BİTLİS EREN ÜNİVERSİTESİ
FEN BİLİMLERİ DERGİSİ

ISSN: 2147-3129 / e-ISSN: 2147-3188

Article Type	: Research Article	**Year**	: 2025
Received	: July 26, 2024	**Volume**	: 14
Revised	: January 6, 2025	**Issue**	: 1
Accepted	: January 24, 2025	**Pages**	: 546-560
DOI	: 10.17798/bitlisfen.1610300		

DEEPFAKE IMAGE DETECTION WITH TRANSFER LEARNING MODELS

Lutfi Emre DEMİR [1], **Yavuz CANBAY [2,*]**

[1]*Kahramanmaraş Sütçü İmam University, Department of Information Systems, Kahramanmaraş, Türkiye*

[2]*Kahramanmaraş Sütçü İmam University, Department of Computer Engineering, Kahramanmaraş, Türkiye*

**** Corresponding Author:*** *yavuzcanbay23@gmail.com*

ABSTRACT

Deepfake is a technology that employs artificial intelligence to generate fake images and videos that closely mimic real ones. The rapid growth and dissemination of digital modifications generate considerable concern in the media, politics, and social networking. Among the concerns faced include the dissemination of misinformation, reputational damage, and threats to physical security. Given these concerns, prompt and reliable identification of Deepfakes is crucial for protecting information security and mitigating its negative impacts. Therefore, the development of effective technologies for Deepfake detection is essential to counter this increasing threat. This study aims to identify Deepfake images and examine the efficiency of transfer learning algorithms in Deepfake identification. This study employed the most commonly utilized transfer learning models, including InceptionV3, EfficientNet, NASNet, ResNet, DenseNet, Xception and ConvNeXt, to perform the detection task. An extensive public dataset of 190,000 images, including both real and artificially generated, was utilized in the study. The performance of each model was assessed by using the metrics of accuracy, precision, recall, and F1-score. DenseNet was the most successful model with a test accuracy of 93%. The obtained results have shown that transfer learning models can effectively detect the Deepfake images, providing a practical approach to the challenge with reasonable performance scores.

Keywords: Deepfake detection, Information security, Transfer learning

1 INTRODUCTION

Advancements in artificial intelligence and deep learning have made Deepfake technology possible, transforming the production of realistic synthetic media content. Deepfake is a combination of "deep learning" and "fake," referring to the technology that processes visual and auditory data to generate impressively realistic imitation images and movies. Although Deepfake offers benefits in diverse domains like entertainment, advertising, and education, they also present significant risks. The utilization of Deepfakes, especially for the malicious aims of misinformation, privacy violation, cyberbullying, and fraud, has raised significant ethical and security issues. Once shared over social media, this type of content can be utilized to influence the community, damage faith in digital interactions, incur economic consequences, and potentially increase social unrest [1].

The rapid proliferation and growing complexity of Deepfake media necessitate the urgent development of reliable detection methods. The consequences of undercover and undetectable Deepfakes are severe, involving electoral manipulation, financial deception, and societal unrest. Consequently, there is an urgent demand for innovative technology capable of consistently and effectively identifying "fake" materials.

In the literature, there exist some studies considering detection of Deepfake media. For instance, Afcharve et al. [2] presents a deep learning method for detecting face forgeries produced by hyper-realistic fake video generation techniques such as Deepfake and Face2Face. For this purpose, two different low-layer and fast networks capable of extracting mesoscopic or medium-scale features of images were designed. The proposed networks provided successful results even in video formats where traditional image analysis techniques are inadequate due to the information degradation caused by data compression. The method was tested on an existing dataset and a custom dataset compiled from online videos. In the tests, over 98% detection success was achieved for Deepfake videos and over 95% for Face2Face videos. The CNN-based model used in the study achieved 98.40% accuracy on the FaceForensics++ dataset.

