these approaches, researchers have also explored the use of other machine learning techniques, such as support vector machines (SVMs) and k-nearest neighbor, for detecting fake faces [21]–[23]. In another hand, Li and Lyu [24] propose in their paper a novel deep learning framework for detecting face forgery in images, based on the idea of contrasting two hypotheses. Zakharov *et al.* [25] propose a method for detecting deep fake images by analysing inconsistencies in the head pose of the subject. Nguyen *et al.* [26] propose a method for detecting images generated by GANs for analysing co-occurrence matrices of pixel values. Rossler *et al.* [27] work provides a comprehensive assessment of the threat posed by deep fake images to face recognition systems and proposes a method for detecting them based on analysing the consistency of facial landmarks. Shu *et al.* [28] proposes a method for detecting anomalies in seasonal key performance indicators (KPIs) using a variational autoencoder, which could potentially be applied to detect deep fake images. These are just a few examples of the many works in the field of deep fake image detection. As the technology behind deep fake images continues to evolve, new detection techniques will likely need to be developed to keep up with the threat they pose. Overall, these studies demonstrate the potential of deep learning models for detecting fake face images and suggest that these models could play a significant role in addressing the problem of fake face generation. However, further research is needed to improve the accuracy of these models and to develop more robust and effective methods for detecting fake face images.

3. METHOD

3.1. Datasets

CelebA-HQ is a large-scale face dataset that contains high-quality images of celebrity faces. The CelebA dataset, consisting of images of celebrities with lower resolution, was developed by researchers from the Chinese University of Hong Kong and Tencent AI Lab. This new dataset is a continuation of CelebA. CelebA-HQ dataset contains 30,000 images with a resolution of 1,024x1,024 pixels, making it one of the highest-resolution face datasets available [29]. The flickr faces HQ (FFHQ) dataset is a high-quality dataset of human faces collected by Nvidia for training and evaluating generative models. The dataset consists of 70,000 high-resolution (1,024x1,024 pixels) images of faces, featuring a diverse range of ages, genders, and ethnicities [30].

3.2. Proposed method

In this particular study, we employed a unique method that involved the utilization of two distinct datasets, namely the CelebA-HQ and FFHQ datasets. The objective was to amalgamate these datasets, resulting in the creation of a third dataset that would serve as a foundation for our research. Through an intricate process, we harnessed the power of a style-based generator model and subjected it to extensive training using this newly formed dataset. The outcome was the generation of a remarkable collection of over 5,000 high-quality images, which were subsequently employed as synthetic label images, commonly referred to as “fake” images.

To establish a solid benchmark for evaluation, we introduced the CelebA-HQ dataset as a source of real image labels, as depicted in Figure 3. By combining both the 5,000 real images from the CelebA-HQ dataset and the 5,000 fake images generated by our style-based generator, we were able to execute comprehensive training and conduct a thorough assessment of our transfer learning models. This approach proved instrumental in enhancing the accuracy and reliability of our research outcomes.

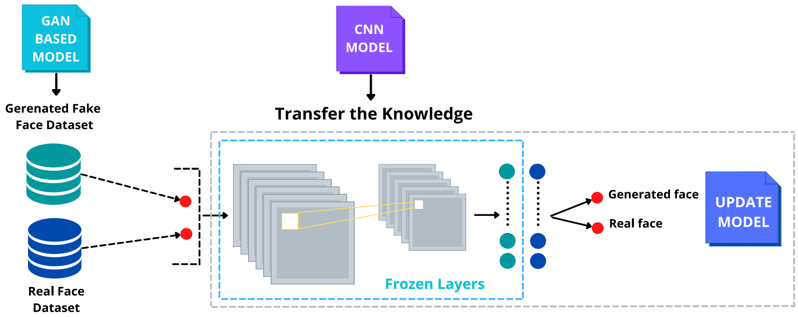


Figure 3. Schema of the method used in this paper

3.3. Evaluation metrics

For the study of deep learning models, four quantitative measures were used to evaluate the performance and efficacy of every model defined by the following functions: accuracy, precision, recall, and F1-score as shown in Figure 4. These metrics are calculated using the following definition:

- True positive (TP): fake images were correctly classified as generated images
- False negative (FN): fake images were incorrectly classified as real images
- True negative (TN): real images were correctly classified as real images
- False positive (FP): real images were incorrectly classified as generated images

