Applied Nanoscience
https://doi.org/10.1007/s13204-021-02072-3

ORIGINAL ARTICLE

Deepfake detection using rationale-augmented convolutional neural network

Saadaldeen Rashid Ahmed Ahmed[1] · Emrullah Sonuç[1]

Received: 11 July 2021 / Accepted: 28 August 2021

Abstract
Deepfake network is a prominent topic of research as an application to various systems about security measures. Although there have been many recent advancements in facial reconstruction, the greatest challenge to overcome has been the means of finding an efficient and quick way to compute facial similarities or matches. This work is created utilizing the rationale-augmented convolutional neural network (CNN) on MATLAB R2019a platform using the Kaggle DeepFake Video dataset with an accuracy of 95.77%. Hence, real-time deepfake facial reconstruction for security purposes is difficult to complete concerning limited hardware and efficiency. This research paper looks into rational augmented CNN state-of-the-art technology utilized for deepfake facial reconstruction via hardware such as webcams and security cameras in real time. Additionally, discuss a history of face reconstruction and provide an overview of how it is accomplished.

Keywords Deepfake · Video · Detection · Segmentation · Facial alignment · Deep learning · Reconstruction

Introduction

There are many uses for facial reconstruction, namely, in video surveillance and aiding in the search of specific criminals from available cameras in public. Additionally, it can also be applied as a safety/security measure for high authority officials to enter restricted premises. As such, it is an extremely active field in research by many data scientists and DL enthusiasts. Given that there exists a surplus of data from the proliferation of social media in recent years, powerful machine learning and statistical models could be built to an alarmingly high accuracy rate. This technology can even be seen in Apple's or Facebook's photo face reconstruction feature when searching for a name; it is commonly asked to tag a specific individual after detecting a face in a photo. This project aims to research the potential usage of deep learning models such as convolutional neural networks to solve facial reconstruction tasks in a real-time video feed (Alhayani and Abdallah 2020; Alhayani and llhan 2021). Additionally, research will be conducted to obtain a greater or more efficient performance of this task using state of the art statistical models while training on a large public/open-source dataset (Alhayani et al. 2021; Al-Hayani and Ilhan 2020).

Research problem

The following section presents the primary and secondary questions that this paper aims to answer:

- Is it possible to classify deepfake images using a rationale-augmented convolutional neural network (CNN) and a measure of the training and validation accuracy and loss percentage?
- Does the approach of deep learning techniques automate faking the real videos for security reasons?
- To develop a system that gives absolute accuracy after evaluating the different types of videos using the CNN model?
- What should be an implementation pipeline for efficiently process the deepfake videos for streaming with a CNN-based deepfake model?
- What aspects need to be considered if further work on the subject is to be performed?

✉ Saadaldeen Rashid Ahmed Ahmed
Saadaljanabi95@gmail.com

Emrullah Sonuç
Esonuc@karabuk.edu.tr

[1] Department of Computer Engineering, Karabuk University, Karabuk, Turkey

Published online: 13 September 2021

Aim of study

Upon completing this research, we hope to have successfully experimented with available state-of-the-art deepfake network learning models to accomplish the facial reconstruction task. With this in mind, the deepfake model should ideally accurately reconstruct the face and any additionally specified faces with greater than 95% accuracy in real time. Additionally, the model should be relatively compact and produce an acceptable frame rate when run in real time. (Kwekha-Rashid et al. 2021; Yahya et al. 2021) I chose to work on this topic because we found the topic to have many use cases in the industry and personal use. Furthermore, this project will be great practice for gaining more experience in the growing field of deep learning and artificial intelligence.

- Provide an overview of previous works and achievements on the reconstruction of different types of faces.
- Apply rationale-augmented CNN methods to the unified Kaggle DeepFake Challenge dataset to perform reconstruction with less data loss and more accuracy.
- Evaluate and compare used rationale-augmented CNN, as well as compare with other studies from the literature.
- Develop a rationale-augmented CNN-based novel methodology with good results for deepfake face reconstruction with Donald Trump and Vladimir Putin.

Literature review

The several works discussed in this project will be based on the papers within the references section. According to Radford et al. (2015), in 2015, DeepFake had acquired state-of-the-art accuracy on the famous RNN benchmark and had later been dramatically improved to 87.80% in the next three years. When referring to Google's Face Net (Arjovsky and Chintala 2017), as shown in Fig. 1, it became apparent that larger datasets contribute largely to the accuracy of such models; additionally, no one other than Google had access to the 260 million images that were used to obtain its high accuracy as mentioned in Chollet (2017). As such, this inspired varying architectures that required fewer data. This project will refer to and utilize several different deep learning architectures mentioned in Mao et al. (2017a), namely, variants of Inception (v1 and v2). In addition to these models, we will also discuss the history of the progress of facial reconstruction and the several different approaches taken to complete the task.

