Scientific Research Publishing

Journal of Computer and Communications, 2023, 11, 1-13
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Deepfake Video Detection Employing Human Facial Features

Daniel Schilling Weiss Nguyen[*], Desmond T. Ademiluyi

Computer Science (DSCS), Aspen University, Phoenix, USA
Email: *dans515e@gmail.com, desmondtoyus@gmail.com

How to cite this paper: Nguyen, D.S.W. and Ademiluyi, D.T. (2023) Deepfake Video Detection Employing Human Facial Features. *Journal of Computer and Communications*, **11**, 1-13.
https://doi.org/10.4236/jcc.2023.1112001

Received: September 10, 2023
Accepted: December 4, 2023
Published: December 7, 2023

Abstract

Deepfake technology can be used to replace people's faces in videos or pictures to show them saying or doing things they never said or did. Deepfake media are often used to extort, defame, and manipulate public opinion. However, despite deepfake technology's risks, current deepfake detection methods lack generalization and are inconsistent when applied to unknown videos, *i.e.*, videos on which they have not been trained. The purpose of this study is to develop a generalizable deepfake detection model by training convoluted neural networks (CNNs) to classify human facial features in videos. The study formulated the research questions: "How effectively does the developed model provide reliable generalizations?" A CNN model was trained to distinguish between real and fake videos using the facial features of human subjects in videos. The model was trained, validated, and tested using the FaceForensiq++ dataset, which contains more than 500,000 frames and subsets of the DFDC dataset, totaling more than 22,000 videos. The study demonstrated high generalizability, as the accuracy of the unknown dataset was only marginally (about 1%) lower than that of the known dataset. The findings of this study indicate that detection systems can be more generalizable, lighter, and faster by focusing on just a small region (the human face) of an entire video.

Keywords

Artificial Intelligence, Convoluted Neural Networks, Deepfake, GANs, Generalization, Deep Learning, Facial Features, Video Frames

1. Introduction

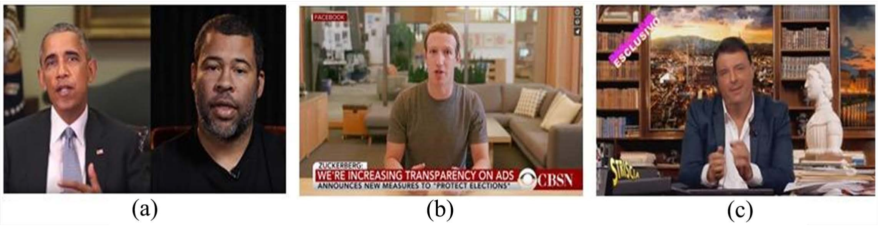
The rise of deepfake technologies has given rise to images and videos that are incredibly realistic but carry the potential for harmful consequences in society

[1]. An analysis by Deeptrace Labs [2], revealed that the prevalence of deepfakes on the internet has nearly doubled in recent years, with women, especially celebrities, being the subjects of 96% of deepfake videos used for sexual fantasies, intimidation, and blackmail. The proliferation of deepfake technologies has significantly impacted the credibility of digital information and has facilitated attacks on corporations, individuals, and other groups. Moreover, there has been a surge in the use of "shallowfakes", which are minimally manipulated recordings. As the researchers [2] pointed out, the hazards posed by deepfakes are no longer hypothetical, and there is an urgent need for a wise and timely reaction.

Trust in visual content is gradually eroding [3] as deepfake generation methods become more realistic and credible over time and more adept at evading detection methods [4] [5]. Deepfake techniques have advanced to the point where humans cannot distinguish between real and fake video content, raising concerns about deepfake technology [6]. Deepfakes could have substantial repercussions for determining the trustworthiness of news articles distributed by the media (see **Figure 1**) and offer a new threat to politics, business, and individual privacy [7]. However, despite advancements in deepfake creation technology, existing deepfake detection algorithms have historically performed poorly on untrained video footage from platforms such as YouTube and Facebook.

