Swapped Face Detection: AI-Based Method and Evaluation for Different Face Swap Algorithms

Mikhail Haleev
SPC RAS
St. Petersburg, Russia
haleev.m@iias.spb.su

Abstract— **Deepfake images have become a major problem in today's digital landscape. Such images are usually created using advanced machine learning techniques. These fake images can deceive viewers, posing risks to privacy, security, and trust. In this paper we introduce an innovative approach to detect deepfake by analyzing facial landmarks and computing corresponding expert features, train multiple classifier models based on three sets of features and achieve an accuracy of 0.682 and F1-score of 0.680. Using the coefficients of the resulting models, we evaluate the importance of features and identify the most important ones.**

I. Introduction

Deepfakes represent a sophisticated form of synthetic media, where advanced algorithms, particularly those involving deep learning and neural networks, are employed to alter or entirely replace segments of an original image or video. These alterations can be so meticulously crafted that the resulting content is often indistinguishable from authentic media to the human eye. The proliferation of deepfakes poses significant challenges, as their realistic nature can be exploited to disseminate misinformation, manipulate public opinion, and cause various forms of harm. Consequently, the development and implementation of robust detection mechanisms are imperative to safeguard against the potential misuse of this technology and to maintain the integrity of information in the digital age.

In papers [1] and [2] say that deepfakes pose significant threats, including fraud through deceptive videos, evidence tampering in legal proceedings, misuse of personal data, and the spread of disinformation to influence public opinion and political processes. These dangers necessitate the development of advanced detection and prevention technologies by law enforcement agencies.

Our proposed method centers on the analysis of facial landmarks, which are crucial in differentiating authentic content from deepfakes. Facial landmarks refer to specific, identifiable points on a face, including the corners of the eyes, the tip of the nose, and the edges of the mouth. These landmarks are instrumental in conveying vital information about facial expressions, geometric structure, and individual identity. By meticulously evaluating the positional and relational significance of these points, our method can detect subtle anomalies and inconsistencies introduced by deepfake manipulations. This approach enhances the accuracy and reliability of deepfake detection, thereby contributing to the broader effort to combat misinformation and protect the integrity of visual media.

We also identified expert features that are most significant for identifying deepfake images, which we calculated based on facial landmarks. Our approach focuses on analyzing facial landmarks and expert features to enhance deepfake detection accuracy.

We compute the following expert features based on these landmarks:

1) *Geometric Distances.* By measuring distances between key facial points, we identify unnatural alterations introduced by deepfake manipulation. For instance, discrepancies in eye-to-eye distance or mouth-to-nose distance can signal potential tampering.
2) *Facial Volume and Area.* Calculating the volume and area enclosed by facial landmarks provides additional insights. Deviations from expected values may indicate deepfake modifications.

We proposed a classifier to be used for deepfake detection and subsequent feature importance analysis. The classifier was trained separately on each set of features and based on their weights, the importance of features within the set was calculated.

1) *Importance of Landmarks.* Assigning weights to each facial point based on its importance allows us to prioritize specific regions. For example, the corners of the mouth and eyes are more critical for accurate detection.
2) *Importance of Expert Features.* Assessing the importance of expert features will allow us to evaluate the key features for detecting deepfakes to use only them in the future.

In the subsequent sections of this paper, we first conduct a review of the current state of research, which serves as the foundational basis for our study. We then provide an in-depth description of our proposed method, detailing each step of the process. This method is employed to generate a robust data set, which is subsequently used to train our classifier model. We calculate and present quality metrics to evaluate the performance of the model. Following this, we assess the significance of various features within the data set, drawing conclusions about their impact on the model's accuracy. Finally, we summarize our findings, discuss the implications of our research, and outline potential directions for further development and refinement of our approach.

II. Related Works

In this section, we review relevant literature on both face swap techniques and deepfake detection methods.

A. Face Swap Detection

In [3], the authors categorize deepfake images into four subgroups: face generation, face replacement, face modification, and expression alteration. Each subgroup corresponds to specific datasets, consisting of images generated solely by GAN architectures. The dataset sizes range from 100,000 to 330,000 images. Notably, emphasis is placed on face replacement in video data.