The study of face reconstruction with deep neural networks had made significant progress with the publication of the United States of America journal paper published in 2017 (Huang et al. 2017). This paper proposed a series of convolutional neural networks that suggested translating specific landmarks into a series of 160-dimensional vectors, also known as an individual's DeepFake. These landmarks included cropped regions centered around an individual's facial features, including left eye, right eye, nose, left mouth corner, right mouth corner. The architecture had also proposed that the target output would be the individual's corresponding identities, including $n > 10{,}000$ unique individuals.

Following the development of the DeepFake model came improvements to the architecture and ultimately produced DeepFake2, DeepFake2+, and DeepFake3. Each of these models made modifications to how patches or landmarks are selected for processing and used different facial landmark detectors as mentioned in Liu et al. (2016). The DeepFake models ultimately reached a respectable performance of 89.53% accuracy on the GAN. However, they began to face problems as the number of classes began to increase.

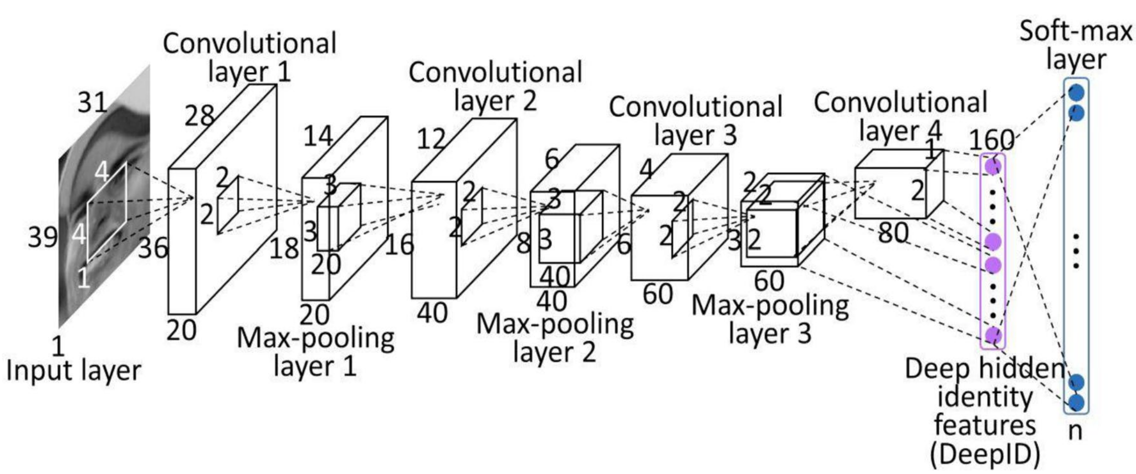

Fig. 1 Different orientations of faces being evaluated for deepfake network (Dang et al. 2018)

With a significant increase in classes and fewer face images for each class, the authors noted “the classifier outputs became diverse and unreliable, therefore cannot be used as features” as mentioned in Miyato et al. (2018). This was likely attributed to the chosen loss function of the architecture; the model had no way of formalizing facial features within similar faces but was training to optimize which class to predict within a specific set of identities (Rashid et al. 2021; Abu-Rumman 2021). As such, it could be expected that using the DeepFake for face verification outside of the ten thousand classes would be much less reliable than that of a DeepFake used for someone inside of the dataset in Russakovsky et al. (2015).

Following the publication of the DeepFake models, Facebook released DeepFake; they claimed it was a model that reached human-level accuracy of facial reconstruction. The particularly unique aspect of Face Net was the face alignment step taken to improve the model’s robustness. By performing 2-D and 3-D alignment on detected faces, it is guaranteed that every face will be analyzed in the same orientation and perspective, allowing for a much more consistent analysis of facial features as described in Huaxiao Mo and Luo (2018); Abu-Rumman et al. 2021). However, this technology is very difficult to implement, and the model required a massive dataset that only Facebook has access to train, as shown in Fig. 2. Ultimately, the model’s architecture was set to produce a 4096-d representation vector trained with cross entropy loss for predicting the probability of associated classes.

In this paper (Kaggle 2021), a deep learning algorithm helps find the catchment basins and ridgelines in the image. In this case, the ridgeline represents the height that separates the two catchment basins. This paper (Hsu et al. 2020) and (Wen et al. 2015) provide a deep learning-based method used to compute the distance from every pixel to every nonzero pixel for deepfake detection. In this paper, the algorithm is then implemented using deep learning methods and functions presented in Korshunova et al. (2017). In this paper (Zhang et al. 2017), the author returns a labeled matrix that consists of positive integer values for different regions and deepfake network for ridgelines. This image is not very useful as there is only one catch basin spanning the entire image. In this paper (Wang et al. 2017), the deepfake network catchment basins are the regions we want to identify the faces. Therefore, we complement our image and apply RNN as method on the complimented image after which we negate the distance to determine the bright catchment basins that represent individual regions. We then apply a deep learning method that will return a labeled matrix consisting of positive values and catchment basins for different faces as presented in Bai (2017). We convert the labeled matrix into an image for face detection and reconstruction to display the image and display the result using deep learning in Zheng et al. (2016). However, before taking the compliment of the image to apply deep learning algorithms, several pre-processing techniques are applied to enhance the region of interest we want to extract. In this paper (Xuan et al. 2019), the LTSM deep learning method followed by morphological erosion of disk size 3 for face reconstruction has been applied to enhance the contrast of the face region. This technique gives accurate results if the image is of reasonable contrast so that the deep learning algorithm can extract the catchment basins efficiently for deepfake network.