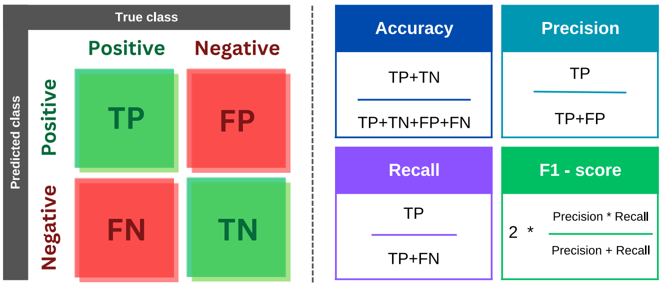


Figure 4. Evaluation metrics used in our study

3.4. Hardware characteristics

For the test of our models, we use the high-performance computing (HPC) infrastructure cluster HPC-MARWAN, with the following characteristics:

- Compute nodes: 2*Intel Xeon Gold 6148 (2.4 GHz/20-core)/192 GB RAM.
- GPU node: 2*NVIDIA P100/192 GB RAM.
- Storage node: 2*Intel Xeon Silver 4114 (2.2 GHz/20-core)/18 * SATA 6 TB.

4. RESULTS AND DISCUSSION

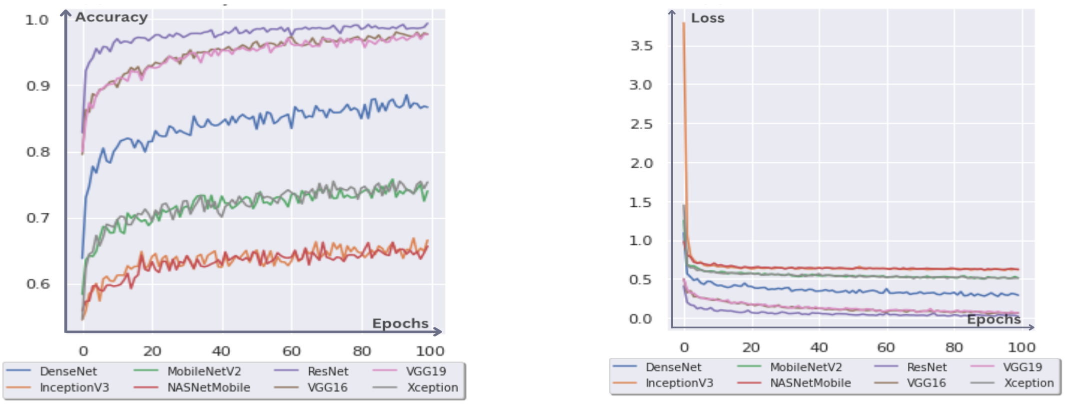
Table 2 presents the results of using various deep learning algorithms for the classification of human face images as real or fake. The performance of each algorithm is evaluated based on several metrics, including accuracy, loss, precision, recall, and F1-score. Based on the results, ResNet50 achieved the highest accuracy of 99.34%, followed by VGG16, VGG19, and DenseNet with accuracies of 97.71%, 97.65%, and 86.65% respectively. The MobileNetV2 and Xception algorithms performed relatively worse with accuracies of 74.00% and 75.34%, while NASNetMobile and InceptionV3 achieved the lowest accuracy scores of 65.65% and 66.56% respectively.

Table 2. Attained results for the CNN models

Deep learning algorithms	Number of parameters	Accuracy (%)	Loss (%)	Precision (%)	Recall (%)	F1-score (%)
ResNet50	24,637,313	99.34	2.92	99.34	99.40	99.37
VGG16	14,977,857	97.71	6.32	97.59	98.00	97.79
VGG19	20,287,553	97.65	6.64	97.60	97.89	97.74
DenseNet	7,562,817	86.65	29.20	86.73	87.62	87.17
MobileNetV2	2,914,369	74.00	51.28	74.11	77.55	75.79
Xception	21,911,081	75.34	51.21	75.51	77.56	76.52
NASNetMobile	4,811,413	65.65	62.52	66.86	67.22	67.04
InceptionV3	22,852,385	66.56	61.73	68.86	71.12	69.97

Moreover, the precision metric, which measures the proportion of true positives among all predicted positives, was highest for ResNet50 with a score of 99.34%, followed by VGG16 and VGG19. The recall metric, which measures the proportion of true positives among all actual positives, was highest for ResNet50 VGG16, VGG19 with scores of 99.40%, 97.71%, and 97.65% respectively. The F1-score, which is a harmonic mean of precision and recall, provides an overall evaluation of the algorithm's performance. ResNet50 achieved the highest F1-score of 99.37%, followed by VGG16 and VGG19 with scores of 97.79% and 97.74% respectively. The remaining algorithms achieved F1 scores below 90%.