Work [4] proposes a unified architecture for detecting all types of deepfake images. The authors provide a concise analysis and comparison of existing models tailored to specific deepfake types. Additionally, they introduce a unified architecture achieving an impressive accuracy of 98% using the publicly available "FaceForensic++" dataset, which includes 1,000 genuine and 4,000 manipulated video clips.

Article [5] specializes in detecting deepfake videos. The researchers utilize deep convolutional neural networks (DCNNs) to compute facial similarity scores. The model achieves an accuracy of 97.5% on open datasets.

The authors of [6] focus on detecting face-swapped images. They present a dataset and conduct a comparative analysis of various architectures. Additionally, the study provides the source code for their approach.

B. DeepFake Detection

Method [7] employs joint unsupervised reconstruction and supervised classification for deepfake detection. Unlike previous approaches, it pays attention not only to explicit neural network errors but also to less obvious signs of overall generation. The method was tested on five different datasets (UADFV, FaceForensics++, Celeb-DF (version 2), DeepFake Detection (DFD), and DeepfakeTIMIT) and outperformed existing solutions on one of the datasets.

Authors of [8] propose model which has two basic modules: GraphNet, which uses graph convolution layers to aggregate and update graph information, and FFN, which has linear layers for the transformation of node features. The effectiveness of the method is assessed using the diverse Deepfake Detection Challenge dataset (DFDC), FaceForensics++ (FF++), World Leaders dataset (WLRD), and the Celeb-DF and achieved up to 76% accuracy.

Leveraging features extracted from three Vision Transformers (DaViT, iFormer, and GPViT), method [9] analyzes video data. The dataset consists of 9 real and 51 fake images. It was tested on FaceForensics++ using four different techniques and achieved accuracy up to 97.72%.

Authors of [10] propose a framework, *ART-AVDF*, that utilizes articulatory representation for accurate audio–visual deepfake detection. In *ART* module, an auditory encoder and a lip encoder are designed to perform audio–visual articulatory representation learning using the self-supervised learning strategy. In *AVDF* module, they utilize the frozen encoders in *ART* module to obtain articulatory embeddings and fuse them with unimodal features, leading to better audio–visual analysis for deepfake detection. They use DFDC and FakeAVCele, DefakeAVMiT and achieve accuracy up to 96%.

Approach [11] breaks down facial images into fundamental elements and computes characteristics based on them such as face texture and naturalness degree. The recognition model used was HDDM, which consists of two models, face texture construction (FTC) and naturalness degree recognition (NDR). It was tested on various datasets (Flickr-Faces-HQ, StyleGan2, FaceForensics++, Celeb-DFv2) and achieved accuracy up to 99.7%.

Using ResNet-Swish-BiLSTM, method [12] labels fake videos. Patterns of identified faces are sized to 224×224 dimensions. It was tested on DFDC and FF++ datasets, achieving 99.13% accuracy. A cross-validation algorithm was also used to evaluate the effectiveness of the method.

Method [13] detects deepfake photos and videos by identifying imperceptible artifacts. Authors of the paper obtain two different types of features, one represents global inconsistencies among the masked patches that help generate intra-inter patch information, and the other one represents spatial global features of the image. Model was tested on FaceForensics++, Celeb-DF (V2), and Deepfakes Image dataset (DFID) with an accuracy of 95.42%.

Authors of [14] propose network consisting of two parts (auxiliary and backbone networks). Authors focus on the inconsistency of illumination between frames. To enhance the illumination inconsistency at feature level, multi-level feature enhancement which consists of FRRC and feature fusion blocks is proposed to recompose features at different levels. Model was tested on FaceForensics++, Celeb-DF and DFDC and achieve accuracy up to 99.11%.

Authors of [15] have proposed a stacking-based ensemble method, where features generated by dual CNN models are stacked followed by the selection of optimal features and elimination of inconsistent features. Method has been tested and validated using FaceForensics++ and Celeb-DF. Two models were used for feature extraction: Xception and EfficientNet-B7. XG Boost Regressor and Random Forest classifier were used as the main models. Method has achieved accuracy up to 98%.