Prajapati et al. [3] aimed to develop deep learning-based generic models to detect Deepfakes, which are synthetic videos created by replacing the face in an original image with the face of another person. The paper proposed a new Generative Adversarial Network (GAN) based model, called MRI-GAN, which detects fake videos using perceptual differences in images. The proposed MRI-GAN model was tested on the Deepfake Detection Challenge dataset while evaluating perceptual differences with structural similarity index. Experimental results showed that the model based on flat frames achieves 91% test accuracy, while the MRI-GAN framework achieves 74% accuracy.

Miao et al. [4] presented a new method that improves the generalization ability and robustness in the field of face forgery detection to prevent the malicious use of face manipulation technology. The proposed model encodes the relationships between patches by extending the Transformer structure with a bag-of-local-feature approach, which allows the model to learn local forgery features without any explicit supervision. The effectiveness of the method had been extensively tested on the FaceForensics++, Celeb-DF and DeeperForensics-1.0 datasets. According to the test results, 87.86% accuracy was achieved on FaceForensics++ dataset, 82.52% AUC on Celeb-DF dataset and 97.01% accuracy on DeeperForensics-1.0 dataset.

Deepfake detection was considered as a fine discrimination-oriented classification problem instead of traditional binary classification to capture fine details in [5]. The proposed multi-attention network-based model combined both low-level textural and high-level semantic features using attentional mechanisms to highlight forgery traces. In order to improve the learning process of the model, regional loss of independence and an attention-guided data augmentation strategy were also added. In the test studies on various datasets, it outperformed other models with 97.60% accuracy on the FaceForensics++ dataset and 0.1679 loss value on the DFDC dataset.

Since there is a challenge between different types of GANs for detecting deepfake images, Kanwal et al. [6] proposed a general solutions. In their work, they aimed to detect deepfake images using Siamese Networks with triplet loss function. The experiments were performed in two different cases. In the first case, the training and test sets were chosen from the same dataset, which consisted of real images from the FFHQ dataset and fake images generated by StyleGAN. In the second case, the training and test sets were chosen from different datasets, in which case the model was trained on the FFHQ dataset and StyleGAN and tested on images generated by PGAN. The results showed that using contrast loss or triplet loss instead of cross-entropy or MSE improves the generalization ability of the model. An accuracy of 94.80% was achieved in the study

In a study by Rafique et al. [7], fake face detection was addressed using two machine learning algorithms based on features extracted by AlexNet and ShuffleNet models. A new image descriptor is also developed to improve the predictive power of the proposed network. The authors claimed that there are differences in compression levels between the original and fake images. In this context, the proposed approach evaluates the difference between the original image and its versions with 85% compression. This method is called Error Level

Analysis (ELA), which reveals the lost details due to the compression level. The images obtained by ELA are given to AlexNet and ShuffleNet models for feature extraction. The generated feature vector was passed to SVM and k-NN classifiers for classification as real or fake. In the experiments, on the Real and Fake Face Detection dataset, ShuffleNet achieved the highest accuracy with 88.2% when combined with k-NN classifier and 87.9% when combining with SVM classifier.

In another study, Nida et al. [8] aimed to detect real and fake face images using transfer learning. Since it is difficult to detect forgery by visual inspection alone, this research uses the "Real and Fake Face Detection" dataset created by the Computational Intelligence Photography Lab, in Yonsei University. In the proposed method, images are normalized as a first step and then preprocessed with Error Level Analysis. These normalized images are trained with various pre-trained deep learning models to classify fake and real faces. The VGG models showed the highest performance with fewer epochs than the other techniques; VGG-16 achieved 91.97% training accuracy and VGG-19 achieved 92.09% training accuracy. Model performances were evaluated by comparing with confusion matrix and existing methods.

Patel et al. [9] proposed an end-to-end method that combines features extracted by various CNN models to detect deepfake videos at frame level. Using the DFDC dataset, frames in videos are processed as individual images and features are extracted from these images. These features were then used for deepfake detection with a Random Forest classifier. Thanks to the features extracted with MobileNet CNN, 90.2% accuracy was achieved.