Overall, the results certifies that ResNet50 performed the best among the tested algorithms, with the highest accuracy, precision, recall, and F1 score. However, the choice of the most appropriate algorithm for a specific task should be based on various factors, including the size and complexity of the dataset, the required level of accuracy, and the available computational resources. The training process was smooth no trace of overfitting or underfitting, as shown in Figure 5, we demonstrate the training process of our models, where Figure 5(a) represents the accuracy and Figure 5(b) shows the loss.



(a) (b)

Figure 5. Training process of transfer learning models: (a) the accuracy of trained models and (b) the loss of trained models

We perform several tests using ResNet50 to predict images outside datasets as shown in Figure 6. Most of the images were predict correctly, however a minority of images failed the classification process, due the complexity, some images may be too complex for the classification model to accurately determine whether they are real or fake. If an image contains excessive visual noise or significant manipulation, it could pose a challenge for the model to accurately classify it. Additionally, deep learning models are limited in their capabilities and may encounter difficulties in accurately classifying images that do not align with their strengths.

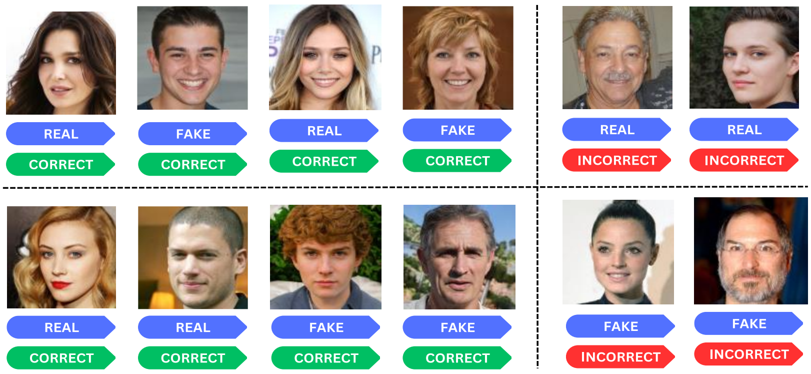


Figure 6. Samples of predicted fake and real images

5. CONCLUSION

In conclusion, the ease of creating and manipulating high-quality images using machine learning algorithms has opened new possibilities for image-based applications. However, it has also resulted in the production of counterfeit images that have the potential to mislead people. The paper presents an in-depth analysis of the challenge of detecting fake face images and proposes the implementation of CNN-based classifiers to address this issue. Some proposed model achieved high accuracy in distinguishing between real and fake images, with ResNet50 being the best-performing algorithm reaching the highest accuracy score of 99.43. furthermore, we provide a valuable contribution towards combating image manipulation and highlight the need for continued research in this area to prevent the malicious use of image manipulation techniques.

ACKNOWLEDGEMENTS

Research grant: this research was supported through computational resources of HPC-MARWAN (www.marwan.ma/hpc) provided by the National Center for Scientific and Technical Research (CNRST), Rabat, Morocco – Phone number (+212537569877).

REFERENCES

[1] M. Berrahal and M. Azizi, “Review of DL-Based Generation Techniques of Augmented Images using Portraits Specification,” in *2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS)*, IEEE, Oct. 2020, pp. 1–8. doi: 10.1109/ICDS50568.2020.9268710.

[2] M. Berrahal and M. Azizi, “Augmented binary multi-labeled CNN for practical facial attribute classification,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 2, pp. 973–979, Aug. 2021, doi: 10.11591/ijeecs.v23.i2.pp973-979.

[3] S. Aurellia and S. F. Rahman, “Face recognition in identifying genetic diseases: a progress review,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 3, pp. 1019–1025, Sep. 2023, doi: 10.11591/ijai.v12.i3.pp1019-1025.

[4] I. Intan, Nurdin, and F. Pangerang, “Facial recognition using multi edge detection and distance measure,” *IAES International Journal of Artificial Intelligence*, vol. 12, no. 3, pp. 1330–1342, 2023, doi: 10.11591/ijai.v12.i3.pp1330-1342.

[5] M. Grari *et al.*, “Using Iot and Ml for Forest Fire Detection, Monitoring, and Prediction: a Literature Review,” *Journal of Theoretical and Applied Information Technology*, vol. 100, no. 19, pp. 5445–5461, 2022.