Method [16] detects deepfakes with using Convolutional Attention Neural Network. Authors of the paper have recognized the significance of the Fourier Transform as a frequency domain representation and encoded facial videos into Matrix Visualization Heatmap (MVHM) for input into image classification networks. The authors found that the spatial attention mechanism increased the performance of the VGG19 network by 9.38 percentage points. Method was tested on DeepFakeTIMIT and achieves an accuracy of 99.2%.

Authors of [17] propose Texture and Artifact Detector (TAD) for deepfake detection, which aims to separate mutually exclusive texture inconsistencies and artifact information, thereby weakening their mutual influence and improving the model's generalization ability. The authors have used FaceForensics++, WildDF, DFDC, CelebDF as datasets and have chosen a set of diverse baseline methods (XceptionNet, EfficientNetB4, EfficientNetB4Att, DSP-FWA, MCX-API) for comparison with the proposed method. Method achieves up to 99% accuracy.

Authors of [18] integrated DeepFake detection with 3D gaze estimation. This integration endows model with the ability to discern forgery videos by distinguishing spatial inconsistencies within eye regions from a gaze perspective between given frames. They proposed a biometric feature integration strategy by introducing Mean Square Error (MSE) and leaky features fusion to regularize our DeepFake detection model. They used FaceForensics++, Celeb-DF and WildDeepfake datasets and achieved accuracy up to 86%.

The multi-scale fusion (MSF) module designed in [19] can obtain forged facial features, and the interactive dual-stream (IDS) module can better integrate feature information in the frequency and spatial domains. To address the problems of low-quality datasets and poor detection performance across datasets, this study proposes multi-scale interactive dual-stream network. Method was tested on FaceFornesics++ and Celeb-DF and achieved accuracy up to 99%.

The paper [20] offers end-to-end transformer-based spatio-temporal model, SFormer, utilizing the transformer architecture to discover global links between several local locations of frame, and temporal relationships among spatial features of consecutive frames. SFormer utilizes a Swin transformer to extract spatial features followed by a transformer block for temporal analysis. Tests are carried out on five face manipulation benchmark datasets named DFD, Celeb-DF, FF++, DFDC and Deeper-Forensics while obtaining an accuracy of 97.81%, 99.1%, 100%, 93.67% and 99.67% respectively.

Authors of [21] propose a Spatial-Frequency Fusion Branch (SFFB), the framework of which is simple and easy to implement. In the training process, they utilize spatial features, frequency domain features and logits for multi-knowledge transfer. They use FaceForensics++ and Celeb-Deepfake and achieve accuracy up to 97%.

Authors of [22] propose a High-Frequency Enhancement (HiFE) network to handle low-quality data. The adaptive frequency-aware features from local sub-networks adopt Block-wise DCT, channel attention mechanism, channel bottleneck module, and inverse Block-wise DCT. Moreover, Multi-level DWT decomposition layers and cascade-residual-based multi-level fusion strategies are designed to realize adaptive global high-frequency enhancement. Method was tested on FaceForensics++, Celeb-DF and OpenForensics and achieved accuracy up to 99%.

C. Conclusion

While most existing methods focus on detecting deepfakes in video data, our work stands out by addressing face replacement exclusively in images. We propose a novel evaluation method for assessing existing face swap algorithms, highlighting their effectiveness and limitations. Our method not only detects deepfakes, but also focuses on the features that served as the basis for detecting deepfakes.

III. METHODOLOGY

We propose to delve into the complex task of understanding feature importance for face classification (Fig. 1). Our method not only detects deepfakes, but also calculates the importance of features to evaluate the algorithm. Here's a description of each step.

We propose to generate Human Face Dataset module from several existing datasets (e.g., CelebA, Labeled Faces in the Wild) containing face images. The dataset contains images of real people (not deepfake generated images). We divide the dataset to base, secondary and original faces for swapping. We select a subset of faces for base faces and secondary one in the proportion *1:1*. Faces from first group act as the canvas onto which features from second group will be swapped. Original faces represent the unaltered state and take part in the forming of the result dataset.