1.1. Statement of the Problem

Most existing deepfake detection algorithms are ineffective when applied to unknown or untrained videos due to their tendency to overfit a specific dataset, leading to a lack of generalization [8]. The existing deepfake detection algorithms lack generalization and are inconsistent when applied to unknown videos, *i.e.*, videos on which they have not been trained. Several deepfake detection algorithms [9] have reported extremely high detection rates. However, those high values were obtained for videos with which the models were already familiar. Most of these approaches perform less well when confronted with previously unseen videos, such as real-world deepfake videos [10]. Hence, researchers have urged further studies to understand why existing deepfake detectors fail and why their detection rates for unknown videos remain poor [5]. Furthermore, the researchers in [10] noted that even the most advanced cutting-edge detection


FACEBOOK
ZUCKERBERG
WE'RE INCREASING TRANSPARENCY ON ADS
ANNOUNCES NEW MEASURES TO "PROTECT ELECTIONS"
CBSN
ESCLUSIVO
STRISCIA

(a) (b) (c)

Figure 1. Deepfake examples: (a) Obama, (b) Mark Zuckerberg, and (c) Matteo Renzi. Reprinted from "DeepFake Detection by Analyzing Convolutional Traces" by Guarnera, Giudice & Battiati, 2020, CVPR workshop. Copyright 2020 by CVPR.

technology available today can be misled into mistaking a counterfeit video for a genuine one. Considering the potentially disastrous repercussions of deepfake videos, developing a generalizable and effective method for recognizing deepfake video footage is crucial [11].

1.2. Research Question and Hypotheses

This research developed a CNN model to classify real and fake videos based on human facial features. To evaluate the model and address the research problems, this research formulated the following questions and hypotheses to assess the model's accuracy:

RQ. Is the developed model generalizable?

H0: The model's average accuracy for known and unknown datasets is not equal.

Ha: The model's average accuracy for known and unknown datasets is equal.

2. Literature Review

Researchers have continued to contribute to the field of deepfake detection. The researchers in [12] developed a method that improves the performance of detection networks by using the VGG19 network as a capsule network. The capsule network consists of main capsules and output capsules, with dynamic routing used to determine whether the acquired characteristics of the main capsules align in real-time. Similarly, the researchers in [1] created a CNN named MesoInception-4 to analyze the mesoscopic characteristics of images, while Wang *et al.* proposed a hierarchical neuron activity model for inclusion in their FakeSpotter, which displays high resilience against common perturbation assaults. In another research [13] developed a strategy for detecting deepfakes by analyzing the Spatio-temporal features of video streams using a CNN/recurrent neural network (RNN) architecture and algorithmic weighting to emphasize the most reliable features of a video-level prediction, which showed promising results in terms of detection accuracy and potential generalization capability.

Researchers have also proposed different approaches to deepfake detection, Güera [14] suggested using RNN to detect deepfake videos, and Zhao [15] used optical flow to capture the apparent fluctuations in facial expressions across successive frames. Masi [16] developed a novel manipulation detection system called SSTNet, which utilizes low-level artifacts and temporal variations to identify manipulations. Ruff [17] proposed a novel loss function to improve the separation of modified face regions from the rest of the face. However, while temporal consistency-based detection systems have made significant progress in recent years, there is still room for improvement in their generalization capabilities.

Nguyen [18] observed that deepfake videos lack global uniformity in light absorption and reflection, resulting in visual irregularities. The researcher used this knowledge to distinguish between genuine and bogus video content, the researchers achieved an accuracy of 98.9%. They utilized a convex polygon form based on the contours of the eyes and the corners of the lips to create more realistic negative scenarios on the computer screen. By analyzing a wide range of deepfake video datasets, the study found significant differences in hue and resolution between the internal face and backdrop parts. The researcher obtained an average accuracy of 91%. Similarly, the researchers in [19] hypothesized that the colors of actual and false camera images were significantly different when comparing the two types of images.