Joshi et al. [10] used the Xception model for the detection of deepfake images and videos. They used a dataset obtained from Kaggle, which contains 90,000 images of real and deepfake content. The model was trained by fine-tuned transfer learning. Data replication and regularization techniques were applied to increase the robustness of the model. With 93.01% accuracy on the test dataset, this method provided an effective solution against deepfake content spread on the internet and social media. The results showed that the Xception model provides a reliable deepfake detection method thanks to its strong classification capabilities.

Liao el al. [11] proposed a Transformer network that uses the self-attention mechanism to model long-range dependencies and global reception domains. A dual-branch feature extraction framework was developed and the extracted features are combined with transformer's encoder structure and the cross-attention mechanism to model them efficiently. The proposed method achieved 83.5%, 70.25% and 78.5% accuracy on Deepfake, Face2Face and NeuralTextures datasets respectively.

Table 1. Comparison of the studies in the literature

Study	Method	Success Rate
[2]	MesoNet	98.40%
[3]	MRI-GAN + SSIM and Simple Frame Model	91% (Simple Frame Model), 74% (MRI-GAN)
[4]	Transformer + Bag-of-Local-Feature	87.86% (FaceForensics++) 82.52% (Celeb-DF), 97.01% (DeeperForensics)
[5]	Multi-attentional Network	97.60%
[6]	Siamese Network	94.80%
[7]	AlexNet, ShuffleNet, ELA, k-NN, SVM	88.2% (ShuffleNet + k-NN) 87.9% (ShuffleNet + SVM)
[8]	VGG16 and VGG19	91.97% VGG16 92.02% VGG19
[9]	Frame-Based MobileNet + Random Forest Classifier	90.2%
[10]	Xception	93.01%
[11]	Transformer	83.5% (Deepfake) 70.25% (Face2Face) 78.5% (NeuralTextures)

The rest of this paper is organized as follows. The Materials and Methods section details the dataset, the computational environment and the transfer learning models employed in this paper. In particular, it also explains the performance metrics used for evaluation indicators. Experimental studies give the details about the experiments and provides a comparative analysis of the model performance in terms of the metrics of accuracy, precision, recall, and F1-score. It also discusses the strengths and weaknesses of each model, and demonstrate that DenseNet performs best among others in the detection of deepfake. Finally, the Conclusion summarizes the achievement of the study's results, presenting the effectiveness of DenseNet for such a real-world application and pointing to the directions of future work, i.e., the use of ensemble method, and video-oriented deepfake detection.

2 MATERIAL AND METHODS

In this section, the materials and methods used in the study are explained.

2.1 Dataset Description

This work used a dataset of 190,000 labeled images obtained from the reference [12], which includes real and GAN-generated deepfake images. The dataset used is a modified and enhanced version of the OpenForensics dataset, utilizing various manipulation methods such as face swapping and reenactment. Each image was created in 256 x 256 JPG format. Some samples from real and fake image are shown in Figure1.

Figure 1. Real and fake image samples.

2.2 Transfer Learning Models

Transfer learning has become a promising method in machine learning, especially when training samples are small or computational power is limited. It is possible to repurpose a model pre-trained on a huge dataset (e.g., ImageNet) for a given task by means of transfer learning, which not only does this save the computational cost, but also to exploit the knowledge contained within pre-trained features. In the following sections, seven leading transfer learning model were briefly explained.

2.2.1 InceptionV3

InceptionV3 stands out among deep learning architectures with its multi-scale feature extraction capability. Thanks to Inception modules, it learns both fine details and general features by using kernel filters of different sizes simultaneously. This is particularly advantageous for detecting manipulations in fake content. Auxiliary classifiers minimize the problem of damped gradients and provide more stable learning during model training. InceptionV3 offers a strong balance between accuracy and computational efficiency when working on large data sets, making it ideal for complex tasks such as deep fake detection [13]. InceptionV3 has the ability to learn features at different scales simultaneously, which is an important factor in deepfake detection. In this study, the purpose of using this model is that it can detect both the overall structure and small anomalies in face manipulations. Moreover, the ability to capture distortions at different frequencies in forged images with versatile convolution kernels makes the model a powerful option for deepfake analysis.