We propose to use several different Swapped Face Generators so that each of them processes the same part of the faces from the first two sets. The result of the generators is Swapped Faces set. The original faces together with the replaced ones are taken in equal proportions and form New Face Swap Dataset. This dataset is used for further analysis and formation of a set of features.

Google Media Pipe Face Mesh is a solution that estimates 3D facial landmarks. It uses machine learning (ML) to infer the 3D surface of a face, requiring only an image without the need for depth mapping. We used it to capture facial points to form a set of facial landmarks and then calculate characteristics.

It was hypothesized that the most relevant features for detecting deepfake images would be those related to both the shape of the face and the ratios of individual distances on the face. The characteristics of the face shape were taken to be area and volume. Also considered were such features as the basic proportions of the facial parts (eyes, mouth and the entire face). To assess the position of parts of the face, it was customary to use ratios that would compare the horizontal dimensions of the face with the vertical ones. For this, the following distances were taken: distance from eye to ear divided by the distance from eye to nose from the right and left sides of the face, distance from eye to eye divided by the distance from eye to nose from the left and right sides of the face, distance from ear to eye divided by the distance from ear to mouth from the left and right sides of the face.

The described features were calculated for each image from the dataset, after which they were divided into three sets of features: facial landmarks, expert features, and a combined set of features. The resulting sets of features determine the final data on which the classifier models will be trained.

After that, we train three models, each on its own set of features. We chose a perceptron as a classifier model due to its learning speed and the ability to obtain model coefficients for subsequent feature importance analysis. The perceptron is a basic algorithm in machine learning used for binary classification. It consists of input features, weights, a summation function, an activation function, and a bias. The perceptron calculates a weighted sum of the inputs, applies an activation function, and outputs a binary result. It learns by adjusting weights and bias based on prediction errors. Despite its simplicity, the perceptron is foundational in the development of more complex neural networks and deep learning models (see Fig. 2).

We calculate accuracy metrics and analyze the model coefficients. Using coefficient normalization, we obtain the importance for each feature. Based on the most important features from the obtained set, a new classifier model is trained.