In these papers (Yang et al. 2016; Bayar and Stamm 2016), there are blob detection and fake face segmentation technique using deep learning that helps in extracting the faces from the images and videos. In this paper (Qian et al. 2015), the deepfake face segmentation technique is less sensitive to light changes and detects small regions, making it useful to extract the sensitive region of interest from

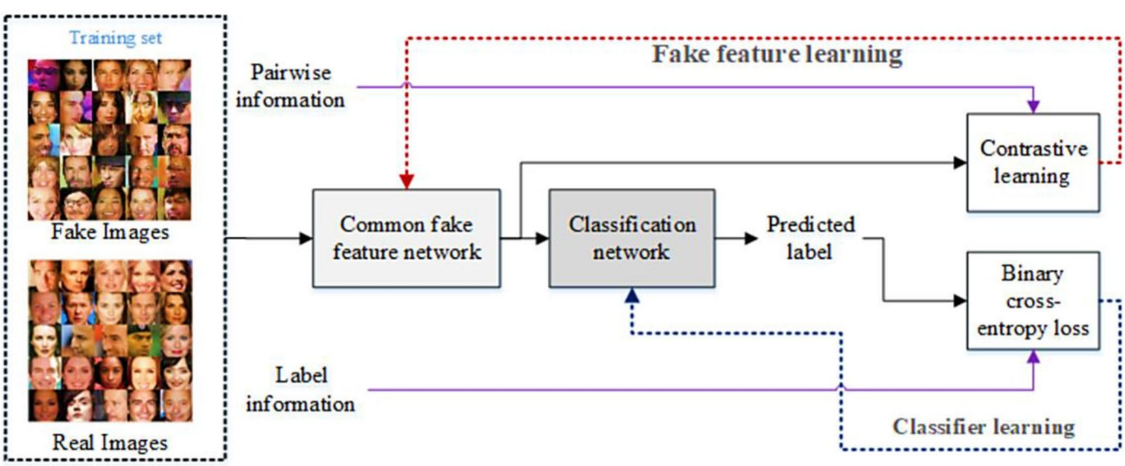

Fig. 2 An architectural diagram for the CNN with deep hidden identity (Sun et al. 2014)

an image. In the paper (Agarwal and Varshney 2019), the author uses the function for deepfake feature analysis to capture the features of the region researchers want to extract. Authors in Arjovsky et al. (2017) then use these features to find that region and plot the ellipse for deepfake network and face reconstruction. However, the face has a certain area range and threshold value passed to collect features only within the specified range. In this paper (Gulrajani et al. 2017), the face threshold value passed is 12, and the area range is 200–5000. For deepfake face threshold values of lesser than 12, more regions were detected, whereas threshold values greater than 12, expected regions sometimes do not appear. Using the threshold value 12, there are situations where more than one face in the image is not extracted. In this paper (Mao et al. 2017b), the author used deep learning because the area range specified is too large to identify the important regions or the threshold value is slightly small, which prevents it from identifying the important regions for face reconstruction using the deepfake network. Therefore, in Keren et al. (2016), the authors must provide accurate results, appropriate threshold value, and area range for deep learning methods.

Methodology

Given the methodological approach to this work using rationale-augmented CNN models, it became clear that the deepfake facial reconstruction task had already been solved concerning the accuracy and most of the available datasets at the time as presented through the flowchart in Fig. 3. However, our models were robust enough to allow for deepfake facial reconstruction outside of the available training/testing dataset (Kaggle 2021).

For a company to implement a facial reconstruction system for security, the current model would require a training session every time a new employee was hired. Additionally, a single forward pass of an image would be computationally expensive and would not allow for real-time face reconstruction. By training deep learning models while using a cross-entropy loss function, the model will not train to quantify new facial features into an encoded vector but rather train to maximize the correct number of classifications within a specific dataset as the architectural details are given in Fig. 4. These are two crucial problems that would need to be solved to create a robust statistical model with high performance.

The problems above were solved by Google researchers with the introduction of a loss metric known as a triplet loss.