The researchers in [20] proposed a noise print strategy for identifying and tracking down deepfake videos and images. The study calculated the degree of resemblance between a person's face and the backdrop by dividing the face and the background into equal pieces and calculating the difference between them. The technique was evaluated on a publicly available dataset, and it reported excellent detection accuracy. Li [21] developed a deep neural network model called the long-term recurrent CNN (LRCN) to detect irregular eye blink frequency in deepfake videos. The researchers in [22] developed a system called Expression Manipulation Detection (EMD) for detecting deepfake videos. EMD employs discriminative feature maps derived from a framework for detecting facial expressions to identify identity changes and facial feature changes in images and videos. This approach was evaluated on the Face2Face and NeuralTextures datasets, and the accuracy was around 99%.

Although several deepfake detection algorithms have reported high detection rates, they tend to perform less effectively when presented with unknown videos. Hence, this work bridged this gap by developing a generalizable deepfake detection model. The next section of the paper will provide a detailed description of the developed deepfake detection model, including its implementation, validation, and testing processes.

3. Methodology

The developed approach in this study was divided into three main phases: feature extraction, classification, and decision-making. In the first phase, videos were segmented into frames, and human faces were detected in each frame. Face alignment, compression, and region of interest (ROI) extraction were then performed on the frames. In the second phase, Convolutional Neural Networks (CNNs) were trained to classify human facial characteristics and differentiate between real and fake videos by analyzing the facial features in the frames. The model utilized this knowledge to classify videos as fake or real. In the third phase, Bayes probability was employed to calculate the final fake score of videos and make predictions about their authenticity. To assess the effectiveness of the model, research questions were formulated, and null and alternative hypotheses were constructed for the question. The results were analyzed using Minitab software, and the statistical significance of the findings was determined using the two-sample t-test, consistent with similar research works [23].

This study employed a quantitative research methodology, where the extracted facial features were converted into numerical form and inputted into the CNNs to determine the authenticity of videos. The model analyzed the relationships between the facial features of the subjects in videos to determine their authenticity. The main steps in building the model are outlined below, and a visual representation can be seen in Figure 2.

1) Faces are recognized and extracted from video frames using the Dlib library. The model then draws a quadratic bounding box around the face region. The images are preprocessed and aligned to reduce any irregularities.

2) CNN models are trained on the preprocessed faces to categorize frames as fake or genuine based on human facial characteristics.

3) The output of step 2 is fed to a Bayesian classifier, which calculates the Bayes probability of video clips by combining the fake scores of all frames. The final score is then compared to a predetermined threshold to determine the authenticity of the video.

The model first collects a diverse dataset of real and deepfake videos. Facial features, such as landmarks, textures, and color information, are then extracted from each frame of the videos. These features are represented in a format that can be processed by machine learning algorithms (CNN). Deep learning-based analysis is then performed on the facial features to identify patterns or differences between real and deepfake videos, such as comparing spatial distribution, texture patterns, or color information. The model is trained on the labeled dataset to classify videos as real or deepfake and model is evaluated using a separate set of testing data to assess its accuracy and performance metrics.

3.1. Sample and Population

In this research, publicly available datasets, specifically DFDC and FaceForensic++, were utilized to train and test the detection model. These datasets contain authentic and deepfake videos and were obtained from credible sources. The datasets were downloaded and saved locally for further processing. Preprocessing of the videos was conducted using Keras, a Python library based on TensorFlow.