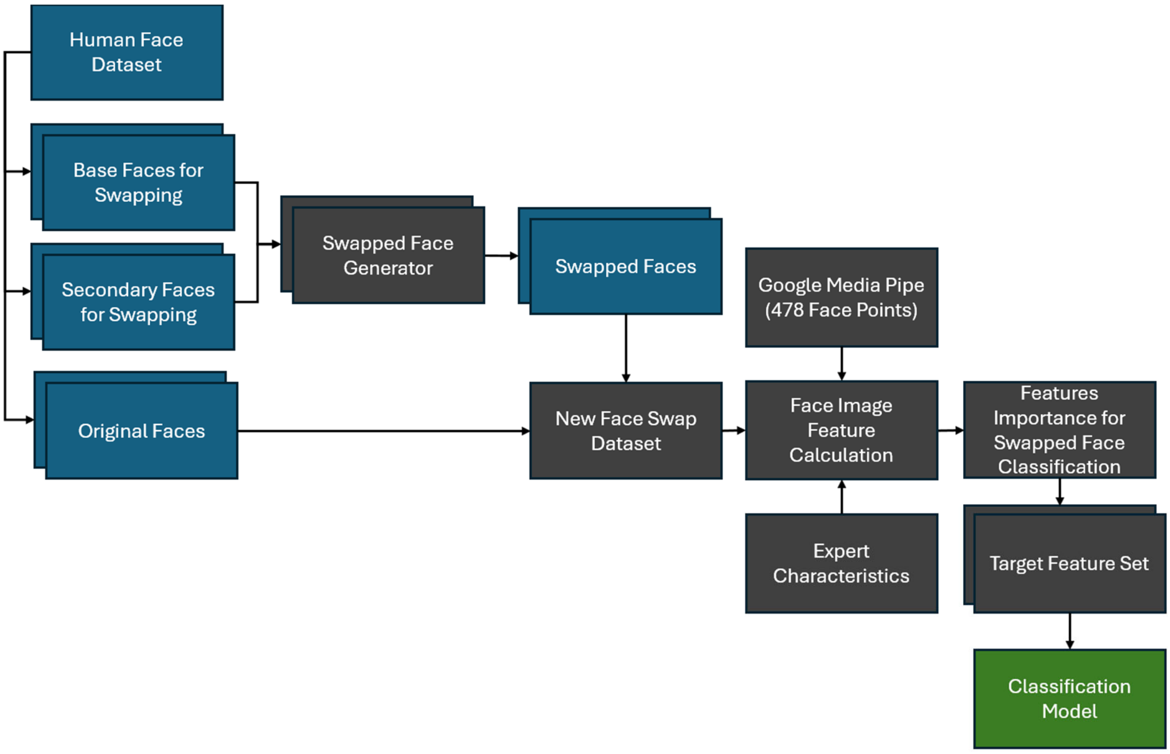

Fig. 1. Approach to identifying feature importance

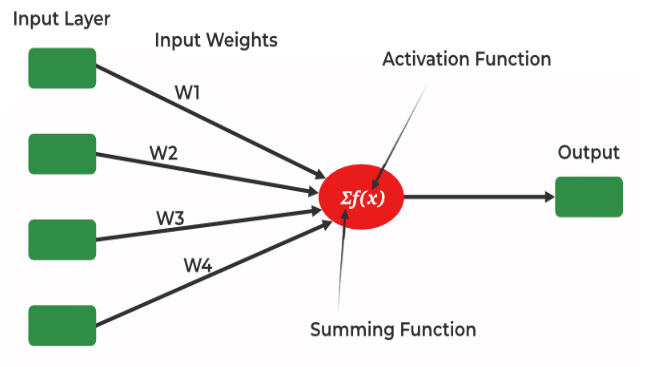

Fig. 2. Perceptron Architecture [23]

IV. RESULTS

This section presents the results of the study such as feature importance, trained models and experiment results. The CelebA dataset was used as the face dataset. A total of 35,100 images were used to form the described image sets.

A. *Train Classification Model*

We traing a classification model in according to methodology presented in Section 3. We train three different models based on input data: exper features (see Section 4B), facial landmarks (see Section 4C), and both of them (combined feature set). As a testing sample, we took 30% of the entire sample (7800 out of 23400) the rest data we used for training. As variable hyperparameters, we used the maximum number of iterations, *eta0* (a constant by which updates are multiplied). Another hyperparameter is whether early stopping of the model will be used during training if the accuracy does not improve for a long time. Different types of regularization (*l2*, *l1*, elasticnet or no regularization) with different coefficients were also used. To identify the most suitable set of hyperparameters, studies were conducted with each set and the best one was taken as the final result.

Models trained on the first two sets have low accuracy, and the first model had no difference in the accuracy of identifying original and deepfake images (Fig. 3), while the second model identified original images more accurately than deepfake (Fig. 4) but still with low accuracy. Based on this, we can judge that the trend is evident, but the accuracy needs to be improved. The third model has greater accuracy than its predecessors and performs equally well on both original images and deepfakes (Fig. 5) and achieved 68.2%.