[6] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, “Deep Learning for Computer Vision: A Brief Review,” *Computational Intelligence and Neuroscience*, pp. 1–13, 2018, doi: 10.1155/2018/7068349.

[7] A. Kherraki, M. Maqbool, and R. El Ouazzani, “Traffic Scene Semantic Segmentation by Using Several Deep Convolutional Neural Networks,” in *2021 3rd IEEE Middle East and North Africa COMMunications Conference, MENACOMM 2021*, 2021, pp. 1–6, doi: 10.1109/MENACOMM50742.2021.9678270.

[8] I. Goodfellow, J. Pouget-Abadie, B. X. Mehdi Mirza, D. Warde-Farley, A. C. Sherjil Ozair, and Y. Bengio, “GAN（Generative Adversarial Nets,” *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, vol. 29, no. 5, p. 177, Oct. 2017, doi: 10.3156/jsoft.29.5_177_2.

[9] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *Journal of Big Data*, vol. 3, no. 1, pp. 1–41, Dec. 2016, doi: 10.1186/s40537-016-0043-6.

[10] N. Rachburee and W. Punlumjeak, “Lotus species classification using transfer learning based on VGG16, ResNet152V2, and MobileNetV2,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 4, pp. 1344–1352, Dec. 2022, doi: 10.11591/ijai.v11.i4.pp1344-1352.

[11] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[12] K. Simonyan and Z. Andrew, “Very deep convolutional networks for large-scale image recognition,” *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Sep. 2015, pp. 1–14, 2015.

[13] G. Huang, Z. Liu, and L. van der Maaten, “Densely Connected Convolutional Networks,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1978, vol. 39, no. 9, pp. 1442–1446.

[14] A. G. Howard *et al.*, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” *Computer Vision and Pattern Recognition*, vol. 14, no. 2, pp. 53–57, 2017.

[15] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 1800-1807, doi: 10.1109/CVPR.2017.195.

[16] B. Zoph and Q. V. Le, “Neural architecture search with reinforcement learning,” *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, 2017, pp. 1-16.

[17] C. Szegedy *et al.*, “Going deeper with convolutions,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

[18] M. Berrahal and M. Azizi, “Optimal text-to-image synthesis model for generating portrait images using generative adversarial network techniques,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 2, pp. 972-979, Feb. 2022, doi: 10.11591/ijeecs.v25.i2.pp972-979.

[19] M. Berrahal and M. Azizi, “Improvement of facial attributes’ estimation using Transfer Learning,” in *2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, IEEE, Mar. 2022, pp. 1–7 doi: 10.1109/IRASET52964.2022.9737845.

[20] X. Zhang, S. Karaman, and S. F. Chang, “Detecting and Simulating Artifacts in GAN Fake Images,” *2019 IEEE International Workshop on Information Forensics and Security, WIFS 2019*, 2019, pp. 1-6, doi: 10.1109/WIFS47025.2019.9035107.

[21] P. He, H. Li, and H. Wang, “Detection of Fake Images Via The Ensemble of Deep Representations from Multi Color Spaces,” in *2019 IEEE International Conference on Image Processing (ICIP)*, IEEE, Sep. 2019, pp. 2299–2303. doi: 10.1109/ICIP.2019.8803740.

[22] A. H. Ali, M. A. Mohammed, R. A. Hasan, M. N. Abbod, M. Sh. Ahmed, and T. Sutikno, “Big data classification based on improved parallel k-nearest neighbor,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 21, no. 1, pp. 235–246, Feb. 2023, doi: 10.12928/telkomnika.v21i1.24290.

[23] F. F. Kharbat, T. Elamsy, A. Mahmoud, and R. Abdullah, “Image Feature Detectors for Deepfake Video Detection,” in *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*, IEEE, Nov. 2019, pp. 1–4. doi: 10.1109/AICCSA47632.2019.9035360.

[24] Y. Li and S. Lyu, “Exposing DeepFake Videos By Detecting Face Warping Artifacts,” 2018.

[25] E. Zakharov, A. Shysheya, E. Burkov, and V. Lempitsky, “Few-shot adversarial learning of realistic neural talking head models,” *Proceedings of the IEEE International Conference on Computer Vision*, Oct 2019, pp. 9458–9467, doi: 10.1109/ICCV.2019.00955.

