



DeepFake Detection Using Deep Learning

Nazneen Mansoor¹ and Alexander Iliev Iliev^{1,2(✉)}

¹ SRH University of Applied Sciences, Berlin, Germany
ailiev@berkeley.edu

² Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria

Abstract. Originating in 2017, DeepFakes utilizes artificial intelligence algorithms, particularly Generative Adversarial Networks (GANs), to generate deceptive videos and images. Despite the wide-ranging applications of GANs, their potential misuse raises ethical concerns. The rapid advancement of technology in multimedia has led to the emergence of malicious activities, such as DeepFake manipulation in politics and entertainment. As manipulated visuals become increasingly realistic, the need for a DeepFake detection approach becomes essential. The research employs two distinct Convolutional Neural Network architectures for classification, emphasizing the escalating importance of DeepFake detection in contemporary society. In this research, DeepFake detection techniques are developed by building two CNN models based on ResNet50 and DenseNet121 architectures. A comparative analysis is done for the two models based on accuracy and F1-score efficiency parameters. Experimental results indicate that the CNN models employed in the study achieve an accuracy of 86% and an F1-score of 0.87, respectively.

Keywords: DeepFakes · DeepFake detection · Deep learning · Convolutional neural network · Artificial intelligence · Machine learning · ResNet50 · DenseNet121

1 Introduction

Over the past years, deep learning technology has proven to be tremendously growing in computer vision. Deep learning techniques have been commonly used in all industries like healthcare, finance, automobile, etc. which provide solutions to complex problems. Nevertheless, the negative aspects of deep learning result in the creation of manipulated videos, photos, or other types of multimedia, which can pose threats to democracy, privacy, and national security. This fake multimedia content is termed “DeepFakes” [13, 14]. The term “DeepFake” which is a combination of ‘deep learning’ and ‘fake’ initially appeared in November 2017 when a Reddit user anonymously published an algorithm to generate fake videos [11]. DeepFakes result from artificial intelligence algorithms that integrate, alter, and superimpose video clips and images to generate realistic fake videos [12]. DeepFakes are generated using digital data which includes videos, images, or audio data to train the deep learning or machine learning models. The training process increases

the chances of creating more DeepFake content in the digital world. As technology advances, the accessibility and the potential to misuse them keep increasing. There are many techniques to create DeepFakes including Generative adversarial networks (GAN), Variational Autoencoders, and diffusion models. The generation of DeepFakes includes training a neural network to explore a vast sample of data to replicate an individual's movements, facial expressions, voice, and other features. With the help of AI tools, and deep learning technology face swapping and face mapping can be achieved using these technologies. GANs are a class of generative algorithms with an architecture that consists of two neural networks, a Generator and a Discriminator [15]. Discriminator has a supervised approach that predicts the data as real or fake, trains on actual data, and feeds the outcome to the generator. The generator, an unsupervised learning network, consists of various hidden layers, activation, and loss functions [16]. The main purpose of the generator is to create fake images based on the discriminator's feedback. GANs are termed adversarial networks because both neural networks are concurrently antagonistic to each other [7]. The backpropagation method is carried out to optimize the weight and bias of the networks until the discriminator is capable of executing the classification tasks. Some commonly used GANs to create fake content are WGAN, DCGAN, DRAGAN, LSGAN, and many more. DeepFake producers can be categorized into four different types: 1) activists or political players 2) fraudsters 3) DeepFake hobbyist community 4) tenable actors like television companies [12]. Figure 1 illustrates a DeepFake generation process by making use of encoder-decoder pairs.

Images created using GANs have extensive applications ranging from text or image to image translation, blending and editing of photographs, face aging, generating new human poses, 3D images, and photo inpainting [7]. However, disagreeing on the benefits of a DeepFake would be biased. The images generated using GAN can effectively deceive the eyes of humans. There are also several DeepFake applications and software that are accessible to the public to create realistic content.

Overall, DeepFake generation algorithms are being developed at an alarming pace and their quality is consistently improving, thus making DeepFake detection tools indispensable in today's society. DeepFakes can be a real threat to society affecting the perceptions and sentiments of people [17]. Politicians, famous celebrities, and other known personalities are impacted by the unethical usage of these AI-based content such as the fabrication of fake news and adult material as well as online impersonation [5]. DeepFakes can also be involved in spreading false information from untrusted sources, money laundering, and data breaching activities. Given these negative implications, DeepFake detection has become crucial in today's society. In this research, a deep learning model is used to classify DeepFakes into real and fake ones. To build the deep learning model, two different CNN architectures are used to compare the performance of these models for detecting DeepFakes. The graph in Fig. 2 represents the number of published papers on the topic of DeepFakes until 2022. Although the entire list of papers is relatively lesser than compared to other domains, the increasing trend proves to be one of the resonant research areas in recent times.

