ITC 3/51

Information Technology
and Control
Vol. 51 / No. 3 / 2022
pp. 563-574
DOI 10.5755/j01.itc.51.3.31510

Deep Learning for Forgery Face Detection Using Fuzzy Fisher Capsule Dual Graph

Received 2022/05/29

Accepted after revision 2022/06/30

http://dx.doi.org/10.5755/j01.itc.51.3.31510

HOW TO CITE: Arunkumar, P. M., Sangeetha, Y., Raja, V., Sangeetha, S.N. (2022). Deep Learning for Forgery Face Detection Using Fuzzy Fisher Capsule Dual Graph. *Information Technology and Control*, 51(3), 563-574. http://dx.doi.org/10.5755/j01.itc.51.3.31510

Deep Learning for Forgery Face Detection Using Fuzzy Fisher Capsule Dual Graph

P. M. Arunkumar

Department of Computer Science and Engineering, Karpagam College of Engineering, Coimbatore, 641032, India; e-mail:pmarunkumara21@outlook.com

Yalamanchili Sangeetha

Department of Information Technology, VR Siddhartha Engineering College, Vijayawada,520007, Andhra Pradesh, India

P. Vishnu Raja

Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, 638060, India

S. N. Sangeetha

Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Sathy, 638401, India

Corresponding author: pmarunkumara21@outlook.com

In digital manipulation, creating fake images/videos or swapping face images/videos with another person is done by using a deep learning algorithm is termed deepfake. Fake pornography is a harmful one because of the inclusion of fake content in the hoaxes, fake news, and fraud things in the financial. The Deep Learning technique is an effective tool in the detection of deepfake images or videos. With the advancement of Generative adversarial networks (GAN) in the deep learning techniques, deep fake has become an essential one in the social media platform. Fake faces may threaten the public, therefore detection of deepfake images/videos is needed. For detecting the forged images/videos, many research works have been done and those methods are inefficient in the detection of new threats or newly created forgery images or videos, and consumption time is high. Therefore, this paper focused on the detection of different types of fake images or videos using Fuzzy Fisher face with Capsule dual graph (FFF-CDG). The data set used in this work is FFHQ, 100K-Faces DFFD, VGG-Face2, and

WildDeepfake. The accuracy for FFHQ data sets, the existing and proposed systems obtained the accuracy of 81.5%, 89.32%, 91.35%, and 95.82%, respectively.

KEYWORDS: Deep fake, deep learning, forgery face detection on image/video, generative adversarial networks, capsule dual graph, fake face detection.

1. Introduction

Manipulating the face refers to the swapping of the source face with the target face, which includes the concept of both identity exchange and expression exchange. The concept of identity exchange form is swapping the whole face of the source image or video and the target image or video with the change of its identity. However, in the expression change form, it can change only facial expressions, not its identity. In the development of manipulating the facial image or video, its result produces fake images or videos that are becoming realistic. To prevent this generation of fake image or videos effective and timely countermeasures is required. To manage the threat of facial manipulating image or videos number of manipulating detectors is implemented. For these various deep learning techniques are used [21, 15, 9, 6].

By using deep learning techniques like GAN and CNN are used to swap face images or videos and it is a more challenging task to preserve the pose, lighting of the photograph, and expressions in the facial image or videos [17]. In the GAN model of training images or videos in the large data set which has high fidelity of synthesis of image or video. To generate fake images/videos in the size of 128 × 128, self-attention GAN and spectral normalization GAN [5, 32, 24] are used. Many research works have been done and applying these techniques will inevitably risk and insecure in detecting the fake image or videos. Therefore, to improve the security and reduce the high inevitably risks this paper proposed Fuzzy Fisher face with a Capsule dual graph (FFF-CDG).

The author introduced the AlexNet convolutional neural network model, which would be regarded as a firm grounding for creating an item testing method based on deep learning in 2012 [9]. Deep learning-based object detection methods are now classified into two types. One is the Two-Stage method, built on R-CNN [6, 17, 5] and TridenNet [32], among other algorithms. The next is the SSD [24, 26, 16] and YOLO-based One-Stage method, which has excellent actual improvement in multi-scale object recognition.

Recent day fake face detection algorithm attains hard challenges in detection and classification. It is a difficult task to isolate the real eye from fake in the face. A variety of multimedia data content is used to tamper with cyber crime applications. Some crimes like false news publishing, digital kidnapping, disinformation, and Ransom ware attacks are very challenging in detecting fake faces. Mostly multimodal techniques are used in the detection of deepfake. Results are evaluated by identifying whether the target data is modified or not. The existing fake face detection model uses AI techniques like two-stream neural network modal, vision transformer, mesonet, etc. The major drawback is important to image regions are less focused and manual image processing is paid less attention.

Deep fake detection is a classification problem where classifiers detect all tempered and genuine videos. After 2017, when the deepfake is used, more algorithms are identified and implemented. Some work classifies manipulated videos and frames with artifacts using neural network techniques. Still, now these researches do not concentrate on effective feature extraction techniques. In this proposed work, features are extracted using a k-mean algorithm with efficient implementation of dimensionality reduction of features using a fuzzy-based fisher face algorithm. The main contribution of this work is:

1. To improve the accuracy of pre-processing in this work, we implement a bilateral filter and by using a k-mean algorithm for extracting features of the image or video.
2. For detecting the fake image or video or real image or video using capsuled dual graph methods accurately.