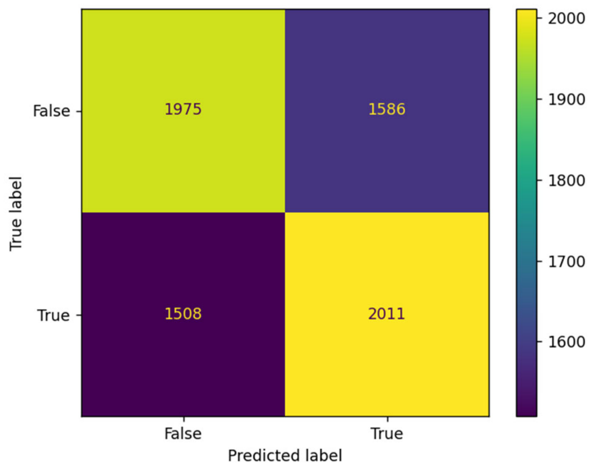

Fig. 3. Confusion matrix for expert features set model

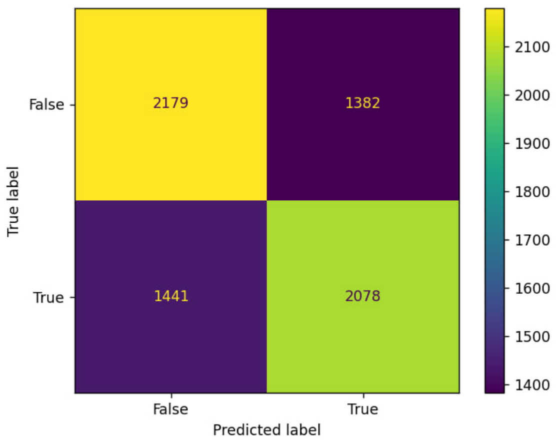

Fig. 4. Confusion matrix for facial landmark features set model

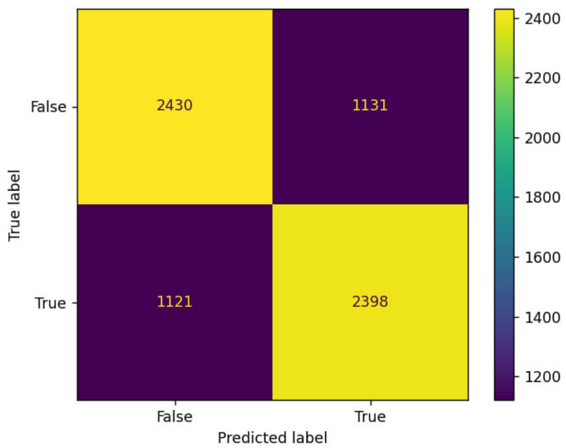

Fig. 5. Confusion matrix for combined features set model

Based on the obtained results (Table 1), it can be judged that the model predicted better on facial points than on expert features, and when combining two sets of features, it showed better results than on each of the sets separately.

TABLE I. METRICS

Feature set	Accuracy	F1-score
Facial landmarks	0.601	0.596
Expert features	0.563	0.565
Combined	0.682	0.680

B. Expert Feature Importance Analysis

Feature importance was obtained as the coefficients of the classification model. Expert feature importance is shown in the Table 2. We see that the most significant features were those related to the shape of the face (volume and face area). Then it is four characteristics that have significant relationship with face swapping: (1) distance from eye to ear divided to distance from eye to nose from right side of the face; (2) face proportions (width divided height); (3) distance from eye to ear divided to distance from eye to nose from left side of the face; (4) distance from eye to eye divided to distance from eye to nose from left side of the face. And the rest characteristics have unsignificant relationship with face swapping.

TABLE II. EXPERT FEATURE IMPORTANCE

#	Feature	Importance
1	Face volume	-0,796
2	Face Area	0,414
3	Distance from eye to ear divided to distance from eye to nose from right side of the face	-0,258
4	Face proportions (width divided height)	-0,224
5	Distance from eye to ear divided to distance from eye to nose from left side of the face	-0,221
6	Distance from eye to eye divided to distance from eye to nose from left side of the face	0,200
7	Left eye proportions (width divided height)	-0,168
8	Distance from ear to eye divided to distance from ear to mouth from left side if the face	0,143
9	Right eye proportions (width divided height)	-0,120
10	Distance from ear to eye divided to distance from ear to mouth from right side if the face	0,101
11	Mouth proportions (width divided height)	0,099
12	Distance from eye to eye divided to distance from eye to nose from right side of the face	0,041

C. Facial Landmarks Importance

Facial landmarks importance has been calculated using coefficients of perceptron model. For visualization, it was decided to normalize the model coefficients, and then compare them with the colors of facial landmarks. The final color of the facial landmark in RGB was calculated as (k * 255, (1 - k) * 255, 0), as a result of which the value of each color was within 255. We show results in Fig. 6. To visualize the importance of facial points, facial points were colored proportionally to their importance. Green symbolized low importance, red symbolized high one. Based on the results obtained we concluded that the classifier model paid attention primarily to the corners of the lips and the contour of the mouth, the area around the eyes and the edges of the face.