In contrast to the previous models utilizing a SoftMax cross-entropy loss function, a metric loss function using triplets allows for a quantifiable way to compare and contrast facial features, as shown in Fig. 5. A triplet consists of three components:

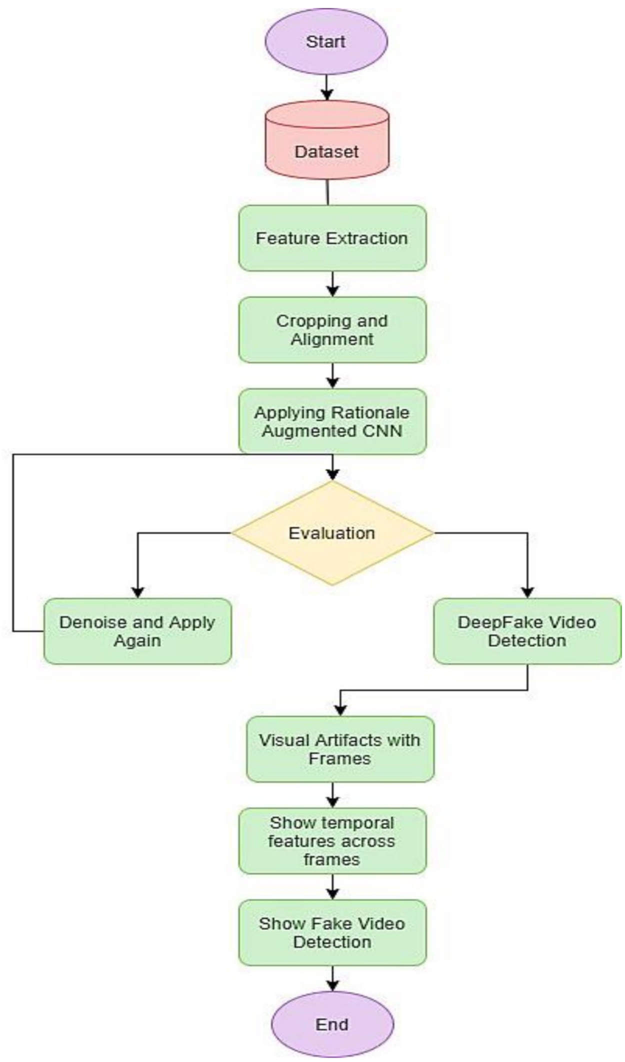

Fig. 3 Flowchart of the methodological approach

- Anchor: image of an individual
- Positive exemplar: another image of the same individual
- Negative exemplar: an image of an individual with similar facial features. Computing a triplet loss.

With a triplet loss function, the neural network will find the optimal weights and bias values to maximize the distance between images of different individuals while minimizing the distance between images of the same individuals. Once the face is detected and an image of a face is cropped, reconstruction becomes a trivial task with a network trained on triplet loss. A forward pass into such a model will produce a 128-dimensional vector that corresponds to the individual's unique identity, as represented in Table 1. By computing the L2 (Euclidean) distance between the 128-dimensional identity vector of two images, this value will represent how similar two faces are to one another. Additionally, specifying a threshold value will determine how sensitive or how different a face

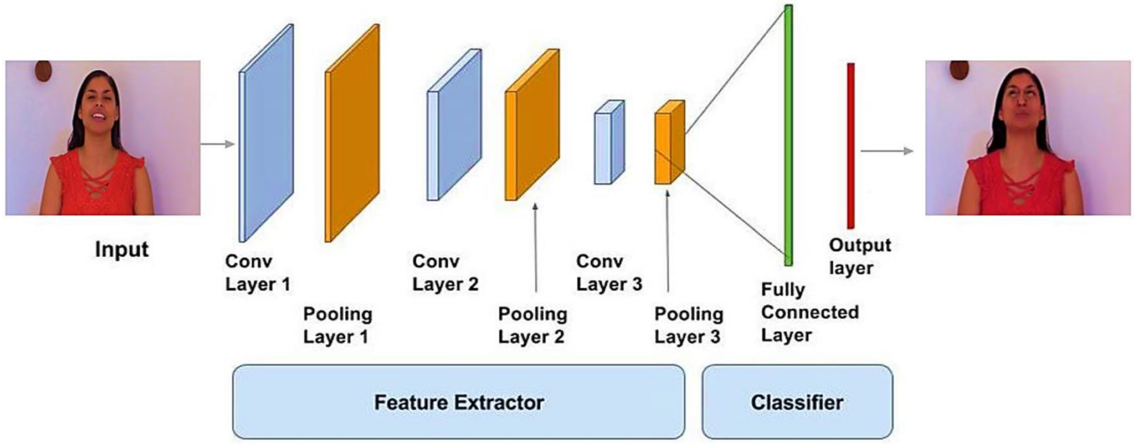

Fig. 4 Architectural diagram of rationale-augmented CNN network for deepfake video detection

Fig. 5 Manually created representation of learning with all the facial feature components in our work