Section 2 of this paper provides an overview of relevant works in the topic of DeepFake detection. Section 3 describes our recommended methodology in detail Sects. 4

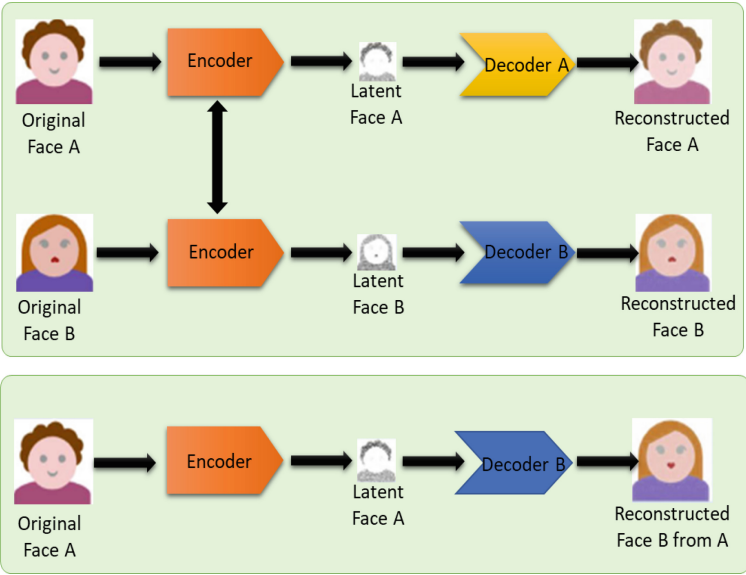


Fig. 1. DeepFake generation model using encoder-decoder [1]

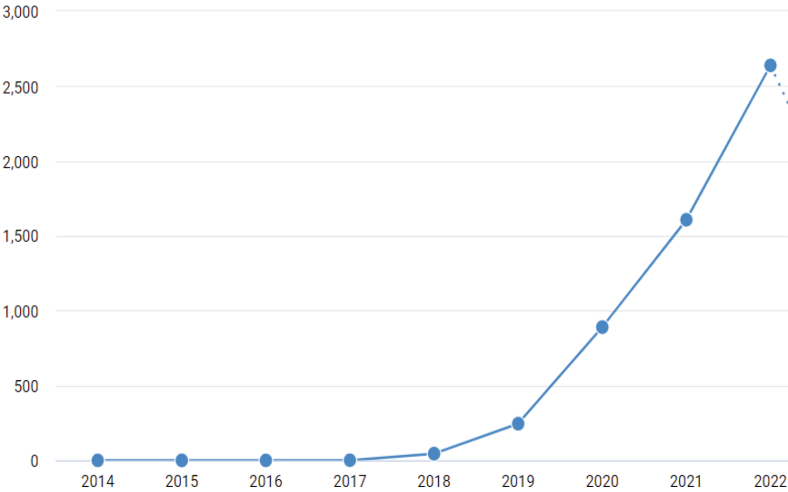


Fig. 2. Number of DeepFakes papers in years from 2014 to 2022, obtained from <https://app.dimensions.ai/> at the end of 2022 with the search keyword “DeepFake detection”

and 5 discuss the results and conclusion, respectively and finally future works is given in Sect. 6.

2 Related Works

With advancements in deep learning technology and the ease of generating fake content, media manipulation has become increasingly popular in recent years. These manipulated multimedia contents are created using the latest AI technologies and tools using deep learning algorithms. The rising number of DeepFake content on various social media platforms demands a DeepFake detection system which could be a solution to various threats in society. Over the past years, there are numerous studies which have been performed by various researchers in the area of DeepFake detection.

It is significant to know the creation of these DeepFakes to understand the different tools and solutions to detect them. A popular deep learning technique for creating DeepFakes is using GANs. GANs are deep learning algorithms that include two neural networks that compete to create new synthetic data. GANs can be separated into three parts [16]:

- Generative – To explain how data is created visually.
- Adversarial – Model training is carried out in an adversarial setting.
- Networks – AI-based deep neural networks are used for training purposes.

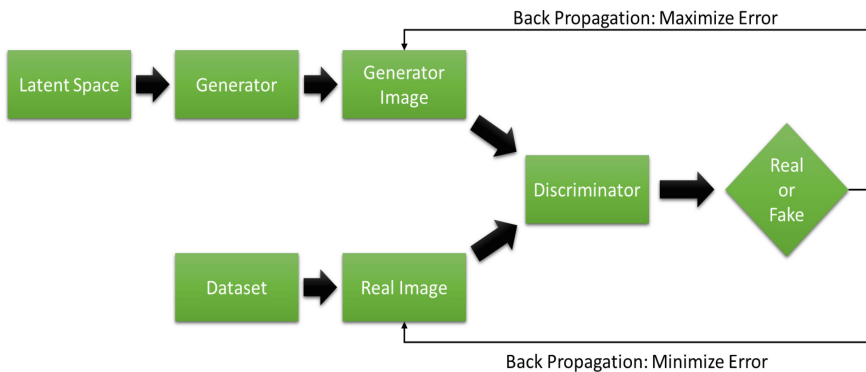


Fig. 3. GAN flow diagram with basic function [18]

Figure 3 depicts a high-level flow diagram of GAN with the basic functions. The generator and discriminator play an antagonistic game during the training phase. The generator's goal is to create samples that deceive the discriminator, whereas the discriminator's goal is to enhance its capacity to differentiate between real and manipulated data. This competitive training forces both networks to develop over time.

Latest advancements in deep learning technology have emerged in creating authentic fake videos effortlessly which used to be done in Hollywood studios in older days. This phenomenon has been reckoned to the outbreak of the crisis in democracy, society, and constitutions and has proven to be a huge imprecation of technological development to human civilization [7]. Sohail [2] proposed a methodology to detect DeepFake media using three different CNN models and a fusion of all three models to enhance generalization capability and robustness. The models were trained and tested on openly

available DeepFake Detection Challenge (DFDC) data consisting of 400 videos and the fusion model attained 99% accuracy. Their proposed approach imposed different image augmentation techniques to make training data diverse which helps the model to learn different sets of features. Vurimi introduced a combination of ResNet CNN architecture and the LSTM (Long Short-Term Memory) approach [3] to detect DeepFake videos. The model is trained in the Celeb-DF v2 dataset which consists of 590 actual YouTube videos and tested with an accuracy rate of 91%. Xueyu in their work [4] developed a novel framework known as “DeepFake disrupter” to protect against DeepFakes using DeepFake detection models mainly Xception, Resnet18, and Resnet50.