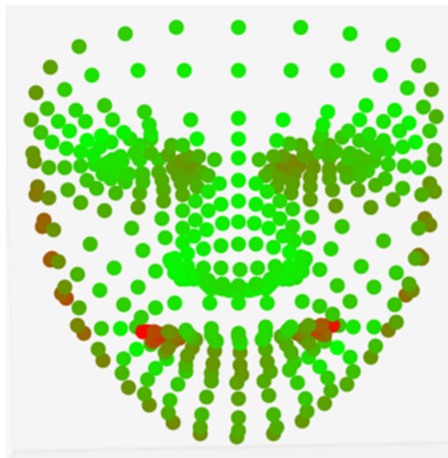

Fig. 6. Facial landmarks importance (green – low importance, red – high importance)

V. CONCLUSION

We analyzed the current state of the research and, based on the findings and expert assessment, formed three approaches to extracting features in images to detect deepfakes. Based on these sets, we trained three classifier models and calculated their quality metrics and visualized confusion matrices.

We analyzed the coefficients of the resulting model and, based on them, calculated the importance of each feature. For the set of facial landmarks, the obtained importances were visualized using the example of a face. We identified the most significant parameters obtained by expert assessment and determined the most significant facial regions for detecting deepfakes.

We found the best classifier achieving an accuracy of 0.682. This accuracy is too low to apply the model to real-world problems, but it demonstrates that even a basic model without complex architecture and selection of additional parameters reveals a trend based on the proposed method. In the future, to increase accuracy, we will use different datasets as image bases and different generators to generate more diversity in the dataset. We also plan to train a more complex model architecture on the current set of features for a more in-depth approach to detecting deepfakes.

REFERENCES

[1] "Europol Report Criminal Use of Deepfake Technology - eucrim." Accessed: Aug. 30, 2024. [Online]. Available: https://eucrim.eu/news/europol-report-criminal-use-of-deepfake-technology/

[2] "Facing reality? Law enforcement and the challenge of deepfakes An Observatory Report from the Europol Innovation Lab", doi: 10.2813/158794.

[3] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, "Deepfakes and beyond: A Survey of face manipulation and fake detection," *Information Fusion*, vol. 64, pp. 131–148, Dec. 2020, doi: 10.1016/J.INFFUS.2020.06.014.

[4] M. Nawaz, A. Javed, and A. Irtaza, "A deep learning model for FaceSwap and face-reenactment deepfakes detection," *Appl Soft Comput*, vol. 162, p. 111854, Sep. 2024, doi: 10.1016/J.ASOC.2024.111854.

[5] P. M. G. I. Reis and R. O. Ribeiro, "A forensic evaluation method for DeepFake detection using DCNN-based facial similarity scores," *Forensic Sci Int*, vol. 358, p. 111747, May 2024, doi: 10.1016/J.FORSCIINT.2023.111747.

[6] X. Ding, Z. Raziei, E. C. Larson, E. V. Olinick, P. Krueger, and M. Hahsler, "Swapped face detection using deep learning and subjective assessment," *EURASIP J Inf Secur*, vol. 2020, no. 1, Dec. 2020, doi: 10.1186/S13635-020-00109-8.

[7] B. Yan, C. T. Li, and X. Lu, "JRC: Deepfake detection via joint reconstruction and classification," *Neurocomputing*, vol. 598, p. 127862, Sep. 2024, doi: 10.1016/J.NEUCOM.2024.127862.

[8] F. Khalid, A. Javed, Q. ul ain, H. Ilyas, and A. Irtaza, "DFGNN: An interpretable and generalized graph neural network for deepfakes detection," *Expert Syst Appl*, vol. 222, p. 119843, Jul. 2023, doi: 10.1016/J.ESWA.2023.119843.

[9] E. Essa, "Feature fusion Vision Transformers using MLP-Mixer for enhanced deepfake detection," *Neurocomputing*, vol. 598, p. 128128, Sep. 2024, doi: 10.1016/J.NEUCOM.2024.128128.

[10] Y. Wang and H. Huang, "Audio–visual deepfake detection using articulatory representation learning," *Computer Vision and Image Understanding*, vol. 248, p. 104133, Nov. 2024, doi: 10.1016/J.CVIU.2024.104133.