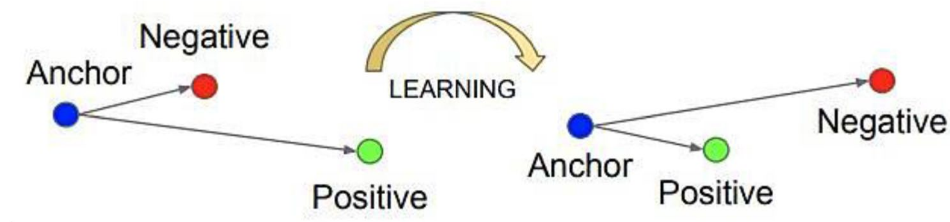

Table 1 The table that displays the values of each layer in CNN for this work

Layer	Size-in	Size-out	Kernel	Param	FLPS
Conv1	$220 \times 220 \times 3$	$110 \times 110 \times 64$	$7 \times 7 \times 3,2$	9 K	115 M
Pool1	$110 \times 110 \times 64$	$55 \times 55 \times 64$	$3 \times 3 \times 64,2$	0	
Rnorm1	$55 \times 55 \times 64$	$55 \times 55 \times 64$		0	
Conv2a	$55 \times 55 \times 64$	$55 \times 55 \times 64$	$1 \times 1 \times 64,1$	4 K	13 M
Conv2	$55 \times 55 \times 64$	$55 \times 55 \times 192$	$3 \times 3 \times 64,1$	111 K	335 M
Rnorm2	$55 \times 55 \times 192$	$55 \times 55 \times 192$		0	
Pool2	$55 \times 55 \times 192$	$28 \times 28 \times 192$	$3 \times 3 \times 192,2$	0	
Conv3a	$28 \times 28 \times 192$	$28 \times 28 \times 192$	$1 \times 1 \times 192,1$	37 K	29 M
Conv3	$28 \times 28 \times 192$	$28 \times 28 \times 384$	$3 \times 3 \times 192,1$	664 K	521 M
Pool3	$28 \times 28 \times 384$	$14 \times 14 \times 384$	$3 \times 3 \times 192,2$	0	
Conv4a	$14 \times 14 \times 384$	$14 \times 14 \times 384$	$1 \times 1 \times 384,1$	148 K	29 M
Conv4	$14 \times 14 \times 384$	$14 \times 14 \times 256$	$3 \times 3 \times 384,1$	885 K	173 M
Conv5a	$14 \times 14 \times 256$	$14 \times 14 \times 256$	$1 \times 1 \times 256,1$	66 K	13 M
Conv5	$14 \times 14 \times 256$	$14 \times 14 \times 256$	$3 \times 3 \times 256,1$	590 K	116 M
Conv6a	$14 \times 14 \times 256$	$14 \times 14 \times 256$	$1 \times 1 \times 256,1$	66 K	13 M
Conv6	$14 \times 14 \times 256$	$14 \times 14 \times 256$	$3 \times 3 \times 256,2$	590 K	116 M
Pool4	$14 \times 14 \times 256$	$7 \times 7 \times 256$		0	
concat	$7 \times 7 \times 256$	$7 \times 7 \times 256$		0	
Fc1	$7 \times 7 \times 256$	$1 \times 32 \times 128$	Maxout $p=2$	103 M	103 M
Fc2	$1 \times 32 \times 128$	$1 \times 32 \times 128$	Maxout $p=2$	34 M	34 M
Fc7128	$1 \times 32 \times 128$	$1 \times 1 \times 128$		524 K	0.5 M
L2	$1 \times 1 \times 128$	$1 \times 1 \times 128$		0	
Total				140 M	1.6B

must be to be considered a match or not a match. Concerning reconstruction, comparing the 128-d feature vector to a set of known encodings until another vector is found within the threshold accomplishes this task with ease.

We began by researching the state-of-the-art models with high accuracy concerning their performance. We prioritized models with faster performance and are smaller in size relative to their accuracy; this will allow for a model that can perform adequately in a real-time environment.

We have used the output vector patches given by the following equations from Keren et al. (2016):

$$c_{ij} = \sigma\left(x_{ij} + W_1 + b_1, \dots, x_{ij} + W_n + b_n\right), \tag{1}$$

where 'W' is the weight matrix, 'b' is the bias term, and 'x' is the number of channels in two dimensions for convoluting the filter. While the equation for sample classification is given by

$$y = argmax_j \frac{1}{r} \sum_{i=1}^{r} p_{ij}, \tag{2}$$

The formula gives the augmented equation for averaging the network with maximum arguments with specified range of filtering with CNN for all deepfake predictions:

$$y = argmax_j max\left(\left\{p_{ij}\right\} 1 \le i \le r\right). \tag{3}$$

Once a model or several models had been chosen, we researched for tweaks to the model's architecture and any impactful preprocessing steps that may contribute to the overall performance/efficiency of the models. By utilizing Deepfake models, a python library suited for this exact task, we were able to experiment with networks trained with triplet loss and had access to pre-trained face reconstruction models via CNN.