2.2.2 EfficientNet

EfficientNet is a model specifically designed for processing large datasets, and the EfficientNet variant used in this study is pre-trained on ImageNet, which is characterized by its computational efficiency in resource-limited environments [14]. EfficientNet is a model that can achieve high accuracy with less computational power thanks to its optimized balance of depth, width and resolution. It is preferred for large datasets used in deepfake image detection due to its fast-training process and low parameter requirement. It is also good at capturing small and subtle changes in manipulated facial images.

2.2.3 ResNet

ResNet achieves training of deep multilayer networks with damped gradient with the help of residual connections. Through these connections, the gradients can be returned more efficiently and models can be trained on deeper layers. ResNet architecture with layered scheme is capable of visualizing the complex visual damages in fake images. The ResNet model employed in the work performed well, maintaining an optimal trade-off between depth and computational cost. ResNet was tested on particularly large and diverse datasets [15]. ResNet offers high generalization capacity by learning from deep layers. In particular, it produces more stable results by minimizing feature loss in the detection of images containing a wide variety of manipulation techniques such as deepfakes. In this study, this model was employed since it can analyze fake images successfully by supporting multi-layer information flow.

2.2.4 DenseNet

DenseNet maximizes information flow by connecting the output of each layer to all subsequent layers. These dense connections allow the model to both avoid gradient loss and learn more effectively with fewer parameters. This feature of DenseNet yields powerful results, especially for large and diverse data sets [16]. Thanks to the strong information sharing between layers, DenseNet can learn fine details in fake images more effectively. Its high accuracy in transfer learning provides an advantage in a task where fine details are critical, such as deepfake detection. In particular, we preferred to use this model since it is effective in determining the differences between real and fake faces by learning small manipulations better.

2.2.5 Xception

Xception is computationally efficient, since it uses less parameters than conventional convolutions. It uses depthwise separable convolutions. In this model, it is possible to learn channel-based as well as spatial features in separate steps, and thus the model is a versatile tool for the identification of low-level as well as high-level manipulations. By its optimized design, Xception can achieve accuracy on difficult problems, including deep fake detection [17]. Xception is a particularly computationally efficient model. Its ability to achieve high accuracy using fewer parameters in the transfer learning process is a critical advantage in the classification of deepfake images. In this study, it is especially preferred because of its ability to detect facial manipulations and fine details. Its ability to better analyze deep forgery traces used in forged content increases its effectiveness in deepfake detection.

2.2.6 NasNet

The main benefits of NasNet include the possibility of finding an automated architecture that is best suited to a given task, along with high resistance to different manipulations. The purpose of this model is to find the best architecture for individual tasks without direct human supervision. Nevertheless, the model applied in the study did not achieve a good performance because other models performed much better. This indicates that the model needs to be fine-tuned for certain tasks e.g., detection of deep fakes [18]. In this paper, since the deepfake problem involves complex and dynamic data, NasNet was used for better performance thanks to its large model search space. Another advantage is that it reduces manual hyperparameter tuning in transfer learning processes and can be easily adapted to different datasets.

2.2.7 ConvNeXt

ConvNeXt integrates some architectural developments proposed in recent years as image transformers into a ResNet-type hierarchical backbone of low, medium and high levels, instead of a traditional CNN-conceptual model. The large kernel size and the application of Layer Normalization help this approach to generalize well in terms of both computational and representation learning performance. It serves as a backbone for computer vision tasks as we are able to efficiently capture local and global spatial features in the model. With an optimized structure, ConvNeXt achieves high performance in challenging areas such as object detection and image classification [19]. In this study, ConvNeXt, one of the latest versions of modern CNN architectures, was chosen for its fast adaptation in transfer learning processes, deep feature learning capability and superior performance on large datasets. The ability to capture subtle textural differences in deepfake images gives this model an advantage in detection processes.