[11] C. J. Chew, Y. C. Lin, Y. C. Chen, Y. Y. Fan, and J. S. Lee, "Preserving manipulated and synthetic Deepfake detection through face texture naturalness," *Journal of Information Security and Applications*, vol. 83, p. 103798, Jun. 2024, doi: 10.1016/J.JISA.2024.103798.

[12] A. Qadir, R. Mahum, M. A. El-Meligy, A. E. Ragab, A. AlSalman, and M. Awais, "An efficient deepfake video detection using robust deep learning," *Heliyon*, vol. 10, no. 5, p. e25757, Mar. 2024, doi: 10.1016/J.HELIYON.2024.E25757.

[13] S. Ganguly, A. Ganguly, S. Mohiuddin, S. Malakar, and R. Sarkar, "ViXNet: Vision Transformer with Xception Network for deepfakes based video and image forgery detection," *Expert Syst Appl*, vol. 210, p. 118423, Dec. 2022, doi: 10.1016/J.ESWA.2022.118423.

[14] C. Zhu, B. Zhang, Q. Yin, C. Yin, and W. Lu, "Deepfake detection via inter-frame inconsistency recomposition and enhancement," *Pattern Recognit*, vol. 147, p. 110077, Mar. 2024, doi: 10.1016/J.PATCOG.2023.110077.

[15] G. Naskar, S. Mohiuddin, S. Malakar, E. Cuevas, and R. Sarkar, "Deepfake detection using deep feature stacking and meta-learning," *Heliyon*, vol. 10, no. 4, p. e25933, Feb. 2024, doi: 10.1016/J.HELIYON.2024.E25933/ASSET/F9A19FF4-05CB-421C-AF36-E3240D24EA06/MAIN.ASSETS/GR011.JPG.

[16] Y. Ni, W. Zeng, P. Xia, G. S. Yang, and R. Tan, "A Deepfake Detection Algorithm Based on Fourier Transform of Biological Signal," *Computers, Materials & Continua*, vol. 79, no. 3, pp. 5295–5312, Jun. 2024, doi: 10.32604/CMC.2024.049911.

[17] J. Gao *et al.*, "Texture and artifact decomposition for improving generalization in deep-learning-based deepfake detection," *Eng Appl Artif Intell*, vol. 133, p. 108450, Jul. 2024, doi: 10.1016/J.ENGAPPAI.2024.108450.

[18] Q. He, C. Peng, D. Liu, N. Wang, and X. Gao, "GazeForensics: DeepFake detection via gaze-guided spatial inconsistency learning," *Neural Networks*, vol. 180, p. 106636, Dec. 2024, doi: 10.1016/J.NEUNET.2024.106636.

[19] Z. Cheng, Y. Wang, Y. Wan, and C. Jiang, "DeepFake detection method based on multi-scale interactive dual-stream network," *J Vis Commun Image Represent*, vol. 104, p. 104263, Oct. 2024, doi: 10.1016/J.JVCIR.2024.104263.

[20] S. Kingra, N. Aggarwal, and N. Kaur, "SFormer: An end-to-end spatio-temporal transformer architecture for deepfake detection," *Forensic Science International: Digital Investigation*, vol. 51, p. 301817, Dec. 2024, doi: 10.1016/J.FSIDI.2024.301817.

[21] B. Wang, X. Wu, F. Wang, Y. Zhang, F. Wei, and Z. Song, "Spatial-frequency feature fusion based deepfake detection through knowledge distillation," *Eng Appl Artif Intell*, vol. 133, p. 108341, Jul. 2024, doi: 10.1016/J.ENGAPPAI.2024.108341.

[22] J. Gao, Z. Xia, G. L. Marcialis, C. Dang, J. Dai, and X. Feng, "DeepFake detection based on high-frequency enhancement network for highly compressed content," *Expert Syst Appl*, vol. 249, p. 123732, Sep. 2024, doi: 10.1016/J.ESWA.2024.123732.

[23] "Perceptron Algorithm for Classification using Sklearn - GeeksforGeeks." Accessed: Sep. 04, 2024. [Online]. Available: https://www.geeksforgeeks.org/sklearn-classification-using-perceptron/?ysclid=m0novhnsyv693925398