Additionally, it is important to consider the dataset that will be used. Based on the papers referenced, it seems that the rationale-augmented CNN (labeled faces in the wild) public dataset is ideal for this task given that they are rotated and cropped in the proper aspect. Lastly, we will attempt to tweak the model's hyper-parameters and apply regularization/dropout to obtain a model accuracy/loss, as the training and validation are shown in Figs. 6, 7. Suppose the opportunity of training a model myself is not possible. In that case, we will explore techniques that can be used to optimize/improve the performance of the face reconstruction process via state-of-the-art machine learning models.

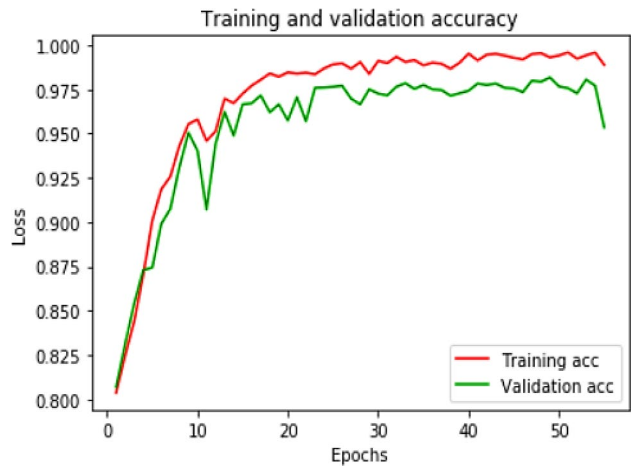

Fig. 6 Training and validation accuracy measured at 50 epochs of training, where the accuracy stands at 95.77%

Fig. 7 Training and validation loss measured at 50 epochs of training where the loss is as low as 0.675 only

Results

We have improved the current state of the deepfake models using rationale-augmented on MATLAB 2019b. The work can also be used from a security standpoint; it is important to consider false instances of a face, such as providing an image via an external screen or paper with a face printed. This can be solved by accounting for depth in a CNN model using a stereo camera or depth sensor. With this CNN model for detection, the model will only run facial reconstruction when an RGB (with depth) face is detected, rather than an RGB without depth image. From an accuracy standpoint, we could use augmented CNN to generate effective triplets by creating faces similar using the Kaggle DeepFake Challenge dataset, yet different feature vectors but different labels as experimental results are shown in Figs. 8, 9, 10. This can be extremely beneficial for researchers and enthusiasts because due to lack of access to the amount of data that companies like Google or Apple can use to train their state-of-the-art models for deepfake networks.

Discussion

Using this as our rationale-augmented CNN algorithm, we begin by applying multiple features of different scales to a sliding-box window across our video of deepfake network. Training a classified as with each feature as a "weak classifier" ultimately contributes to deciding whether or not a feature contributes to a facial feature. The classifier learns the features that have the strongest correlation to a human face. Additionally, regions of the image with little to no features are quickly skipped over and are not necessary for future computations. Furthermore, as its name suggests, a cascading CNN is an architecture composed of a series of CNNs that produce a bounding box prediction of the face's region. We displayed the pipeline of how a large set of bounding boxes are refined into a few accurate bounding boxes through forwarding passes throughout the networks and the comparison with other techniques presented in Table 2 below.

Fig. 8 The input video sample images that are inserted into the deepfake network for evaluation