Features play a significant role in the classification task in the case of deep learning models. Mousa [5] implemented a CNN-based deep-learning model using mouth features to detect DeepFake videos. They conducted the experiments on Celeb-DF videos and used MoviePy, an open-source application to edit and cut videos specifically to extract mouth features. Artem [6] put forward an approach to compare different EfficientNets models and their performance for DeepFake detection. The authors did a comparison of models on DFDC training data and the study proves that no strong correlation is between the performance of the model and its size. The accuracy rate was higher for EfficientNet B4 and B5 models. Andreas implemented a model [8] to train the FaceForensics dataset which contains facial forgeries using a deep learning-based approach in a supervised fashion. Also, they created a large-scale dataset consisting of manipulated facial images which can be used for improving DeepFake detection models. Ying Xu [9] introduced a different approach to generalize DeepFake detection, unlike other DeepFake detection techniques. They proposed a generalizable model for detecting unknown and novel DeepFake using supervised contrastive learning. Additionally, the proposed approach aids in examining the learned features for explainability.

In the case of deep learning face-related tasks, the quality and quantity of the training data have a great impact on the outcome. The amount of time taken for collecting and labeling massive data quantity of data requires huge labor and remains expensive. Harshil [10] implemented a CNN approach for detecting DeepFake on the CelebDF dataset by including various data augmentation techniques as part of pre-processing. Some of the augmentation techniques used by authors in [10] include GridMask, traditional augmentations, style transfer, and face morphing. The positive and negative effects of each augmentation technique were analyzed and compared for improving the performance. Ipshita [7] contributed to the detection and classification of DeepFake images using pre-trained models such as InceptionV3 and VGG-16. Furthermore, as part of the research work fake images were generated using GAN, and localization and segmentation of fake facial images are done using a segment-based model.

Generally, we can apprehend that researches in the area of detecting fake images are still in scope as the generation of these manipulated images which look realistic is tremendously increasing in the current world. Although we have a walkthrough of the latest developments in this field of research, there are numerous study areas to address and explore.

3 Design and Methodology

In this study, the researchers employ popular pre-trained convolutional neural network (CNN) architectures, namely ResNet-50, and DenseNet121 to detect DeepFake images. DeepFake detection is a crucial application in the age of synthetic media, as it seeks to differentiate between real and manipulated content. The selection of these pre-trained models is significant because they are renowned for their effectiveness in image classification tasks. The researchers employ Celeb-DF, a dataset that presumably contains both real and DeepFake images of celebrities, as the training and testing data. This dataset provides a diverse set of facial images, making it suitable for evaluating the robustness and generalization capability of the DeepFake detection models. By utilizing CNN architectures, the researchers aim to demonstrate the effectiveness of convolutional neural networks in achieving high accuracy in identifying DeepFake content. CNNs have proven to be highly proficient in capturing complex patterns and features within images, which is vital for detecting subtle manipulations commonly found in DeepFake videos and images. The utilization of these pre-trained models further accelerates the development process and enhances the model's ability to discern real from fake.

In the proposed work, the classification task is performed using two CNN models, ResNet-50 and DenseNet121, with a comparative analysis based on model accuracy to evaluate their effectiveness in differentiating between real and fake images. The overall process has been executed using the Jupyter Notebook platform on which deep learning algorithms can be trained at increased speed.

3.1 Dataset

The training dataset for the two Convolutional Neural Network (CNN) architectures plays a crucial role in developing deep learning models for image classification tasks. In this case, the dataset consists of a mix of fake and real images, sourced from the Celeb-DF dataset [19]. This dataset is instrumental in advancing the field of deep learning and computer vision, particularly in addressing the challenges posed by DeepFake videos, which have gained notoriety for their potential to spread disinformation and manipulate digital content.

The Celeb-DF dataset is a valuable resource for researchers in the field of deep learning and computer vision, containing a substantial collection of both DeepFake and real videos. Specifically, the dataset comprises 5,639 DeepFake videos and 590 real videos, providing a comprehensive representation of both synthetic and authentic content. Each video in the collection lasts approximately 13 seconds on average and has a standard frame rate of 30 frames per second. The authenticity of the real videos in the Celeb-DF dataset is notable, as they are curated from publicly accessible YouTube videos. These videos encompass 59 celebrity interviews, which include individuals from diverse ethnic backgrounds, age groups, and genders. This diversity is essential for ensuring that the CNN models are capable of generalizing their learning across various demographics and characteristics.

In the context of the DeepFake videos within the Celeb-DF dataset, they are the result of a face-swapping technique applied to 59 distinct subjects. This technique involves digitally manipulating the facial features of one individual to make them appear as another

person, which is a common method for creating DeepFake content. These synthetic videos pose a significant challenge in the detection of manipulated content, making them an ideal test case for evaluating the effectiveness of the CNN architectures. The dataset contains a substantial number of images, with a total of 19,457 individual frames extracted from the videos. To ensure proper training, validation, and testing of the CNN models, a balanced split of the dataset is employed. Specifically, 80% of the data, amounting to 15,565 images, is allocated for training the models. This large training set enables the models to learn the underlying patterns and features necessary for distinguishing between real and DeepFake content. Furthermore, 10% of the dataset is set aside for validation purposes. This validation set is instrumental during the training process as it helps monitor the model’s performance and adjust hyperparameters to prevent overfitting, ensuring that the model generalizes well to new, unseen data. Finally, the remaining 10% of the dataset is reserved for testing the models. This test set serves as the ultimate evaluation, assessing the CNN architectures’ ability to accurately classify real and fake images (see Fig. 4).