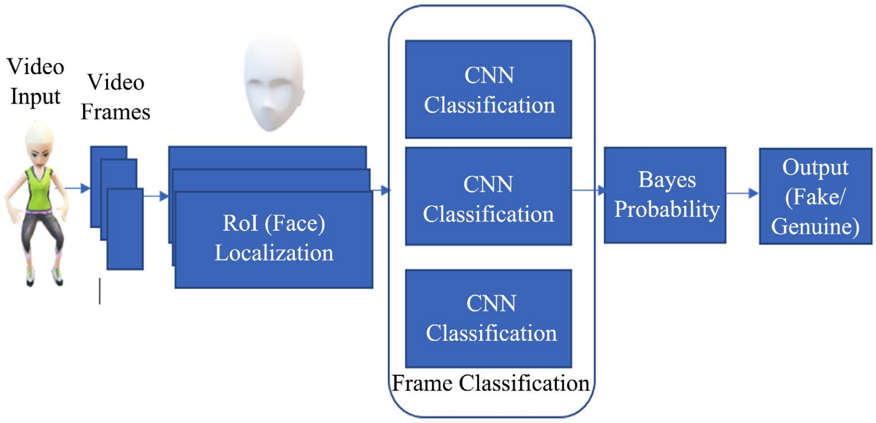


Figure 2. A visual illustration of the developed deepfake detection model.

Over 25,000 videos were preprocessed to adjust aspect ratio, lighting, and geometric consistency, as well as to extract facial features. Image preprocessing techniques such as frame normalization, rotation, cropping, contrast and brightness adjustment, and Gaussian blur were applied to enhance the quality of the videos. This preprocessing step was crucial for the success of the research. The data was organized into fake and real subfolders for each dataset, and Figure 3 illustrates the hierarchical structure of the datasets used in this study.

The training dataset utilized in this study was obtained from the FaceForensics++ video dataset, which consisted of over 500,000 frames from 5000 videos. Permission was obtained from the authors to use this dataset for research purposes. The validation set included 2000 real and fake videos from Subset "08" of the DFDC dataset, which were saved in the "validation" folder. The testing dataset utilized ten subsets from the DFDC dataset, totaling approximately 21,000 videos, and were saved in the "testing" folder. The model was trained and validated concurrently to prevent overfitting.

3.2. Instrumentation

The developed CNN models used multiple convolutional layers, batch normalization, and pooling for frame classification, building on previous research. Frames were converted into numerical data, and CNNs were trained using this data as input. The model utilized convolution filters with different widths for faster feature learning, and learnable weights and biases were assigned to human facial features for training. Batch normalization and dropout techniques were implemented to prevent overfitting. Accuracy, sensitivity, and specificity were used as evaluation metrics.

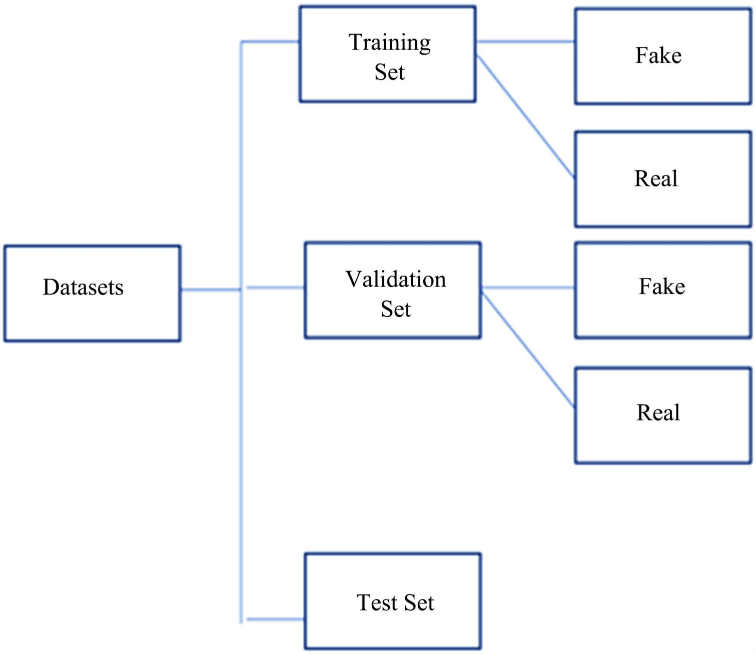


Figure 3. The hierarchical structure of the datasets used in this study.