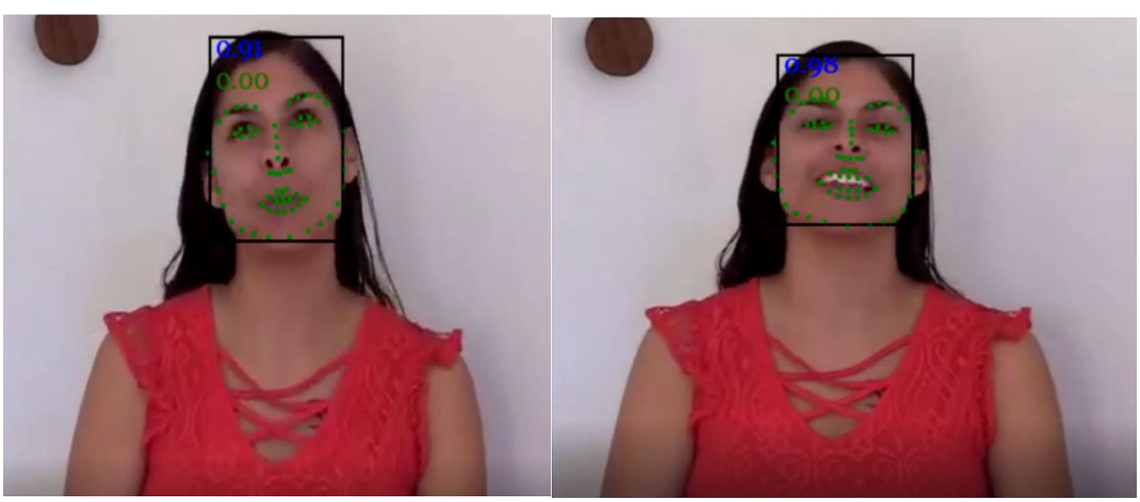

Fig. 9 The video image sample during training and validation of deepfake network

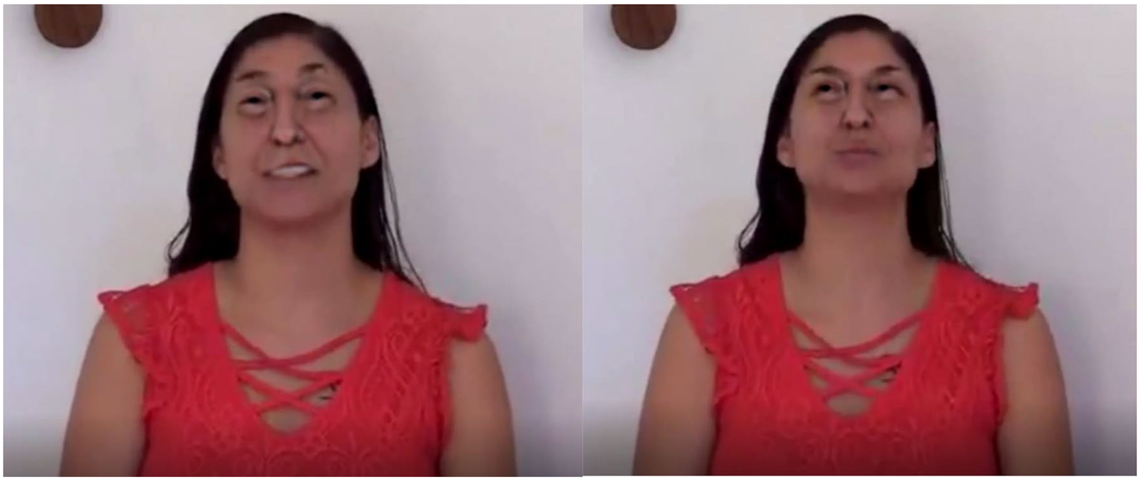

Fig. 10 Generated deepfake network after initialization of deepfake model of "Donald Trump" to an input video to create an annotated video

Table 2 Comparison of different techniques that have been performed earlier on deepfake network

Article	Technique	Accuracy (%)
Wubet (2020)	Transfer learning	93.23
Alhayani and Abdallah (2020)	Generative adversarial network	85.62
Proposed	Rationale-augmented CNN	95.77

Conclusion

A deepfake network framework could assume significant research in worldwide security by battling wrongdoing, and in the end, security concerns by facial reconstruction. This deepfake framework could be a wide assortment of utilizations that incorporate visual observation and access control in extraordinary conditions like face detection and reconstruction. This work is created utilizing the rationale-augmented convolutional neural network (CNN) on the MATLAB R2019a platform using the Kaggle DeepFake Video dataset. The contrast between the calculation expenses of the two actualized techniques is low; the CNN strategy is near ongoing. Low computational expense is fundamental because of the connection between the faces. We applied the filter of Donald Trump to deepfake as a quicker relationship technique has been executed. It has been composed a few times in this work; the dataset is tremendous and generally excellent for grouping and division quality. The CNN executed technique is very basic. This work is its blend with an increasingly deepfake arrangement with division with expanded accuracy of 95.77% being accomplished along these lines.

Funding This study was self-funded.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

References

Abu Rumman A (2021) Transformational leadership and human capital within the disruptive business environment of academia. WJET 13(2):178–187. https://doi.org/10.18844/wjet.v13i2.5652

Abu-Rumman A, Al Shraah A, Al-Madi F et al (2021) Entrepreneurial networks, entrepreneurial orientation, and performance of small and medium enterprises: are dynamic capabilities the missing link? J Innov Entrep 10:29. https://doi.org/10.1186/s13731-021-00170-8

Agarwal S, Varshney LR (2019) Limits of deepfake detection: a robust estimation viewpoint. arXiv preprint. arXiv:1905.03493

Alhayani B, Abdallah AA (2020) Manufacturing intelligent Corvus corone module for a secured two way image transmission under WSN. EC Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/EC-02-2020-0107

Alhayani BSA, llhan H, (2021) Visual sensor intelligent module based image transmission in industrial manufacturing for monitoring and manipulation problems. J Intell Manuf 32:597–610. https://doi.org/10.1007/s10845-020-01590-1

Alhayani B, Abbas ST, Mohammed HJ et al (2021) Intelligent secured two-way image transmission using corvus corone module over WSN. Wireless Pers Commun. https://doi.org/10.1007/s11277-021-08484-2

Al-Hayani B, Ilhan H (2020) Efficient cooperative image transmission in one-way multi-hop sensor network. Int J Electr Eng Educ 57(4):321–339

Arjovsky M, Chintala S, Bottou L (2017) Wasserstein generative adversarial networks. In Proceedings of the International Conference on Machine Learning, Sydney, 6–11 August 2017, 214–223.