The paper has been organized as follows: Section 2 describes the review of the literature, Section 3 introduces deepfake detection using FFF-CDG, Section 4 discusses the experimented results and Section 5 concludes the paper with future directions.

2. Review of Literature

Generating face synthesis images or videos required only editing skills and a lot of time to implement this tool in the swapping of face images or videos like deepfakes [26]. The significant techniques of deep learning have got in the various applications of computer vision. The advances in the deep generative models and analysis of realistic content of fake images or videos are also referred to as deepfakes. The current deepfake detection techniques are implemented using binary classification problems in which authentic images or videos are distinguished from fake images or videos using two-class CNN [16]. The binary classifier needs a large volume of the database for detecting fake and real images or videos and training the network model. A huge number of fake images or videos is available, but the limitations are in terms of assigning the benchmark for justifying the different detection of fake images or videos. To address this issue [18] proposed a GAN model based on Faceswap-GAN with the deepfake data set of 620 videos. In the forensics model, swapping Face images or video is a challenging task for preserving the pose, lighting of the photographs, and expressions in the face of the image or video [33]. This paper [4] proposed the concept of a bag of words technique for extracting the features and these extracted features are fed into the different classifiers like random forest [28], SVM [34], and multi-layered perceptron [31].

A face detection system using the fisher face algorithm is specifically designed to recognize the face image or video by matching the results of its feature extraction. This paper proposed a model for detecting the deepfake concept using the technique of the head's position inconsistently. The algorithms are used for creating the face of various persons without modifying their originality in the aspect of expression. Hence, it creates mismatched facial landmarks. The location of the landmark varies from the fake face

Table 1

Survey on detection of fake images or videos

Author	Input Data Type	Type of Classifier	Name of Data Set
Chintha et al. (2020) [3]	Video	Bidirectional Convolutional recurrent LSTM network	Face Forensics++, Celeb-DF
Agarwal et al. (2020) [11]	Video	CNN	Instagram and YouTube
Fernandes et al. (2020) [23]	Video	ResNet50 model with pre-trained on VGGFace2	VidTIMIT, COHFACE, Deepfake TIMIT
Mittal et al. (2020) [2]	Video	Siamese network	Deepfake TIMIT and DFDC
Agarwal et al. (2020) [8]	Video	Rules-based extracting features and its behavior.	Face Forensics++, Google/Jigsaw deepfake detection data set DFDC and Celeb-DF.
Ciftci et al. (2020) [14]	Video	CNN	UADFV, Face Forensics , Face Forensics++ , Celeb-DF
Hsu et al. (2020) [12]	Image	CNN	Celebi , and DUGAN, WGAN, WGAN-GP
Gandhi et al. (2020) [13]	Image	VGG and ResNet	CelebA
Guarnera et al. (2020) [19]	Image	KNN, SVM	CelebA
Li et al. (2020) [19]	Image	CNN	Face Forensics++, Deepfake Detection (DFD), DFDC and Celeb-DF.
Wang et al. (2020) [28]	Image	ResNet-50	Face Forensics++
Li et al. (2019) [20]	Video	VGG16 ResNet50,	UADFV, Deepfake TIMIT
Nguyen et al. (2019) [34]	Video/ Image	Capsule networks	Face Forensics
Yang et al. (2019) [31]	Video/ Image	SVM	UADFV, DARPA MediFor GAN Image/Video

of the image or video to the real face of the image or video [10].

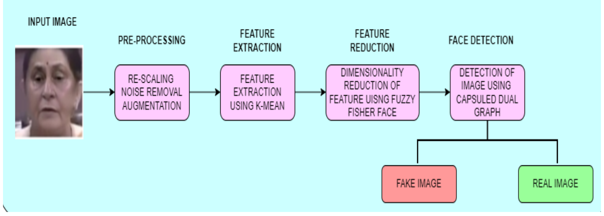
The concept of deepfake is super realistic, high in pervasive, providing authenticity of images or videos in trustworthiness has become a challenging task. Deepfakes are implemented by exploiting AI algorithms. In which images or videos are captured from the camera and detecting the image or videos using deep learning techniques [30]. Recently, many research works have been focused on the analysis of deepfake which shows that various artifacts in the concept of deepfake. Developing the automated detection of deepfake images or videos using the AI model is developed [27, 22, 7].

The exsisting algorithms has tendency to predict the fake images, but major concern is how accurate the prediction is? Also sometimes training data accuracy is not high which leads test data in low performance. To ensure fake image, there is need of intelligent algorithm for both training and testing.

3. Proposed FFF-CDG Methodology

In this section, the architecture of FFF-CDG is given in Figure 1. In the pre-processing phase, Rescaling Image, noise removal, face image augmentation. The noises in the images are removed in the pre-processing stage. Pre-processed images are given as input to the feature extraction stage. The k-mean technique is used in clustering related features. Further extracted features are processed with dimensionality reduction techniques with fuzzy fisher face algorithms. Finally, fake detection is done by a capsuled dual graph algorithm.