The model's sensitivity, which represents the number of correctly classified fake videos divided by all fake videos, and specificity, which is the fraction of real videos correctly classified, were calculated to evaluate the model's performance. These metrics can be mathematically expressed in terms of False Positive (FP), False Negative (FN), and True Positive (TP) as follows:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{Number of fake video assessments}}{\text{Total number of all fake videos assesed}} \tag{1}$$

$$\text{Sensitivity} = \frac{\text{TN}}{\text{TN} + \text{FP}} = \frac{\text{Number of real video assessments}}{\text{Total number of all real videos assessed}} \tag{2}$$

Hence,

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}} = \frac{\text{Correct Assessments}}{\text{Total number of assessments by the model}} \tag{3}$$

3.3. Data Analysis

The analysis strategy employed in this study involved using the Minitab software to conduct two-sample t-tests and evaluate the statistical significance of the model's accuracy. The research questions were addressed by comparing the model's results with those of other deepfake detection methods, assessing the model's performance on the training and testing datasets. A significance level (alpha) of 0.05 was used to determine statistical significance, with p-values below this threshold resulting in rejection of the null hypotheses [24].

3.4. Participants

For this investigation, publicly available data from the DFDC dataset was acquired and utilized. The DFDC dataset contains over 100,000 deepfake and genuine videos, totaling approximately 470 GB [25]. The videos are available for download as a single large file or as 50 smaller subsets, each approximately 10 GB [25]. Due to the size of the full dataset and constraints related to download speed and system resources, this study randomly selected 10 subsets, resulting in a total of 20,000 videos. The sample population used in this research aligns with the demographics of the entire dataset population.

The sample size of 20,863 videos was utilized in three evaluations to answer the research questions. Similar studies in literature have employed datasets with varying sizes, ranging from 15 to 50,000 videos. According to Das [26], the average number of samples needed to establish confidence in biometric research is approximately 15,000. Hence, the sample size of 20,863 videos in this study is deemed sufficient.

3.5. Research Question

This research question aims to assess the model's generalizability and answer the primary research question:

RQ1. Is the developed model generalizable?

Put another way, can the developed FakeLooks model maintain the accuracy it demonstrated on the "testing" data when applied to a variety of different datasets? If it can, then the model is generalizable; if it cannot, this mean that the model overfits the training dataset.

Consequently, this study framed the following null and alternative hypotheses to answer this research question:

H_0: The model's average accuracy for known and unknown datasets is not equal.

H_a: The model's average accuracy for known and unknown datasets is equal.

To test these hypotheses, this study evaluated the performance of the model on the testing dataset, which the model was already familiar with due to its use in the initial training. The training dataset consisted of over 25,000 videos, including variations of scale, lighting, and orientation that were introduced into the initial 5000 videos obtained from FaceForensics++ dataset. The model was run on both datasets, and the results are presented in Table 1.

Using the Minitab software, we obtained p-value was 0.025 (see Table 2), which is less than the alpha value of 0.05 indicating that the impact of this study is statistically significant. Therefore, based on these findings, this research rejects the null hypothesis.

4. Discussion

Previous studies have reported reduced model accuracy when tested on different datasets. In this study, the evaluation aimed to investigate if the developed Fakelloks model maintains accuracy when tested on a different dataset compared to the training dataset. The performance of the model was compared on the training dataset (FaceForensics++) and the testing dataset (subsets of the DFDC dataset) by running all the videos through the model and recording the performance. The results were then compared to identify any significant reduction in accuracy for the testing dataset. Figure 3 provides a visual comparison of the results for both datasets. The accuracy of the model when tested on the same dataset used for training was found to be 99.98% (training), while on a different (testing) dataset, it achieved an accuracy of 98.39%. As shown in Figure 4, the variation in accuracy was only about 1%. This demonstrates that unlike other methods, this approach showed great generalization beyond the dataset used for model training.

Table 1. Model results for known and unknown datasets.

Dataset	Accuracy (%)	Sensitivity	Specificity
Training	99.98	99.91	9.96
Testing	98.39	96.88	8.10

Table 2. Estimation for pair difference for known and unknown datasets.