Arjovsky M, Chintala S, Bottou L (2017) Wasserstein generative adversarial networks. In International Conference on Machine Learning, 214–223

Bai S (2017) Growing random forest on deep convolutional neural networks for scene categorization. Expert Syst Appl 71:279–287
Bayar B, Stamm MC (2016) A deep learning approach to universal image manipulation detection using a new convolutional layer. In Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security. ACM, 5–10
Chollet, F (2017) Xception: deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, 21–26 July 2017, 1610–2357
Dang L, Hassan S, Im S, Lee J, Lee S, Moon H (2018) Deep learning based computer generated face identification using convolutional neural network. Appl Sci 8:2610
Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC (2017) Improved training of Wasserstein GANs. In Advances in Neural Information Processing Systems, 5767–5777
Hsu C-C, Zhuang Y-X, Lee C-Y (2020) Deep fake image detection based on pairwise learning. Appl Sci 10(1):370. https://doi.org/10.3390/app10010370
Huang G, Liu Z, Van Der Maaten L, Weinberger K (2017) Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017, 4700–4708
Huaxiao Mo BC, Luo W (2018) Fake Faces Identification via Convolutional Neural Network. In Proceedings of the ACM Workshop on Information Hiding and Multimedia Security, Innsbruck, Austria, 20–22 June 2018, 43–47
Kaggle (2021) Available Online: https://www.kaggle.com/c/deepfake-detection-challenge/data
Keren G, Deng J, Pohjalainen J (2016) Convolutional Neural Networks with Data Augmentation for Classifying Speakers' Native Language, Chair of Complex & Intelligent Systems, University of Passau, Germany Department of Computing, Imperial College London, UK
Korshunova I, Shi W, Dambre J, Theis L (2017) Fast face-swap using convolutional neural networks. In Proceedings of the IEEE International Conference on Computer Vision, 3677–3685
Kwekha-Rashid AS, Abduljabbar HN, Alhayani B (2021) Coronavirus disease (COVID-19) cases analysis using machine-learning applications. Appl Nanosci. https://doi.org/10.1007/s13204-021-01868-7
Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC (2016) Ssd: Single shot multibox detector. In European Conference on Computer Vision; Springer: Cham, Switzerland. pp. 21–37
Mao X, Li Q, Xie H, Lau RY, Wang Z, Smolley SP (2017a) Least squares generative adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017, 2813–2821
Mao, X., Li, Q., Xie, H., Lau, R. Y., Wang, Z., and Paul Smolley, S. (2017b). Least squares generative adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2794–2802).
Miyato T, Kataoka T, Koyama M, Yoshida Y (2018) Spectral normalization for generative adversarial networks. In Proceedings of the International Conference on Learning Representations, Vancouver, 30 April–3 May 2018
Qian Y, Dong J, Wang W, Tan T (2015) Deep learning for stage analysis via convolutional neural networks. In Media Watermarking, Security, and Forensics 2015 9409: 94090J
Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv 2015. arXiv:1511.06434
Rashid AS, Tout K, Yakan A (2021) The critical human behavior factors and their impact on knowledge management system–cycles. Business Process Management Journal, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/BPMJ-11-2020-0508
Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M et al (2015) ImageNet large scale visual recognition challenge. Int J Comput vis 115:211–252
Sun Y et al (2014) Deep Learning Face Representation from Predicting 10,000 Classes. Obtained From: http://mmlab.ie.cuhk.edu.hk/pdf/YiSun_CVPR14.pdf
Wang X, Thome N, Cord M (2017) Gaze latent support vector machine for image classification improved by weakly supervised region selection. Pattern Recogn 72:59–71
Wen D, Han H, Jain AK (2015) Face spoof detection with image distortion analysis. IEEE Trans Inf Forensics Secur 10(4):746–761
Wubet WM (2020) The Deepfake Challenges and Deepfake Video Detection. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278–3075, Volume-9 Issue-6. https://doi.org/10.35940/ijitee. E2779.049620
Xuan X, Peng B, Dong J, Wang W (2019) On the generalization of GAN image forensics. arXiv preprint. arXiv:1902.11153
Yahya W, Ziming K, Juan W et al (2021) Study the influence of using guide vanes blades on the performance of cross-flow wind turbine. Appl Nanosci. https://doi.org/10.1007/s13204-021-01918-0
Yang P, Ni R, Zhao Y (2016) Recapture image forensics based on Laplacian convolutional neural networks. In International Workshop on Digital Watermarking. Springer, Charm, 119–128
Zheng L, Duffner S, Idrissi K, Garcia C, Baskurt A (2016) Siamese multi-layer perceptrons for dimensionality reduction and face identification. Multimed Tools Appl 75(9):5055–5073
Zhang Y, Zheng L, Thing VL (2017) Automated face swapping and its detection. In 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP), 15–19. IEEE