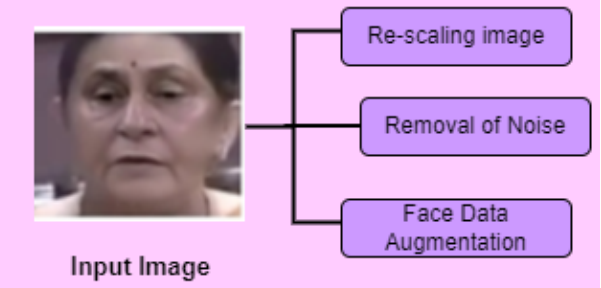
Figure 1
Framework of FFF-CDG


INPUT IMAGE
PRE-PROCESSING
FEATURE EXTRACTION
FEATURE REDUCTION
FACE DETECTION
RE-SCALING NOISE REMOVAL AUGMENTATION
FEATURE EXTRACTION USING K-MEAN
DIMENSIONALITY REDUCTION OF FEATURE UISNG FUZZY FISHER FACE
DETECTION OF IMAGE USING CAPSULED DUAL GRAPH
FAKE IMAGE
REAL IMAGE

3.1. Pre-Processing

For classifying the deepfake of the image in an efficient manner pre-processing is needed which enhances the image for further processing. The steps involved in pre-processing are given in Figure 2.

Figure 2
Pre-processing



3.1.1. Rescaling Image or Video

The image or video form data set contains RGB band values of 0 to 255. Processing the image in this band value is not an applicable one for applying directly to the proposed work of FFF-CDG. Therefore, rescaling the input image between 0 and 1 using the scaling factor of 1/255 is done.

3.1.2. Noise Removal

To improve the efficiency in the detection of fake images and real images or video, the noise is reduced from raw input face images by using a bilateral filter. A bilateral filter is a technique of spatial smoothing. It is a kind of nonlinear filter. It is a collection of spatial information of the image, the similarity of color information, and has a similarity of the image with its pixel values. It reduces the noise and smoothening of the image but maintains the details of the edge in the image. The formula used for this is given below:

$$img_{(i,j)} = \frac{\sum(p,q) \in p(i,j) f(p,q) w(i.j,p,q)}{\sum(p,q) \in p(i,j) w(i.j,p,q)}, \quad (1)$$

where, (i, j) is the position of pixel, $img_{(i,j)}$ denotes the output image, $f(p, q)$ is input image and $w(i. j, pin, q)$ is the calculation of two gaussian function value. In the bilateral filter, using spatial proximity weight is calculated and similarity pixel value is multiplied and weight of pixels is convolved in the image. In this way noise will reduce in the image with keeping the edges of the image.

3.1.3. Face Image or Video Zooming Augmentation

To detect the face in the image effectively, zooming augmentation is needed. The zooming range of the image is 0.2 parametric values. The parametric range will be 1- value to 1+ value.

3.2. Face Feature Extraction

Feature extraction is the main component in the detection of face. To achieve a high level of detection of face extracting features like the position of nose, eyes, chin, and mouth. In this work k-mean algorithm is used. It is a classical distance-based algorithm. Calculate the distance between two points if it is smaller then it shows the similarity of pixel values. Finally, get one class of the face image or video. Divide any g group of objects into p point groups. To get a high similarity of pixels by calculating the average of data objects near p point groups. It is calculated by using

$$\mu_p = \frac{\sum_{i=1}^{n}\{c^i=p\}x^i}{\sum_{i=1}^{n}\{c^i=p\}}, \quad (2)$$

where c^i denotes the nearest point group in all datapoints. i and p. μ_p is the centre point in point group.

3.3. Proposed Face Detection Using FFF-CDG

3.3.1. Fuzzy Fisher Face

In order to evaluate the detection of fake image or videos in this work Fuzzy Fisher faces an algorithm with Convolutional long short-term memory (FFF-CDG). The detection of fake face images or videos by evaluating within-class and between-class pixels in the image. In the fuzzy-based concept of the fisher face algorithm, which partitions the fake image [1, 29] and real image or video using a given set of feature vector values. These feature vector values are transformed by principal component analysis (PCA). Let feature set of values are $Y = \{y_1, y_2, y_3, ..., y_n\}$ Fuzzy C-Means partition these vector values which specify the degrees of membership of each vector values to the classes. Then the partition matrix is represented by $P = [\mu_{ij}]$ for i=1,2,...,m and j=1,2,...n which satisfy the two properties like

$$\sum_{i=1}^{c} \mu_{ij} = 1 \quad (3)$$

$$0 < \sum_{j=1}^{n} \mu_{ij} < n. \quad (4)$$

The membership values are close to 0.5 which exhibits the vector values of high membership vector values. For computing the membership degree by using the following steps:

Step 1: In the training process calculate the matrix of Euclidean distance between pair of two features of vector values.

Step 2: Assign infinity to all entries in the matrix.

Step 3: Sort them in ascending order.

Step 4: Gather all class labels for the nearest neighbor's pattern.