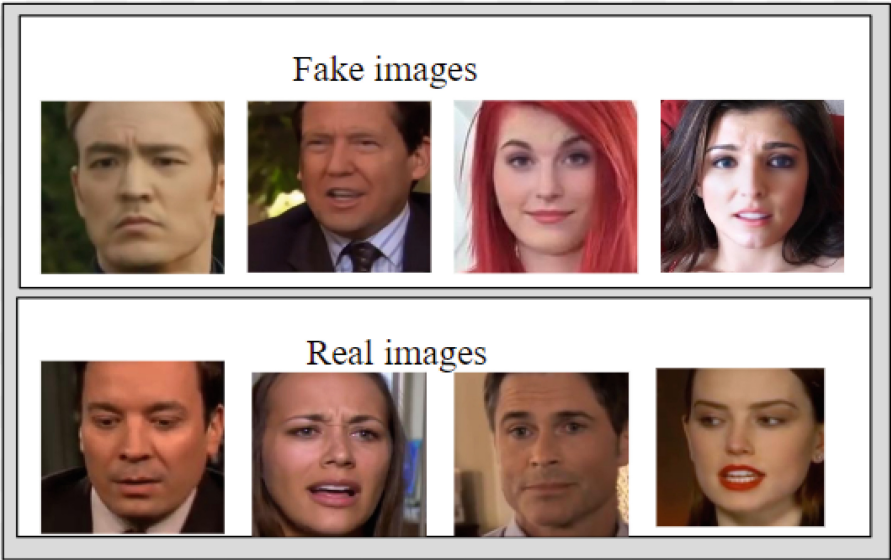


Fig. 4. Real and fake images from training dataset [19]

3.2 Model Architecture

Convolutional neural networks (CNNs) have emerged as powerful tools for addressing this challenge, demonstrating superior performance in distinguishing between real and manipulated images. In this elaboration, we will delve into the key aspects of this research, including the architecture of DeepFake detection models using CNNs and the utilization of three prominent CNN architectures: ResNet-50, and DenseNet121. Moreover, we will explore the adaptations made to these architectures to suit the Celeb-DF

dataset, which differs from the ImageFolder dataset, and the implications of these developments in the broader context of computer vision. Figure 5 provides a visualization of the CNN architecture designed specifically for DeepFake detection models. One of the significant accomplishments in this domain is the detection of manipulated images using the Celeb-DF dataset, a benchmark dataset commonly employed for evaluating DeepFake detection algorithms. To tackle this task effectively, two well-established CNN architectures—ResNet-50, and DenseNet121 have been harnessed.

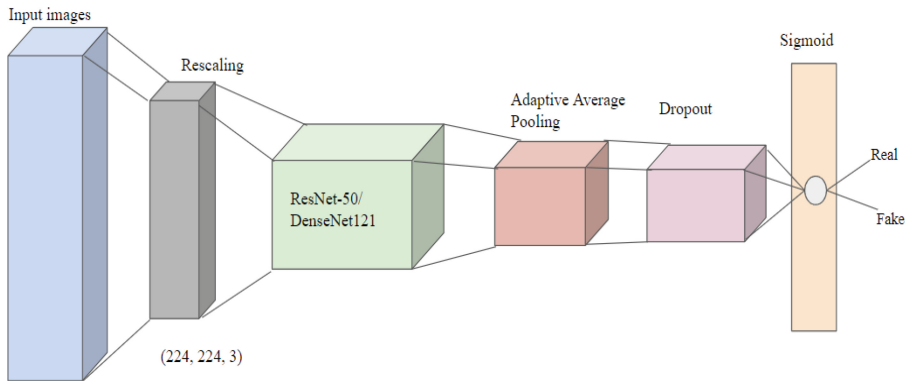


Fig. 5. CNN architecture of DeepFake detection model

A critical consideration in using the Celeb-DF dataset lies in the variance of image dimensions compared to datasets like ImageFolder. To address this discrepancy, certain modifications have been introduced to the conventional CNN architectures. These adaptations ensure that the models can handle the unique characteristics of the Celeb-DF dataset. Additionally, the output layer of the network has undergone alterations. Unlike the ImageFolder data, where multiple classes are present, this research focuses on binary classification—distinguishing between real and fake images. As a result, the updated architecture is optimized for processing coloured images with a resolution of 224×224 pixels. To ensure uniformity and optimal data pre-processing, the image data are normalized to a scale ranging from 0 to 1 through a rescaling layer. Following this pre-processing step, the rescaled images are passed through the base models: ResNet-50, Inception V3, and VGG-16. The output from these base models is then subjected to an adaptive average pooling layer, followed by a dropout layer with a dropout rate of 50%. Finally, a sigmoid function is applied in the output layer, which is tailored for binary classification tasks, serving as the model's decision-making mechanism to classify images as either real or fake.

ResNet50. ResNet-50 represents an evolution of the ResNet architecture and is a CNN with 48 convolution layers, one average pooling layer, and one max pooling layer, totalling 50 layers in depth. After the introduction of AlexNet, ResNet has been one of the most influential architectures in the realm of deep learning. ResNet's innovation lies in its establishment of residual connections between different layers, an approach that has led to reduced loss, enhanced knowledge acquisition, and improved training

performance. When a layer features a residual connection, it indicates that the layer's output is a convolutional combination of both its input and the input from an earlier layer. This architectural breakthrough has had a profound impact, enabling the creation of exceptionally deep networks while addressing challenges like vanishing gradients. ResNet-50, in particular, has gained popularity for its ability to capture complex image features effectively, making it an ideal candidate for DeepFake detection.