2.3 Model Implementation

This study employed transfer learning with fine-tuning to optimize efficiency and accuracy. Rather than training models from scratch, we utilized pre-trained deep networks to preserve general feature representations while fine-tuning the deeper layers for enhanced adaption to deepfake-specific characteristics. This methodology enabled us to decrease computing cost, prevent overfitting, and enhance classification precision on our dataset. Our findings validate its efficiency, with DenseNet having 93% accuracy, illustrating the benefits of integrating transfer learning with fine-tuning for deepfake detection. While there exist many variations of CNNs, we preferred to use mostly used transfer learning techniques such as InceptionV3, EfficientNet, NasNet, ResNet, DenseNet, Xception and ConvNeXt in our study.

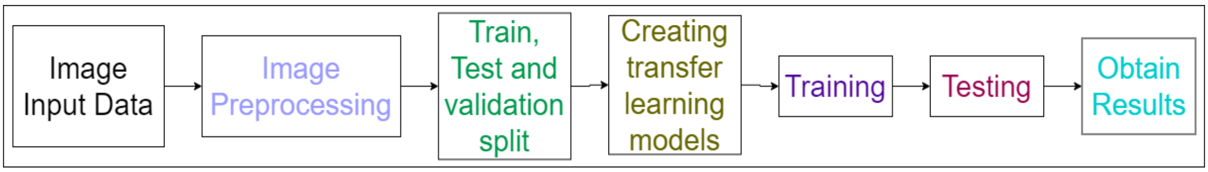

Figure 2. The workflow of the model implementation.

Figure 2 presents the workflow of the experimental implementations. In the workflow of the model implementation, firstly, the image dataset was loaded for the processing purpose. Every image was labelled as "fake" or "real" and the dataset was structured taking these labels into account. Data loaded was formatted according to the desired format for each model layer. Then in preprocessing, the size of each image was standardized to 224x224. The aim of resizing image is making the model run fast, stable and effectively. The dataset was divided into 75% training, 20% validation and 5% test. During this process, the balanced structure of the dataset, with 50% fake and 50% real images, was maintained, meaning the number of images in each class was kept equal. Subsequently, transfer learning models were developed and fine-tuned using the training dataset, followed by validation with the validation dataset. The models' performances through the trained model were evaluated using the test dataset. Accuracy, precision, recall and F1-score were used as performance indicators. Those metrics allowed to account for the goodness of the model and its performance by class. The results were then presented in tables and confusion matrix in the following subsections.

Sigmoid activation function was used due to its application in binary classification and that the output of the network gets transformed into a value between 0 and 1. In addition, the Adam optimization algorithm was used for weight optimization of the learned model. Binary Cross-Entropy was chosen as loss function as it is normally applied in the case of binary classification problems. Each model was trained during the learning phase for 10 epochs. Batch size was determined as 32. The learning rate was left at its default value.

3 EXPERIMENTAL RESULTS AND DISCUSSION

Accurate detection and classification of Deepfake images is crucial. In this section, the performances of seven transfer learning methods were obtained and then compared. A 190,000 labelled image dataset was used in the experiments. Accuracy, precision, recall and F1-Score metrics were employed for evaluations [20-23].

According to the obtained results, DenseNet achieved the highest test accuracy demonstrating its ability to accurately differentiate real and manipulated content. InceptionV3 and Xception followed closely with test accuracies of 92.60% and 92.15%, respectively. The performance metrics obtained in the experimental studies are presented in Table 2.

Table 2. Performance results of the transfer learning algorithms.

Model	**Train Accuracy (%)**	**Test Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
InceptionV3	99.17	92.60	93	92	92
EfficientNet	99.31	91.67	92	91	91
ResNet	99.12	90.54	91	90	90
DenseNet	98.90	93.00	93	93	93
Xception	99.33	92.15	92	92	92
NasNet	99.30	86.16	88	86	86
ConvNeXt	99.33	91.36	92	91	91