Step 5: Calculate the membership grade to class " i for j^{th} pattern by using:

$$me\mu_{ij} = \begin{cases} 0.51 + 0.49(n_{ij}/k) \ if\ i\ is\ the\ same\ as\ jth\ pattern's\ label \\ 0.49(n_{ij}/k)\ if\ i\ is\ \neq\ same\ as\ jth\ pattern's\ label \end{cases} \quad (5)$$

where n_{ij} the number is neighbours in the j^{th} pattern of i^{th} class. By this way, we have applied fuzzified to the membership grade for all label patterns.

3.3.2. Deepfake Detection Using Capsule Dual Graph (CDG)

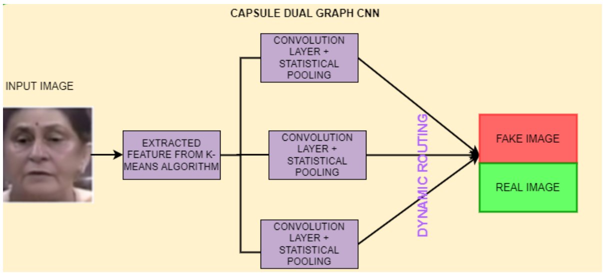
The proposed detection system of deepfake detection using Capsule dual graph CNN consists of three capsules. To detect fake and real face image or videos two capsules are needed. The dual graph CNN is represented in one capsule to perform the detection. The features extracted from Section 3.2 are given as input to this detection model. The dual graph neural network is the variance of traditional neural network with graph. Each node of the graph represents the features of the face image or video. Dual graph consists of two CNN and the input data is represented as matrix $Y \in R^{m \times n}$ with the data points points $Y = \{d_1, d_2, ... d_i, d_{i+1}, d_n\}$. The structure of the graph is denoted as an adjacency matrix called $Adj \in R^{n \times n}$. Vector features and adjacency matrix Adj are the DGCNN model inputs. Output of local consistency for hidden layer i is declared in j Equation (6).

$$conv_{LC}^{i}(P) = z^i \sigma\left(P^{-\frac{1}{2}}\overline{AdJ}P^{-\frac{1}{2}}z^{i-1}r^i\right) \quad (6)$$

where, $\overline{AdJ} = Adj' \times I_n$, I- identity matrix, $\overline{AdJ}$ – adjacency matrix with self-loops, P- is the adjacency matrix in normalized form, z- output, r-parametric value for training and σ - activation function (ReLU). Figure 3 shows that work flow of capsule dual graph in CNN.

Figure 3

Capsule dual graph of CNN


CAPSULE DUAL GRAPH CNN
INPUT IMAGE
EXTRACTED FEATURE FROM K-MEANS ALGORITHM
CONVOLUTION LAYER + STATISTICAL POOLING
CONVOLUTION LAYER + STATISTICAL POOLING
CONVOLUTION LAYER + STATISTICAL POOLING
DYNAMIC ROUTING
FAKE IMAGE
REAL IMAGE

This Figure 3 contains three main capsules and produces two output capsules for real image and fake image or videos. Features extracted from k-means algorithm are fed as input and it is distributed to three main capsules. In the three main capsules includes statistical pooling which is mainly used for detecting the forgery image or videos. The outputs of three main capsules are dynamically routing to output capsules. This output capsules are detecting the image or video is real or fake. The Algorithm 1 describes the capsule dual graph.

Algorithm 1. Dynamic Routing Capsules Dual Graph

```
Step 1: Procedure Dynamic Routing ($out_{j/i}$, $Wt$, $it$)
Step 2: for $i$ in range do    // DGCNN
  Step 3: $\widehat{Wt} \leftarrow Wt + \text{rand}(\text{size}(Wt))$
  Step 4: $\widehat{out}_j \leftarrow \widehat{wt}_j \; squash(out_{j/i})$ where $wt_i \in R^{m\times n}$
  Step 5: For all input $i$ capsules and output capsule $j$ do
    Step 6: $bi_{ij} \leftarrow 0$
  Step 7: End For
  Step 8: For $it$ iterations do
    Step 9: For all input capsules $i$ do
      Step 10: $inc_i \leftarrow softmax(bi_i)$
    Step 11: For all output capsules $j$ do
      Step 12: $outs_j \leftarrow \sum_i inc_{ij} \, \widehat{out}_{j/i}$
    Step 13: For all output capsules $j$ do
      Step 14: $o_j \leftarrow squash(sq_j) = \frac{\|s_i\|^2}{1+\|s_i\|^2} \times \frac{s_i}{\|s_i\|}$
    Step 15: for all input $i$ capsules and output capsules do
      Step 16: $bi_{ij} \leftarrow bi_{ij} + \widehat{out}_{j/i}.o_j$
  Step 17: End For
  Step 18: return $o_j$
Step 19: End For
```

In the Algorithm 1, the dynamic routing three main output capsules $out_{j/i}$ for it iterations are evaluated. To improve the efficiency of the algorithm by slightly add Gaussian noise to the 3-D weight value of tensor Wt and implementing the squash as in Equation (7) before process the routing by all iterations. This added Gaussian noise helps to reduce over-fitting of the graph. The outputs of main capsules are calculated as:

$$L = -(x\log(\hat{x}) + (1 - x) \; \log(1 - \hat{x})), \tag{9}$$

where, x is the ground truth value of label and $\hat{x}$ is the predicted label calculated using Equation (9). The output capsule o_j by using:

$$\hat{x} = \frac{1}{m} \sum_i softmax\left(\begin{bmatrix} o_1^T \\ o_2^T \end{bmatrix}\right) \tag{10}$$

By using Equation (10) the output capsules length separates the two output capsules like detection of real image and fake image or video for all dimensions.