DenseNet121. DenseNet121 is a convolutional neural network architecture widely used for image classification tasks. It stands out for its dense connectivity pattern, where each layer receives input from all preceding layers, promoting feature reuse and efficient parameter utilization. With 121 layers, DenseNet achieves impressive accuracy while addressing issues like vanishing gradients. This design improves information flow, making DenseNet121 effective for tasks such as scene understanding and object recognition in computer vision.

For evaluating CNN-based classifiers (ResNet50, and DenseNet121) detection efficiency, a test dataset is formed which includes images from the Celeb-DF dataset. The performance metric used to compare the DeepFake detection models are accuracy, recall, F1-score, and precision. Cross-Entropy loss is selected as the loss function and Adam optimizer of learning rate 0.0001. Also, for ResNet-50 and DenseNet121, the batch size is opted as 32 for training, validation, and test datasets. After conducting multiple experiments with different sets of epochs, the epoch value is set to 15. Hyperparameter tuning has been done to choose the parameters after doing multiple experiments with various loss functions, optimizers, and batch size combinations. The experimented batch size values are 16, 32, and 64 with Adam, and Stochastic gradient descent optimizers and different sets of learning rates ranging from 0.01 to 0.00001.

4 Results and Discussion

In the proposed work, DeepFake detection is performed using two commonly used CNN architectures: Resnet50, and DenseNet121. The performance metrics used to compare all three models include accuracy, precision, recall, and F1-score. The results from the test data are represented in Table 1.

Figure 6 shows the learning curves obtained by training the CNN models with different hyperparameters and includes both accuracy and loss plots with an increasing number of epochs. The learning curve plots show the results obtained over 15 epochs. The optimizer used for all three models is the same (Adam optimizer) with a learning rate of 0.0001 and batch size of 32. All the hyperparameters are chosen after conducting multiple experiments on the models.

The results detail the efficiency comparison of different CNN architectures for detecting DeepFake images using various performance metrics and learning curves. When models are evaluated on the same datasets they were trained on, nearly all of the studies exhibit outstanding performance scores, but if the models are tested on unseen data, they appear to struggle. When working with real-time data which is unbalanced, the training, validation, and testing accuracies are not effective performance indicators of models due to the vast difference between false negatives and false positives. Hence, it is relevant

Table 1. Comparative analysis

Model	Accuracy (%)	Precision	Recall	F1-score
ResNet-50	86.3	0.861	0.917	0.888
DenseNet121	86.5	0.931	0.833	0.879

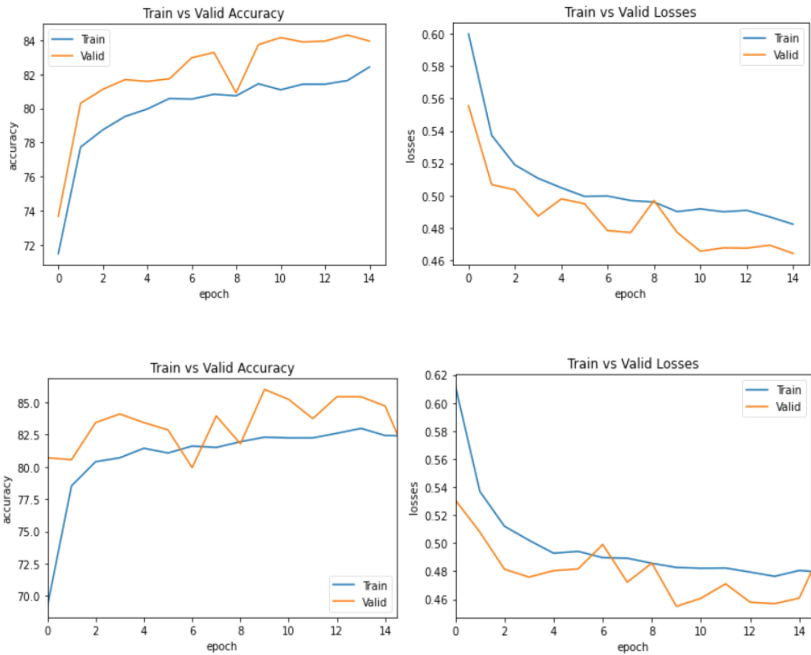


Fig. 6. Learning curves of DeepFake detection models

to include other metrics to evaluate the efficiency of the models other than accuracy. The results from the experiments show both the models exhibit similar performances with an accuracy range of 86% and F1-score of 0.88. Although there are fluctuations in the validation curve of DenseNet121 results, both training and validation accuracies and loss functions converge to the same point after 15 epochs. For choosing the number of epochs the learning loss curve pattern is studied and by using the early stopping technique, overfitting issues are controlled to a great extent.

From the experiments, it is clear that the validation data has lots of fluctuations compared to the training data which can be practically possible with real-time data. Ultimately, by performing hyperparameter tuning the results were improved for all three models. The fluctuations in the validation data can be improved by implementing different cross-validation techniques for data split which is not performed during the study. In terms of training time for 15 epochs, DenseNet121 has less training time compared to ResNet50.