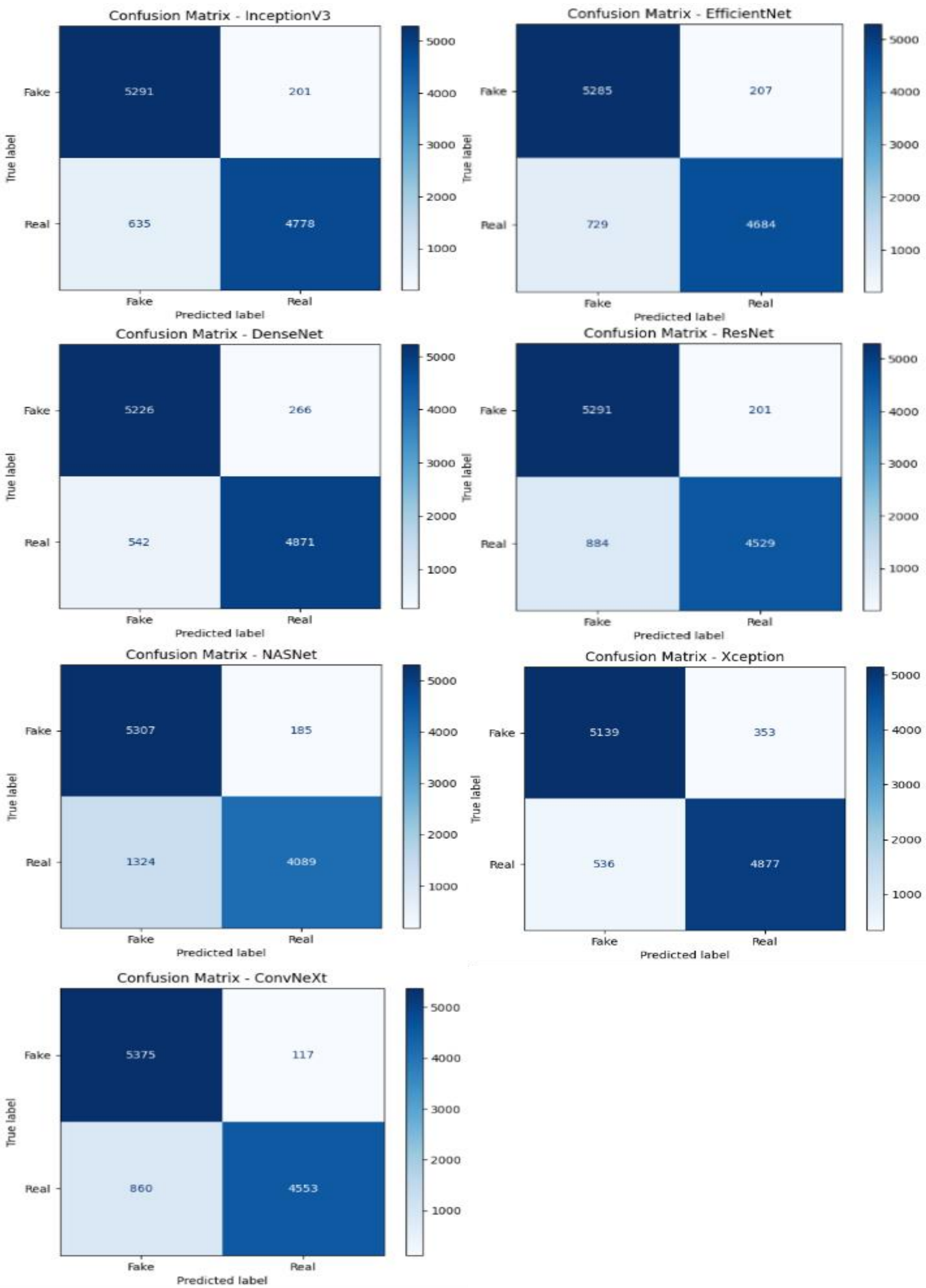

Figure 3. The confusion matrices of the transfer learning models.

For a more detailed analysis of model performances, the confusion matrix of all the models were given in Figure 3. These matrices summarize true positives, false positives, true negatives, and false negatives, providing a more nuanced picture of the classification behavior of each of the models. As can be seen from Table 2, the best test accuracy is obtained by DenseNet across all the categories with scarce misclassifications. Conversely, NasNet provides generally worse detection test results.