[26] T. T. Nguyen *et al.*, “Deep learning for deepfakes creation and detection: A survey,” *Computer Vision and Image Understanding*, vol. 223, pp. 1–20, Oct. 2022, doi: 10.1016/j.cviu.2022.103525.

[27] A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Niessner, “FaceForensics++: Learning to detect manipulated facial images,” *Proceedings of the IEEE International Conference on Computer Vision*, Oct 2019, pp. 1–11, doi: 10.1109/ICCV.2019.00009.

[28] Y. Shu, T. Gao, Z. Zhang, and J. Zhang, “A General KPI Anomaly Detection Using Attention Models,” in *2022 IEEE International Conference on Services Computing (SCC)*, IEEE, Jul. 2022, pp. 114–119. doi: 10.1109/SCC55611.2022.00027.

[29] C.-H. Lee, Z. Liu, L. Wu, and P. Luo, “CelebAMask-HQ Dataset,” *MMlab*, 2020. [Online]. Available: https://mmlab.ie.cuhk.edu.hk/projects/CelebA/CelebAMask_HQ.html. Access date: May 09, 2022.

[30] Karras *et al.*, “FFHQ (Flickr-Faces-HQ),” *Paperswithcode*. [Online]. Available: https://paperswithcode.com/dataset/ffhq. Access date: May 06, 2023.

BIOGRAPHIES OF AUTHORS

Mohammed Berrahal has a Ph.D. in computer Science at Mohammed First University in Oujda, Morocco, where he is conducting research on security and law enforcement applications utilizing deep learning. He holds an M.Sc. in internet of things from National School of Computer Science and Systems Analysis (ENSIAS), Mohammed 5 University in Rabat, Morocco (2018) and a B.Sc. in computer engineering from ESTO, Mohammed First University in Oujda, Morocco (2016). Furthermore, he is certified in artificial intelligence, 3D modeling, and programming. Additionally, he has served as a reviewer for a number of international conferences and journals. He is currently employed at Mohammed First University as an administrative assistant. He can be contacted at email: m.berrahal@ump.ac.ma.

Mohammed Boukabous has a Ph.D. in Computer Science at Mohammed First University in Oujda, Morocco, where he is conducting research in security intelligence using deep learning algorithms in exchanged messages. He holds a M.Sc. degree in internet of things from Sidi Mohamed Ben Abdellah University in Fez, Morocco (2019), as well as a B.Sc. degree in Computer Engineering from Mohammed First University (2016). Furthermore, he holds several certifications in natural language processing, artificial intelligence, security intelligence, big data, and cybersecurity. Additionally, he served as a reviewer for various international conferences. He is currently employed at Mohammed First University as an administrative. He can be contacted at email: m.boukabous@ump.ac.ma.

Mimoun Yandouzi received Ph.D. in Computer Science at Mohammed First University in Oujda, Morocco, where he is conducting research on the use of computer vision and deep learning techniques for the analysis of drone data, particularly in the case of forest fire detection. He holds a degree in Computer Engineering from the School of Mineral Industry in Rabat, Morocco (2001). Furthermore, he holds several certifications in artificial intelligence, computer vision, cloud computing, big data, and data mining. He also acted as a reviewer for several international conferences. He is currently employed at Mohammed First University as a professor at the ENSA Engineering School. He can be contacted at email: m.yandouzi@ump.ac.ma.

Mounir Grari has a Ph.D. in Computer Engineering at Mohammed First University in Oujda, Morocco, where he is conducting research on the use of the internet of things and machine learning in the detection and monitoring of forest fires. He holds an engineering degree in Computer Science from EMI, University Mohammed 5 in Rabat, Morocco (2002). Furthermore, he is certified in artificial intelligence, 3D modeling, and programming. Additionally, he has served as a reviewer for a number of international conferences and journals. He is currently employed at Mohammed First University as secretary general of the College of Technology. He can be contacted at email: m.graril@ump.ac.ma.

Idriss Idrissi is a professor at the Higher School of Technology (ESTO) of Mohammed First University, Oujda, Morocco, where he is researching internet of things security using deep learning. He has an M.Sc. degree in internet of things from Sidi Mohamed Ben Abdellah University in Fez, Morocco (2019), a B.Sc. degree in Computer Engineering from Mohammed First University (2016). Additionally, he holds several certifications in networking, artificial intelligence, cybersecurity, and programming. Also, he was a reviewer for various international conferences and journals. He can be contacted at email: idrissi@ump.ac.ma.