5 Conclusion

With the advent of technology in the multimedia sphere, several malicious activities, such as the misuse of forged videos and images in different domains like politics and entertainment, have evolved as a negative element of it. This development in multimedia has made manipulated and real images nearly identical, leading to extensive research in this domain. The most commonly used phrase, ‘DeepFake,’ has the capability of producing hyper-realistic visuals. This research work aims in building a DeepFake detection technique using two CNN architectures.

The experiments showcase comparable performances of the models, with an accuracy range of 86% and an F1-score of 0.88. Noteworthy is the observation of fluctuations in the validation curve of DenseNet121, although both training and validation accuracies, along with loss functions, converge after 15 epochs. The adoption of the early stopping technique successfully addresses overfitting concerns. In terms of training time, the study notes that DenseNet121 outperforms ResNet50 in efficiency, completing 15 epochs in less time.

This research work provides an overview of how the models differentiate the images from real or manipulated ones by building DeepFake detection models using CNN architectures. These DeepFake Detection models serve as crucial safeguards, helping mitigate the risks associated with malicious misinformation, identity theft, and fraudulent activities.

6 Future Work

Despite the availability of various DeepFake detection tools or models to identify forged images or videos, current research often lacks reasoning or explanation for image classification. Interpretability plays a pivotal role in building trust with end-users and stakeholders, which is essential for the broader adoption of DeepFake detection technology. In light of these challenges, it has become increasingly evident that there is a need for the development of interpretation methods specifically tailored for neural networks used in DeepFake detection. To enhance the reliability and trustworthiness of DeepFake detection systems in the future, it can be suggested to develop more explainable and interpretable models by making use of Explainable AI methods.

References

1. Nguyen, T.T., et al.: Deep learning for DeepFakes creation and detection: a survey. *Comput. Vis. Image Underst.* **223**, C (2022). <https://doi.org/10.1016/j.cviu.2022.103525>
2. Khan, S.A., Artusi, A., Dai, H.: Adversarially robust DeepFake media detection using fused convolutional neural network predictions (2021)
3. Vamsi, V.V.V.N.S., et al.: DeepFake detection in digital media forensics. *Glob. Transit. Proc.* **3**(1), 74–79 (2022). <https://doi.org/10.1016/j.gltp.2022.04.017>
4. Wang, X., Huang, J., Ma, S., Nepal, S., Xu, C.: DeepFake disrupter: the detector of DeepFake is my friend. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, pp. 14900–14909 (2022). <https://doi.org/10.1109/CVPR52688.2022.01450>

5. Jafar, M.T., Ababneh, M., Al-Zoube, M., Elhassan, A.: Forensics and analysis of DeepFake videos. In: 2020 11th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, pp. 053–058 (2020). <https://doi.org/10.1109/ICICS49469.2020.239493>
6. Pokroy, A.A., Egorov, A.D.: EfficientNets for DeepFake detection: comparison of pretrained models. In: 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), St. Petersburg, Moscow, Russia, pp. 598–600 (2021). <https://doi.org/10.1109/ElConRus51938.2021.9396092>
7. Rong, I.: Detection and segmentation of DeepFake face images generated by GANs using segmented based CNN, 1 January 2020. <https://esource.dbs.ie/handle/10788/4244>
8. Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., Niessner, M.: FaceForensics++: learning to detect manipulated facial images. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), pp. 1–11 (2019). <https://doi.org/10.1109/ICCV.2019.00009>
9. Xu, Y., Raja, K., Pedersen, M.: Supervised contrastive learning for generalizable and explainable DeepFakes detection. In: 2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW), Waikoloa, HI, USA, pp. 379–389 (2022). <https://doi.org/10.1109/WACVW54805.2022.00044>
10. Improving DeepFake Video Detection Using Data Augmentation Techniques | Research Square (n.d.). <https://doi.org/10.21203/rs.3.rs-1844392/v1>
11. DeepFake Technology - Organization for Social Media Safety: Organization for Social Media Safety (n.d.). <https://www.socialmediasafety.org/advocacy/deepfake-technology/>
12. Westerlund, M.: The Emergence of DeepFake Technology: A Review | TIM Review, 1 November 2019. <https://timreview.ca/article/1282>
13. What are DeepFakes? How fake AI-powered audio and video warps our perception of reality. Business Insider (n.d.). <https://www.businessinsider.com/guides/tech/what-is-deepfake>
14. Team, G.L.: All You Need to Know About DeepFake AI. Great Learning Blog: Free Resources What Matters to Shape Your Career! 14 May 2021. <https://www.mygreatlearning.com/blog/all-you-need-to-know-about-deepfake-ai/>
15. Bao, J., Chen, D., Wen, F., Li, H., Hua, G.: CVAE-GAN: fine-grained image generation through asymmetric training. In: 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, pp. 2764–2773 (2017). <https://doi.org/10.1109/ICCV.2017.299>
16. Agrawal, R.: An End-to-End Introduction to Generative Adversarial Networks (GANs). Analytics Vidhya, 20 October 2021. <https://www.analyticsvidhya.com/blog/2021/10/an-end-to-end-introduction-to-generative-adversarial-networksgans/>
17. Box, B.B.: How DeepFake Technology Impact the People in Our Society? Medium, 3 February 2020. <https://becominghuman.ai/how-deepfake-technology-impact-the-people-in-our-society-e071df4ffc5c>
18. P.: Inside the Generative Adversarial Networks (GAN) architecture. Medium, 19 November 2019. https://medium.com/@Packt_Pub/inside-the-generative-adversarial-networks-gan-architecture-2435afbd6b3b
19. Ali, L., Alnajjar, F., Jassmi, H.A., Gocho, M., Khan, W., Serhani, M.A.: Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures. *Sensors* **21**, 1688 (2021). <https://doi.org/10.3390/s21051688>