According to the results provided in Figure 3, InceptionV3 presented 635 FP and 201 FN, EfficientNet had 729 FP and 207 FN, DenseNet resulted 542 FP and 266 FN, ResNet presented 884 FP and 201 FN, NasNet produced 1324 FP and 185 FN, Xception had 536 FP and 353 FN and finally ConvNeXt resulted in 860 FP and 117 FN. These results show that DenseNet is the most successful model in terms of total number of FP and FN.

The dataset employed in this work, which contains 190,000 GAN-generated images, is one of the largest studied in the deepfake detection tasks. This extensive and multimodal dataset offered a valuable chance to evaluate the models' generalization across different kinds of manipulations. Within these conditions, DenseNet performed well and resulted in the best performance.

DenseNet architecture has the merit of feature reuse, and thus has been effective as its datasets are large and training data is heterogeneous. The dense connectivity between layers allowed the model to efficiently capture complex patterns and subtle inconsistencies in deepfake manipulations. InceptionV3 and Xception also performed well because of their ability to capture multi-scale features. These models accounted for the variety of distortions in the dataset, however, because of their slightly lower F1 scores as compared with DenseNet, there is still potential to be fully exploited in the context of also dealing with subtle artifacts. ConvNeXt presented the highest training accuracy of 99.33%, highlighting its capability to adapt to the training data, followed by EfficientNet with 99.31% accuracy. But, test accuracy of ConvNeXt (91.36%) is a little overfitting, highlighting the needs for methods including regularization, and advanced data augmentation to promote generalization. In the ConvNeXt model, performance metrics show 91% test accuracy and an F1-score of 91%. ResNet and NasNet have the worse testing performances. While ResNet's deep architecture provided some benefits, it may not have fully leveraged the dataset's diversity, leading to less effective generalization. That reliance on automated architecture design of NasNet might have made it less able to retain the complex features necessary for this task. The large dataset played a crucial role in evaluating the models' robustness. Models like DenseNet are resistant to these conditions, while others like NasNet failed to generalize effectively. However, the variety of manipulations in the dataset probably uncovered weaknesses in architectures that are not well-suited for fine-grained pattern recognition. The strong performance of DenseNet shows that it is suitable for real-world deployment, which is especially important for applications where understanding the situation's accuracy and recall is crucial. The capacity to process large amounts of data quickly makes it a good candidate for real application deepfake detection systems.

4 CONCLUSION

This study evaluated seven transfer learning models InceptionV3, EfficientNet, ResNet, DenseNet, Xception, NasNet and ConvNeXt for the task of deepfake detection using a dataset of 190,000 images including both real and fake images. The main objective was to compare these architectures in terms of performance in the detection of manipulation in terms of accuracy, precision, recall and F1-score.

DenseNet was obtained as the best-performing model, achieving the highest test accuracy among all evaluated architectures. Its densely connected layers and feature reuse capability allowed it to excel in identifying subtle manipulations present in deepfake images. This robust performance, underlines its usefulness for real-world applications.

The dataset played an important role in shaping the outcomes of the models. Containing many kinds of manipulations, it offered an evaluation not only of model robustness, but also of model generalization power. The large number of samples in the dataset also validated the comparative analysis, providing a stable used as a benchmark. However, while DenseNet demonstrated superior performance, other models like InceptionV3, Xception and ConvNeXt also achieved commendable results, with high accuracy, indicating their capability to handle such tasks effectively.

Future studies will investigate ensemble methodologies to enhance robustness through a combination of different models. Furthermore, we plan to expand our methodology to encompass video-based deepfake detection, and using more advanced models to detect deepfake images.

Acknowledgements

We thank the anonymous reviewers for their valuable feedback and support.

Conflict of Interest Statement

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The study is complied with research and publication ethics.

Artificial Intelligence (AI) Contribution Statement

This manuscript was entirely written, edited, analyzed, and prepared without the assistance of any artificial intelligence (AI) tools. All content, including text, data analysis, and figures, was solely generated by the authors.