4. Experimental Result and Discussions

The proposed deep learning based deepfake detection system with efficient feature extraction and detection process is experimented with fake and real image or videos of public data sets such as FFHQ, 100K-Faces, VGGFace2 and WildDeepfake. The proposed system is implemented using the machine learning library called PyTorch.

4.1. Data Set Description

4.1.1. Flickr-Faces-HQ, FFHQ

Flickr-Faces-HQ, FFHQ data set contains a group of 70,000 face images with a high-quality resolution generated by generative adversarial networks (GAN).

4.1.2. 100K-Faces

100K-Faces data set contains 100,000 unique human face images generated using StyleGAN

4.1.3. VGGFace2

The data set VGGFace2 contains large amount of face images or videos from various types of nearly nine thousand different subjects, with an average of more than 300 images or videos per subject. Image or videos were collected from the Google search engine with information of illumination, ethnicity, age, and occupation.

4.1.4. WildDeepfake

It is a deepfake detection real world data set collected from internet. This data set subjects of real and fake are collected from internet sources and consists of diverse scenes, each scene consists of more persons with rich facial expressions.

4.2. Evaluation Metrics

The proposed FFF-CDG based Capsule Dual graph deep fake detection is evaluated with various evaluation metrics such as Accuracy, Sensitivity, Specificity, ROC and error detection rate. The Deep Fake detection system is compared with standard deep fake detection approaches such as VGG19, ResNet and MobileNet.

Accuracy

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} X100 \tag{11}$$

Sensitivity

$$Sensitivity = \frac{TP}{TP+FN} X100 \tag{12}$$

Specificity

It is used to evaluate the rate between True Negative (TN) and True Positive (TP)

$$Specificty = \frac{TN}{TN+FP} X100 \tag{13}$$

Half Total Error Rate (HTER)

$$\frac{FRR+FAR}{2} \tag{14}$$

where, *FRR* is False Rejection Rate and *FAR* is False Acceptance Rate.

The Accuracy comparison of the proposed FFF-CDG deepfake detection is shown in Table 2 for various data sets. With the baseline of various approaches such as VGG1 face2, ResNet, Mobile Net and experimented with proposed efficient deep fake image or video feature extraction with capsule dual graph for various data sets.

Table 2

Accuracy comparison of proposed vs traditional baseline system for various data sets

Methods	Data sets			
	FFHQ	100K Faces	VGG face2	Wild Deep Fake
VGG19	81.5	74.78	87.43	89.25
ResNet	89.32	79.13	89.78	86.25
Mobile Net	91.35	90.21	90.01	96.75
Proposed FFF-CDG	95.82	96.34	97.12	98.23

For FFHQ data sets, the existing and proposed system obtained the accuracy of 81.5%, 89.32%, 91.35% and 95.82%, respectively. For 100K faces data sets, the approaches obtained the accuracy of 74.78 %, 79.13%, 90.21% and 96.34% sequentially. For VGGFace 2 data set, accuracy values are 87.43%, 89.78%, 90.01%, 97.12% and for WildDeepfake data set, 89.25%, 86.52%, 96.75% and 98.91%, respectively. The analyzed results proves that the proposed system acquired improved percentage than the traditional baseline deep fake detection systems. The sensitivity, specificity, ROC comparison of various deep fake detection system using four different data sets are evaluated and the results are shown in Table 3 and Figure 8.

Table 3

Sensitivity and Specificity analysis of proposed system in different data sets

Methods	Sensitivity %			
	VGG 19	Res Net	Mobile Net	Proposed FFF-CDG
FFHQ	81.5	74.78	87.43	89.25
100K-Faces	89.32	79.13	89.78	86.25
VGG -FACE2	91.35	90.21	90.01	96.75
WildDeepfake	95.82	96.34	97.12	98.23
Methods	Specificity %			
	VGG face2	Res Net	Mobile Net	Proposed FFF-CDG
FFHQ	80.34	83.15	85.28	93.34
100K-Faces	84.2	86.2	89.54	94.67
VGG -FACE2	87.1	87.1	89.1	95.16
WildDeepfake	86.1	89.4	86.3	96.32

Table 3 shows that sensitivity, specificity analysis of proposed FFF-CDG compared with existing algorithms and various data sets of FFHQ, 100K-Faces, VGG-face 2, WildDeepfake. The proposed FFF-CDG-got sensitivity score as 91.14 % in FFHQ data set, 89.85% in 100K-Faces data set, 88.13 % in VGG-FACE2 data set and 93.15% in WildDeepfake data set. Similarly, for Specificity of proposed work FFF-CDG got 93.34% in FFHQ data set, 94.67% in 100K-Faces data set, 95.16% in VGG-FACE2 data set and 96.32% in WildDeepfake data set. The analysis of training and testing of data set in the face image or video and produce the validation face input data in the terms of accuracy and loss metric information for the deepfake face image or video data set. Our proposed work FFF-CDGmodel in epochs is shown in Figure 4.