N	Mean	StDev	95% CI for μ Difference	**t-Value**	**p-Value**
2	99.1	1.12	(89.086, 109, 284)	124.79	0.025

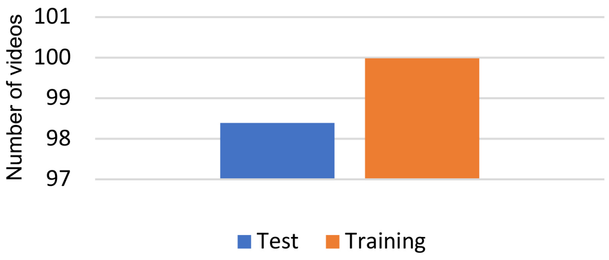


Figure 4. Analysis of results from the training and test datasets.

Figure 4 shows that the model's performance demonstrated good generalization, as the observed discrepancy in precision was only about 1%. Therefore, this study rejects the null hypothesis, as the obtained p-value of 0.025 is less than the significance level of 0.05 used in this study. This result demonstrates that human facial features are sufficient for detecting deepfake videos, which may have implications for the development of lighter and faster deepfake detection systems. However, this study has a few limitations which will be highlighted in the next section.

Limitations

Some of the limitations identified in this research include:

- Limited to videos with human faces: The model developed in this study focuses only on detecting deepfake videos with human faces and may not perform well on videos without human faces or non-facial feature alterations. This limitation may reduce the model's effectiveness in scenarios where non-human objects or scenes are involved.
- High computational costs: The model requires significant computational resources, which may limit its ability to process large videos in real-time applications. The computational costs associated with the model may also pose challenges in terms of scalability and deployment in resource-constrained environments.
- Variability with software and hardware configurations, datasets, and experimental settings: The model's performance may vary depending on the specific software and hardware configurations used for implementation, as well as the characteristics of the datasets and experimental settings employed. This variability may impact the reproducibility and generalization of the model's results across different setups.
- Limited subsets of the dataset: The study utilized limited subsets of the DFDC dataset for model evaluation, which may affect the reproducibility and generalization of the results. The use of limited datasets may not fully capture the diversity and complexity of real-world scenarios, potentially limiting the robustness of the model's performance.
- Lack of interpretability: The model does not provide explanations or justifications for its predictions, which may limit the ability of stakeholders to understand the basis for the model's conclusions. This lack of interpretability

may raise concerns about the transparency and trustworthiness of the model's outputs.

5. Conclusion and Recommendations

5.1. Recommendation

Further research is needed to validate the performance of the model on different data sources and consider potential biases in real-world scenarios. Transfer learning techniques could be explored to reduce training time and improve the model's performance. Comparison with other existing methods and thorough investigation of potential biases in training data are necessary for making conclusive claims about the model's effectiveness. Future research could also expand the model's evaluation on different backgrounds, lighting conditions, and orientations, and focus on providing explanations for the model's predictions to enhance interpretability. Furthermore, it is important to note that deepfake detection methods based on human facial features may have limitations, as deepfake technology continues to evolve and improve. Therefore, a combination of multiple detection methods, including both human facial features and other approaches such as audio analysis and metadata examination, may be more effective in accurately detecting deepfake videos.

5.2. Conclusion

This study demonstrates that human facial features can be effectively used for detecting deepfake videos, and the model developed in this study shows high precision and generalization, with only slight accuracy reduction on unknown datasets. Despite the limitations, the results of this study suggest that the approach outperforms existing deepfake detection techniques in terms of generalization and has implications for the advancement of deepfake detection research. However, regular updates and improvements to the model may be necessary to adapt to new deepfake generation techniques and enhance its detection accuracy.

Acknowledgements

We extend our heartfelt appreciation to the esteemed faculty and diligent staff of Aspen University for their unwavering support throughout this research endeavor. Their dedication and expertise have been invaluable in shaping the outcome of this study. Furthermore, we would like to express our deepest gratitude to God almighty for His boundless blessings and guidance. His divine presence has been a source of strength and inspiration, guiding us every step of the way.