Contributions of the Authors

L.E. Demir: Methodology, formal analysis, visualization, data curation, writing & editing.

Y. Canbay: Conceptualization, methodology, writing, visualization, supervision.

REFERENCES

[1] M. S. Rana, M. N. Nobi, B. Murali, and A. H. Sung, "Deepfake detection: A Systematic Literature Review," *IEEE Access,* vol. 10, pp. 25494-25513, 2022.

[2] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, "Mesonet: A Compact Facial Video Forgery Detection Network," in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2018: IEEE, pp. 1-7.

[3] P. Prajapati and C. Pollett, "Mri-gan: A Generalized Approach to Detect Deepfakes Using Perceptual Image Assessment," *arXiv preprint arXiv:2203.00108,* 2022.

[4] C. Miao, Q. Chu, W. Li, T. Gong, W. Zhuang, and N. Yu, "Towards Generalizable and Robust Face Manipulation Detection via Bag-Of-Local-Feature," *arXiv preprint arXiv:2103.07915,* 2021.

[5] 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*, 2021, pp. 2185-2194.

[6] S. Kanwal, S. Tehsin, and S. Saif, "Exposing AI Generated Deepfake Images Using Siamese Network With Triplet Loss," *Computing and Informatics,* vol. 41, no. 6, pp. 1541–1562, 2022.

[7] R. Rafique, M. Nawaz, H. Kibriya, and M. Masood, "Deepfake Detection Using Error Level Analysis and Deep Learning," in *2021 4th International Conference on Computing & Information Sciences (ICCIS)*, 2021: IEEE, pp. 1-4.

[8] 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: IEEE, pp. 271-275.

[9] M. Patel, A. Gupta, S. Tanwar, and M. Obaidat, "Trans-DF: A Transfer Learning-Based End-To-End Deepfake Detector," in *2020 IEEE 5th International Conference on Computing Communication And Automation (ICCCA)*, 2020: IEEE, pp. 796-801.

[10] P. Joshi and V. Nivethitha, "Deep Fake Image Detection using Xception Architecture," in *2024 5th International Conference on Recent Trends in Computer Science and Technology (ICRTCST)*, 2024: IEEE, pp. 533-537.

[11] M. Liao and M. Chen, "A New Deepfake Detection Method by Vision Transformers," in *International Conference on Algorithms, High Performance Computing, and Artificial Intelligence (AHPCAI 2024)*, 2024, vol. 13403: SPIE, pp. 953-957.

[12] M. Karki. *deepfake and real images*: https://www.kaggle.com/datasets/manjilkarki/deepfake-and-real-images, (23.12.2024).

[13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818-2826.

[14] M. Tan and Q. Le, "Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks," in *International Conference on Machine Learning*, 2019: PMLR, pp. 6105-6114.

[15] 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*, 2016, pp. 770-778.

[16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4700-4708.

[17] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in *Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1251-1258.

[18] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8697-8710.

[19] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A Convnet for the 2020s," in *Proceedings of the Ieee/Cvf Conference On Computer Vision and Pattern Recognition*, 2022, pp. 11976-11986.

[20] D. M. Powers, "Evaluation: From Precision, Recall and F-Measure to Roc, Informedness, Markedness and Correlation," *Arxiv Preprint Arxiv:2010.16061,* 2020.

[21] A. Utku, Z. Ayaz, D. Çiftçi, and M. A. Akcayol, "Deep Learning Based Classification for Hoverflies (Diptera: Syrphidae)," *Journal of the Entomological Research Society,* Vol. 25, No. 3, Pp. 529-544, 2023.

[22] Y. Canbay, S. Adsiz, and P. Canbay, "Privacy-Preserving Transfer Learning Framework for Kidney Disease Detection," *Applied Sciences,* Vol. 14, No. 19, P. 8629, 2024.

[23] Y. Kaya, Z. Yiner, M. Kaya, and F. Kuncan, "A New Approach to Covid-19 Detection from X-Ray Images Using Angle Transformation with Googlenet and Lstm," *Measurement Science and Technology,* Vol. 33, No. 12, P. 124011, 2022.