It is observed from Figure 4 that in our FFHQ, 100K-Faces, DFFD, CASIA-webface data set gives less validation loss and good validation accuracy for FFF-CDGmodel. Figure 5 shows the AUROC value comparison of various Deepfake detection system with various data sets.

Figure 4

Training and Validation Accuracy and Loss in FFF-CDG

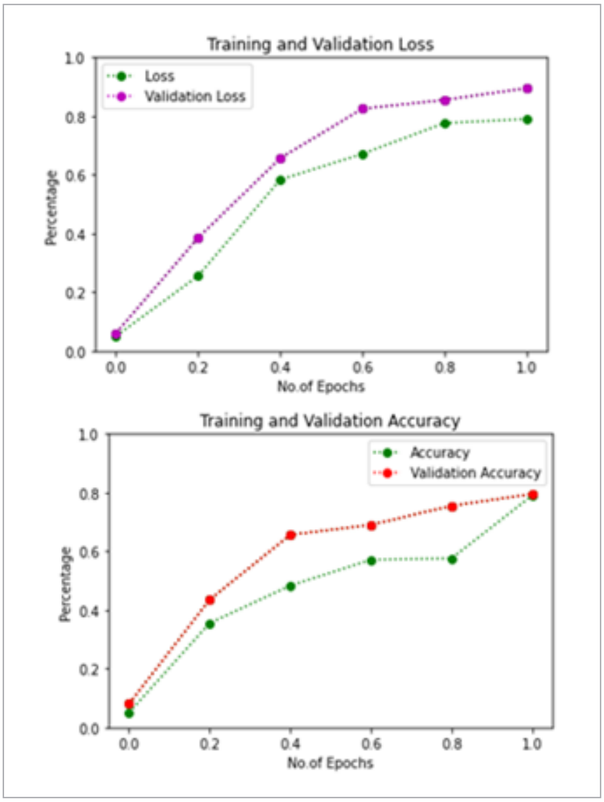
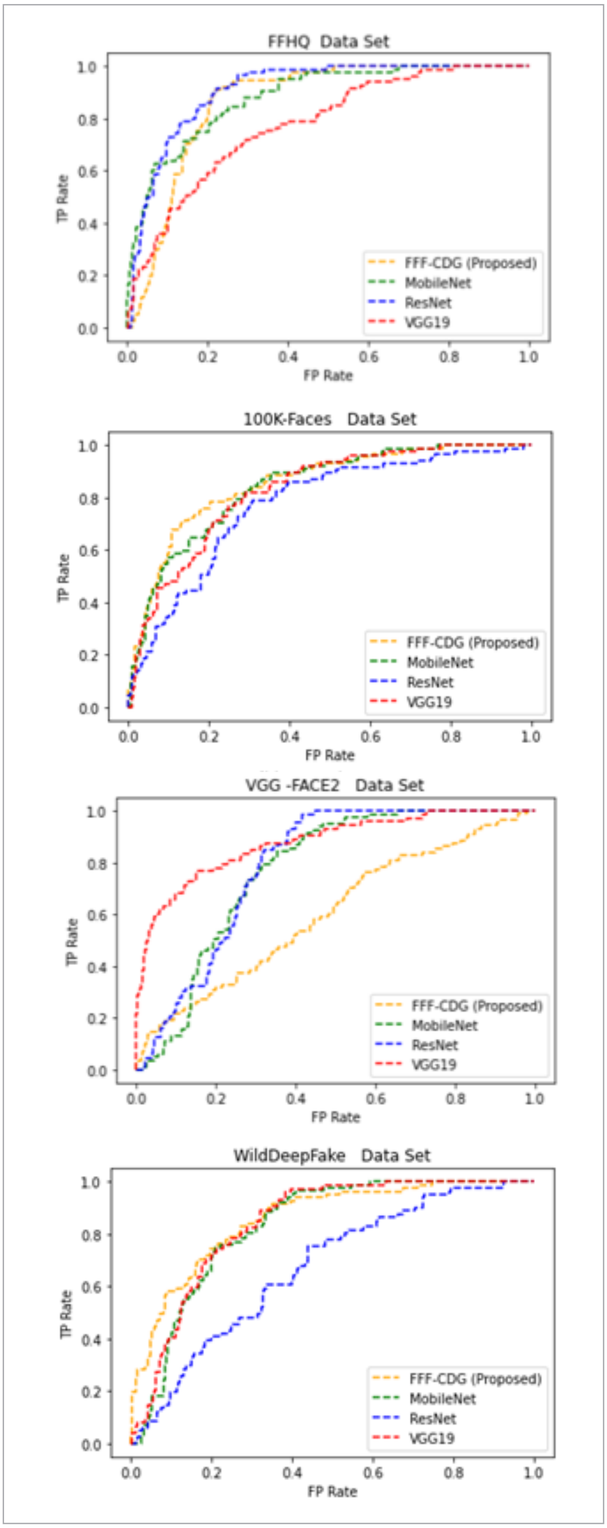


Figure 5

AUROC



From the analysis, the proposed system obtained improved ROC value of 84.2% for FFHQ data set, 84.1% for 100K-faces data set, 86.12% for VGG-FACE2 data set and 86.74% for WildDeepfake data set. Various baseline VGG19 systems were obtained for the data sets such as FFHQ, 100K-Faces, VGG-FACE2 and WildDeepfake as 82.1%, 83.1%, 85.53% and 81.12%. Baseline ResNet approach secured 89.2%, 83.3%, 81.47% and 82.78%. Baseline MobileNet approach obtained 81.21%, 82.1%, 81.72% and 81.54% correspondingly. With the comparison, the proposed DF detection system secured improved ROC value compared to traditional baseline systems. Table 4 presents the performance of proposed methods with different data sets in terms of error detection rate.