We would also like to acknowledge the valuable insights and feedback received from our colleagues and mentors, whose contributions have enriched our work. Finally, we would like to extend our gratitude to our families and friends for their unwavering encouragement and support, which has been instrumental in our academic journey. Thank you all for your unwavering support and con-

tributions towards the completion of this research project.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Afchar, D., Nozick, V., Yamagishi, J. and Echizen, I. (2018) MesoNet: A Compact Facial Video Forgery Detection Network. 2018 *IEEE International Workshop on Information Forensics and Security* (*WIFS*), Hong Kong, 11-13 December 2018, 1-7. https://doi.org/10.1109/WIFS.2018.8630761

[2] Ajder, H., Patrini, G., Cavalli, F. and Cullen, L. (2019) The State of Deepfakes: Landscape, Threats, and Impact. https://docslib.org/doc/12559428/the-state-of-deepfakes-landscape-threats-and-impact-henry-ajder-giorgio-patrini-francesco-cavalli-and-laurence-cullen-september-2019

[3] Maras, M. and Alexandrou, A. (2018) Determining Authenticity of Video Evidence in the Age of Artificial Intelligence and in the Wake of Deepfake Videos. *The International Journal of Evidence & Proof*, **23**, 255-262. https://doi.org/10.1177/1365712718807226

[4] Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales A. and Ortega-Garcia, J. (2020) DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection. *Information Fusion*, **64**, 131-148. https://doi.org/10.1016/j.inffus.2020.06.014

[5] Chintha, A. (2020) Employing Optical Flow on Convolutional Recurrent Structures for Deepfake Detection. Master's Thesis, Rochester Institute of Technology, Rochester. https://scholarworks.rit.edu/theses

[6] Alkazhami, A. (2020) Facial Identity Embeddings for Deepfake Detection in Videos. Master's Thesis, Linköping University, Linköping.

[7] Matern, F., Riess, C. and Stamminger, M. (2019) Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations. 2019 *IEEE Winter Applications of Computer Vision Workshops* (*WACVW*), Waikoloa, 7-11 January 2019, 83-92. https://doi.org/10.1109/WACVW.2019.00020

[8] Du, R.H., Liang, L.I. and Yang, C.Q. (2020) Predictors of Mortality for Patients with COVID-19 Pneumonia Caused by SARS-CoV-2: A Prospective Cohort Study. *European Respiratory Journal*, **55**, Article ID: 2000524. https://doi.org/10.1183/13993003.00524-2020

[9] Chen, H.S., Rouhsedaghat, M., Ghani, H., Hu, S.W., You, S.Y. and Kuo, C.C.J. (2021) Defakehop: A Light-Weight High-Performance Deepfake Detector. 2021 *IEEE International Conference on Multimedia and Expo* (*ICME*), Shenzhen, 5-9 July 2021, 1-6.

[10] Hussain, S., Neekhara, P., Jere, M., Koushanfar, F. and McAuley, J. (2020) Adversarial Deepfakes: Evaluating Vulnerability of Deepfake Detectors to Adversarial Examples. 2021 *IEEE Winter Conference on Applications of Computer Vision* (*WACV*), Waikoloa, 3-8 January 2021, 3347-3356. https://arxiv.org/abs/2002.12749 https://doi.org/10.1109/WACV48630.2021.00339

[11] Hulzebosch, N., Ibrahimi, S. and Worring, M. (2020) Detecting CNN-Generated Facial Images in Real-World Scenarios. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Seattle, 14-19 June 2020, 2729-2738.

https://doi.org/10.1109/CVPRW50498.2020.00329

[12] Nguyen, H.H., Yamagishi, J. and Echizen, I. (2019) Capsule-Forensics: Using Capsule Networks to Detect Forged Images and Videos. 2019 *IEEE International Conference on Acoustics*, *Speech and Signal Processing* (*ICASSP*), Brighton, 12-17 May 2019, 2307-2311. https://doi.org/10.1109/ICASSP.2019.8682602

[13] Sabir, E., Cheng, J., Jaiswal, A., AbdAlmageed, W., Masi, I. and Natarajan, P. (2019) Recurrent Convolutional Strategies for Face Manipulation Detection in Videos. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Long Beach, 16-17 June 2019, 80-87.