Lecture Notes in Networks and Systems 1018

Kohei Arai *Editor*

Intelligent Computing

Proceedings of the 2024 Computing
Conference, Volume 3



Springer

Series Editor

Janusz Kacprzyk , *Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*

Advisory Editors

Fernando Gomide, *Department of Computer Engineering and Automation—DCA, School of Electrical and Computer Engineering—FEEC, University of Campinas—UNICAMP, São Paulo, Brazil*

Okay Kaynak, *Department of Electrical and Electronic Engineering, Bogazici University, Istanbul, Türkiye*

Derong Liu, *Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, USA*

Institute of Automation, Chinese Academy of Sciences, Beijing, USA

Witold Pedrycz, *Department of Electrical and Computer Engineering, University of Alberta, Alberta, Canada*

Systems Research Institute, Polish Academy of Sciences, Warsaw, Canada

Marios M. Polycarpou, *Department of Electrical and Computer Engineering, KIOS Research Center for Intelligent Systems and Networks, University of Cyprus, Nicosia, Cyprus*

Imre J. Rudas, *Óbuda University, Budapest, Hungary*

Jun Wang, *Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong*

The series “Lecture Notes in Networks and Systems” publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the worldwide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Kohei Arai
Editor

Intelligent Computing

Proceedings of the 2024 Computing
Conference, Volume 3

Editor
Kohei Arai
Faculty of Science and Engineering
Saga University
Saga, Japan

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-3-031-62268-7 ISBN 978-3-031-62269-4 (eBook)
<https://doi.org/10.1007/978-3-031-62269-4>

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Preface

It is with great pleasure that we introduce the proceedings of the Computing Conference 2024, held on July 11 and 12, 2024. This conference served as a platform for researchers and professionals from around the globe to convene and exchange ideas at the forefront of computing and its diverse applications. The enthusiasm and dedication displayed by participants underscored the significance of this event in fostering collaboration and advancing the field.

We received an overwhelming total of 457 contributions from esteemed scholars and practitioners. These submissions underwent a rigorous double peer-review process, facilitated by experts in their respective domains. After careful evaluation and deliberation, a total of 165 papers were selected for publication in these proceedings.

The diverse array of topics covered in these papers reflects the breadth and depth of contemporary computing research, spanning areas such as artificial intelligence, machine learning, cybersecurity, data science, and beyond. Each paper represents a valuable contribution to the collective knowledge base of the computing community, offering insights, innovations, and solutions to pressing challenges.

We extend our heartfelt gratitude to all authors, reviewers, organizers, and attendees whose efforts and contributions made this conference a resounding success. It is our hope that the insights shared and connections forged during this event will continue to inspire and propel advancements in computing for years to come.

Regards,
Kohei Arai

Contents

Selective Coordination Study in a Renewable Energy-Based Microgrid	1
<i>Matthew Zorumski, Hessam Keshtkar, and Farideh Doost Mohammadi</i>	
SSIVD-Net: A Novel Salient Super Image Classification and Detection Technique for Weaponized Violence	16
<i>Toluwani Aremu, Li Zhiyuan, Reem Alameeri, Mustaqeem Khan, and Abdulmotaleb El Saddik</i>	
An Intelligent Environment Application Case to Manage Comfort Preferences, at an University Residence	36
<i>Pedro Filipe Oliveira and Paulo Matos</i>	
Enhancing Explainability in Mobility Data Science Through a Combination of Methods	45
<i>Georgios Makridis, Vasileios Koukos, Georgios Fatouros, Maria Margarita Separdani, and Dimosthenis Kyriazis</i>	
A Trust Model for Informed Agent Collaboration in Complex Tasks	61
<i>Ricardo Barbosa, Ricardo Santos, and Paulo Novais</i>	
Offensive Language Detection from Arabic Texts	77
<i>Arafat A. Awajan</i>	
Identifying Potential Customer Using Graph Social Media Analytics	92
<i>Siti Nur Aisyah Anuar, Noryanti Muhammad, and Mohd Izhar Firdaus</i>	
Categorizing Philanthropic CSR Activities Through Hybrid Approach of Fuzzy Clustering and Topic Modeling	104
<i>Nik Siti Madiah Nik Mangsor, Syerina Azlin Md Nasir, Shuzlina Abdul-Rahman, and Zurina Ismail</i>	
A Ride-Hailing Company Supply Demand Prediction Using Recurrent Neural Networks, GRU and LSTM	123
<i>Sahand Fathi, Soheil Fathi, and Vahid Balali</i>	
A Comprehensive Literature Review on Slope Entropy Algorithm: Bridging Past Insights with Future Directions	134
<i>Juan Seguí Moreno and Antonio Molina Picó</i>	