Table 4

Error detection rate

DF detection methods	Data sets			
	FFHQ	100K-Faces	VGG-Face2	Wild Deep Fake
VGG19	12.41	17.32	15.33	14.42
ResNet	13.45	15.23	12.15	11.23
MobileNet	11.25	13.22	12.65	12.21
Proposed FFF-CDG	5.34	6.45	6.12	5.78

The proposed approach obtained minimum error rate of 5.34 for the data set FFHQ data set, 6.12 for VGG-Face2, 6.45 for 100K-faces and 5.78 for WildDeepfake data sets. The error rate is minimum compared to the baseline deepfake detection methods such as VGG19, ResNet and MobileNet. Figure 6 illustrates the half total error rate (HTER) of various approaches with respect to various data sets.

The proposed DF detection system secured minimum HTER of 1.89 for FFHQ data set, 2.67 for 100K-faces data set, 2.8 for VGG-FACE 2 data set and 3.0 for the WildDeepfake data set. These HTER are minimum compared to traditional baseline systems such as VGG19, ResNet and MobileNet. Figure 7 shows the computational time comparison.

The proposed system secured 9.2 ms for VGG-Face2 data set, 10.89 ms for FFHQ, 11.44 ms for wild deepfake data set and 12.42 ms for 100K-Faces data set. This is minimum compared to other traditional baseline DF detection system where VGG19 secured the computational time for the data sets as 15.2 ms, 16.4 ms, 10.45 ms and 12.56 ms. ResNet obtained 16.5 ms, 18.76 ms, 11.65 ms and 13.33 ms and MobileNet secured 14.8 ms, 16.55 ms, 10.56 ms and 12.21 ms, respectively. Hence, all the kind of evaluation results proves that proposed FFF-CDG obtained improved sensitivity, specificity, accuracy and minimum error rate, EER and computation time. This proves that proposed system is efficient detection of DF image or videos with improved accuracy and minimum error.

Figure 6

HTER comparison of DF detection systems

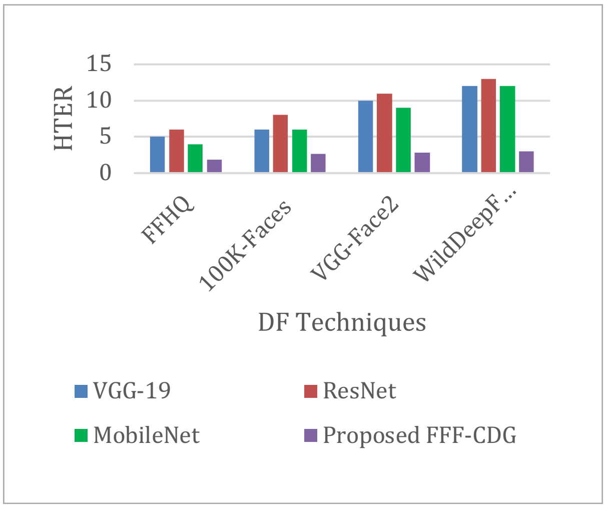
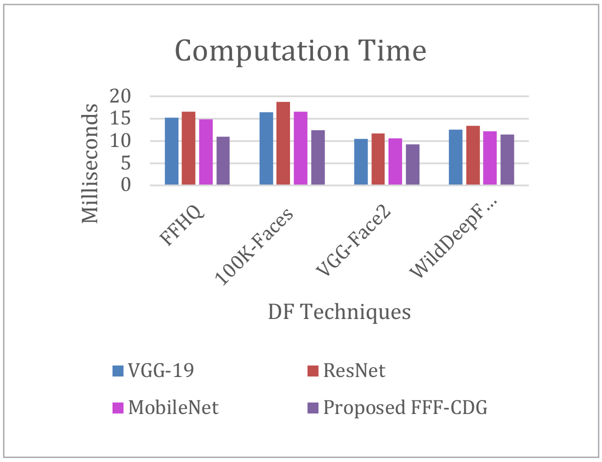