[14] Güera, D. and Delph, J. (2018) Deepfake Video Detection Using Recurrent Neural Networks. 2018 15*th IEEE International Conference on Advanced Video and Signal Based Surveillance* (*AVSS*), Auckland, 27-30 November 2018, 1-6. https://doi.org/10.1109/AVSS.2018.8639163

[15] Zhao, H., Zhou, W., Chen, D., Wei, T., Zhang, W. and Yu, N. (2021) Multi-Attentional Deepfake Detection. *IEEE*/*CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 20-25 June 2021, 2185-2194. https://doi.org/10.1109/CVPR46437.2021.00222

[16] Masi, I., Killekar, A., Mascarenhas, R.M., Gurudatt, S.P. and AbdAlmageed, W. (2020) Two-Branch Recurrent Network for Isolating Deepfakes in Videos. In: Vedaldi, A., Bischof, H., Brox, T. and Frahm, J.M., Eds., *ECCV* 2020: *Computer Vision—ECCV* 2020, Springer, Cham, 667-684. https://doi.org/10.1007/978-3-030-58571-6_39

[17] Ruff, C.B., Burgess, M.L. and Squyres, N. (2018) Lower Limb Articular Scaling and Body Mass Estimation in Pliocene and Pleistocene Hominins. *Journal of Human Evolution*, **115**, 85-111. https://doi.org/10.1016/j.jhevol.2017.10.014

[18] Nguyen, T., Nguyen, Q., Nguyen, C., Nguyen, D. and Nahavandi, S. (2021) Deep Learning for Deepfakes Creation and Detection. *Computer Vision and Image Understanding*, **223**, Article ID: 103525. https://doi.org/10.1016/j.cviu.2022.103525

[19] McCloskey, S. and Albright, M. (2018) Detecting GAN-Generated Imagery Using Color Cues. arXiv: 1812.08247.

[20] Cozzolino, D. and Verdoliva, L. (2019) Noiseprint: ACNN-Based Camera Model Fingerprint. *IEEE Transactions on Information Forensics and Security*, **15**, 144-159. https://doi.org/10.1109/TIFS.2019.2916364

[21] Li, Y., Chang, M.C. and Lyu, S. (2018) In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking. 2018 *IEEE International Workshop on Information Forensics and Security* (*WIFS*), Hong Kong, 11-13 December 2018, 1-7. https://doi.org/10.1109/WIFS.2018.8630787

[22] Mazaheri, G. and Roy-Chowdhury, A. (2022) Detection and Localization of Facial Expression Manipulations. 2022 *IEEE*/*CVF Winter Conference on Applications of Computer Vision* (*WACV*), Waikoloa, 3-8 January 2022, 2773-2783. https://doi.org/10.1109/WACV51458.2022.00283

[23] Zhao, Y., Ban, L. and Lang, H. (2019) Capturing the Persistence of Facial Expression Features for Deepfake Video Detection. In: Zhou, J., Luo, X., Shen, Q. and Xu, Z., Eds., *ICICS* 2019: *Information and Communications Security*, Springer, Cham, 630-645. https://doi.org/10.1007/978-3-030-41579-2_37

[24] Greenland, S., Senn, S.J. and Rothman, K.J. (2016) Statistical Tests, P Values, Confidence Intervals, and Power: A Guide to Misinterpretations. *European Journal of Epidemiology*, **31**, 337-350. https://doi.org/10.1007/s10654-016-0149-3

[25] Dolhansky, B., Bitton, J., Pflaum, B., Lu, J., Howes, R., Wang, M. and Ferrer, C. (2020)

The Deepfake Detection Challenge Dataset. arXiv: 2006.07397.

[26] Das, S., Zhu, Y. and Jain, A. (2006) Validating a Biometric Authentication System: Sample Size Requirements. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**, 1902-1319. https://doi.org/10.1109/TPAMI.2006.255