Engineering Research Strategies for Investigating Long-Term Automation Effects, Behavioural Adaptation and Change Processes: Experts' Views	145
<i>Naomi Y. Mbelekani and Klaus Bengler</i>	
Cardiac MRI Semantic Segmentation for Ventricles and Myocardium Using Deep Learning	169
<i>Racheal Mukisa and Arvind K. Bansal</i>	
Problems of Building Digital Twins of Escalators at Subway Stations Based on Machine Learning	189
<i>Alexey Subbotin, Nataly Zhukova, and Mikhail Gudilov</i>	
DeepFake Detection Using Deep Learning	202
<i>Nazneen Mansoor and Alexander Iliev Iliev</i>	
AI Based Commercial Decisions: The Cryptocurrency Market Case	214
<i>Sujata Joshi, Mohit Satya, and Menachem Domb</i>	
Applying Formal Methods to Elicit Specifications for Grid Data Analysis Applications Using Machine Learning Algorithms	224
<i>Vinitha Hannah Subburaj and Anitha Sarah Subburaj</i>	
ChatGPT and Artificial Intelligence in Higher Education: Literature Review Powered by Artificial Intelligence	240
<i>Andrej Čep and Andrija Bernik</i>	
Weak Relation Enforcement for Kinematic-Informed Long-Term Stock Prediction with Artificial Neural Networks	249
<i>Stanislav Selitskiy</i>	
A Modified Ant Colony Optimization with KNN for High-Dimensional Data Classification	262
<i>Gideon Popoola, Gerard Shu Fuhnwi, Janet O. Agbaje, and Kayode Fesomade</i>	
Cross Approach Between Modern Artificial Intelligence and Emergency Medicine: A Review	278
<i>Kokou Edjinedja, Oussama Barakat, Thibaut Desmettre, Tania Marx, Omar Elfahim, and Charlotte Bredy-Maux</i>	
GraphRank Pro+: Advancing Talent Analytics Through Knowledge Graphs and Sentiment-Enhanced Skill Profiling	298
<i>Sirisha Velampalli and Chandrashekar Muniyappa</i>	

The Advanced Optimal Control Problem and Approach to Its Numerical Solving	310
<i>Askhat Diveev</i>	
SGAN: Appliance Signatures Data Generation for NILM Applications Using GANs	325
<i>Christina Gkoutroumpi, Nikolaos Virtsionis Gkalinikis, and Dimitrios Vrakas</i>	
Multi-modal Deep Learning for Detecting Toxicity in Transcribed-Audio Conversations	340
<i>Ismail El Sayad, Josue Gourde, Jake Pott, Sachin Muthayan, and Simranjit Singh</i>	
Action Unit Recognition: Leveraging Weak Supervision with Large Loss Rejection	349
<i>Oseluole Enabor and Hima Vadapalli</i>	
Performance Analysis of Synthetic Events via Visual Object Trackers	364
<i>Mohamad Alansari, Hamad AlRemeithi, Sara Alansari, Naoufel Werghi, and Sajid Javed</i>	
Flight Plan Optimization, Scanning and 3D Reconstruction Using an Unmanned Aerial Vehicle	385
<i>Andrija Bernik, Boris Tunjić, and Andrej Čep</i>	
Text Extraction and Recognition in Manga Comics Using Image Processing Techniques	408
<i>Naser Al-Ibrahim, Mohamed Al-Ibrahim, and Waleed Al-Awadhi</i>	
A Comparative Study on Plant Diseases Using Object Detection Models	419
<i>Abu Adnan Sadi, Ziaul Hossain, Ashfaq Uddin Ahmed, and Md. Tazin Morshed Shad</i>	
Gender Identification Using 2D Ear Biometric	439
<i>Lavanya Srinivasan</i>	
Detecting COVID-19 in Computed Tomography Images: A Novel Approach Utilizing Segmentation with UNet Architecture, Lung Extraction, and CNN Classifier	450
<i>Kenan Morani, Esra Kaya Ayana, Dimitrios Kollias, and Devrim Unay</i>	
Attention U-Net Oriented Towards 3D Depth Estimation	466
<i>Leonel Jaime Ocsa Sánchez and Juan Carlos Gutiérrez Cáceres</i>	

Multi-camera Tracking Based on Spatio-Temporal Association in Small Overlapping Regions	484
<i>Lap Quoc Tran, Manh Cong Pham, and Quang Nhat Nguyen</i>	
Development of UX Evaluation Tool in Digital Media Content Production: Focusing on the Mobile Environment	504
<i>Sokjong Jang and Joungkyung Cho</i>	
Optimizing E-Sports Revenue: A Novel Data Driven Approach to Predicting Merchandise Sales Through Data Analytics and Machine Learning	522
<i>Md Abu Sufian, Jayasree Varadarajan, Mahesh Hanumanthu, Lakshmisakuntala Katneni, Ayeshsa Jamil, Vinay Lal, and Jason Boomer</i>	
Efficient Graph Encoder Embedding for Large Sparse Graphs in Python	568
<i>Xihan Qin and Cencheng Shen</i>	
Predictive Savings Service	578
<i>Dimitar Kazakov and Ventsislav Nikolov</i>	
Malaysian Local Fruits E-Marketplace System with Intelligent Promotional Feature	588
<i>Check-Yee Law, Yong-Wee Sek, Choo-Chuan Tay, Lee-Ying Chong, and Abdul Hanif Bin Mislam</i>	
Theory of Programming Adoption	602
<i>Isaac Atta Senior Ampofo, Isaac Atta Junior Ampofo, Beatrice Ampofo, Eugene Louis Batie Badzongoly, Francis Ohene Boateng, and William Asiedu</i>	
Energy Aware Workload Scheduling Metrics for Execution of Parallel Application in Heterogeneous Cloud Computing Platform	618
<i>K. N. Divyaprabha and T. S. B. Sudarshan</i>	
True Random Numbers from Pairs of SRAM Devices	630
<i>Manuel Aguilar Rios, Michael Partridge, Saloni Jain, and Bertrand Cambou</i>	
Benchmarking Metaheuristic-Integrated QAOA Against Quantum Annealing	651
<i>Arul Rhik Mazumder, Anuvab Sen, and Udayon Sen</i>	

Quantum Circuits for Fixed Matching Substring Problems 667
Domenico Cantone, Simone Faro, Arianna Pavone, and Caterina Viola

Author Index 687