Figure 7

Computation time



5. Conclusion

This paper demonstrated the deep learning method for Deep Fake detection of image or videos in an effective way. From the extracted features, face image or videos are used for DF detection in a fine-tuned structure. The features from the face image or videos are extracted using proposed k-mean algorithm. The most relevant features are extracted by enhancing the k-mean based feature extraction with nearest point group approach. These relevant extracted features are then fed as input into the capsule network for deepfake detection. There are five capsules which include three for input capsule and two output capsules to represent the fake and real image or video. Experimental analysis with various baseline DF detection approaches such as VGG19, ResNet and MobileNet using the benchmark DF image or video data sets includes FFHQ, 100K-faces, VGG-Face2, WildDeepfake demonstrated that the proposed DF detection approach secures improved accuracy for FFHQ data sets, the existing and proposed system obtained the accuracy of 81.5%, 89.32%, 91.35% and 95.82% respectively. For 100K faces data sets, the approaches obtained the accuracy of 74.78 %, 79.13%, 90.21% and 96.34% sequentially. For VGG-Face 2 data set, accuracy values are 87.43%, 89.78%, 90.01%, 97.12% and for WildDeepfake data set, 89.25%, 86.52%, 96.75% and 98.91%, respectively. In terms of half total error rate, the proposed DF detection system secured minimum HTER of 1.89 for FFHQ data set, 2.67 for 100K-faces data set, 2.8 for VGG-FACE 2 data set and 3.0 for wild deepfake data set. The proposed system secured 9.2 ms for VGG-Face2 data set, 10.89 ms for FFHQ, 11.44 ms for wild deepfake data set and 12.42 ms for 100K-Faces data set. Hence, all evaluation proves that proposed DF detection method is general and effective to detect wide range of fake image or video attacks. In future, the proposed system is extended up to classifiers in a different network for the analysis of fake and real image or video in the data set. Limitation of the study is used ML techniques is not up to 100 percent accuracy in fake detection.

References

1. Abayomi-alli, O. O., Damaševičius, R., Maskeliūnas, R., Misra, S. Few-shot Learning with a Novel Voronoi Tessellation-based Image Augmentation Method for Facial Palsy Detection. Electronics, 2021, 10(8), 978. https://doi.org/10.3390/electronics10080978

2. Agarwal, S., El-Gaaly, T., Farid, H., Lim, S. N. Detecting Deep-fake Videos from Appearance and Behavior. arXiv preprint arXiv:2004.14491, 2020. https://doi.org/10.1109/WIFS49906.2020.9360904

3. Agarwal, S., Farid, H., Fried, O., Agrawala, M. Detecting Deep-fake Videos from Phoneme-viseme mismatches. Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, 660-661. https://doi.org/10.1109/CVPRW50498.2020.00338

4. Bai, S. Growing Random Forest on Deep Convolutional Neural Networks for Scene Categorization. Expert Systems with Applications, 2017, 71, 279-287. https://doi.org/10.1016/j.eswa.2016.10.038

5. Brock, A., Donahueand, J., Simonyan, K. Large Scale GAN Training for High Fidelity Natural Image Synthesis. arXivpreprint arXiv:1809.11096,2018.

6. Chang, X., Wu, J., Yang, T., Feng, G. Deepfake Face Image detection Based on Improved VGG Convolutional Neural Network. Proceedings 39th Chinese ControlConference, Shenyang, China, 2020. https://doi.org/10.23919/CCC50068.2020.9189596

7. Chintha, A., Thai, B., Sohrawardi, S. J., Bhatt, K. M., Hickerson, A. et al. Recurrent Convolutional Structures for Audio Spoof and Video Deepfake Detection. IEEE Journal of Selected Topics in Signal Processing, 2020, 85. https://doi.org/10.1109/JSTSP.2020.2999185

8. Ciftci, U. A., Demirand, I., Yin, L. Fake catcher: Detection of Synthetic Portrait Videos Using Biological Signals. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020. https://doi.org/10.1109/TPAMI.2020.3009287

9. ĐorCevi, M., Milivojevi, M., Gavrovska, A. J. Deepfake Video Production and SIFT-based Analysis. Proceedings of 2019 27th Telecommunications Forum TELFOR, Belgrade, Serbia, 2019, 3. https://doi.org/10.5937/telfor2001022Q

10. Du, M., Pentyala, S., Li, Y., Hu, Y. Towards Generalizable Deepfake Detection with Locality-aware Autoencoder. Proceedings of 29th ACM International Conference on Information & Knowledge Management, Gold Coast, Australia, 2020, 325-334. https://doi.org/10.1145/3340531.3411892

11. Fernandes, S., Raj, S., Ewetz, R., SinghPannu, J., KumarJha, J. et al. Detecting Deepfake Videos Using Attribution-based Confidence Metric. Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, 308-309. https://doi.org/10.1109/CVPRW50498.2020.00162

12. Gandhi, A., Jain, S. Adversarial Perturbations Fooldeepfake Detectors. arXiv preprint arXiv2003.10596, 2020. https://doi.org/10.1109/IJCNN48605.2020.9207034

13. Guarnera, L., Giudice, O., Battiato, S. Deepfake Detection by Analyzing Convolutional Traces. Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, 666-667. ttps://doi.org/10.1109/CVPRW50498.2020.00341

14. Hsu, C. C., Zhuang, Y. X., Lee, C. Y. Deepfake Image Detection Based on Pairwise Learning. Applied Sciences, 2020, 10(1). https://doi.org/10.3390/app10010370

15. Joel, F., Eisenhofer, T., Schönherr, L., Fischer, A., Kolossa, D. et al. Leveraging Frequency Analysis for Deep Fake Image Recognition. Proceedings of International Conference on Machine Learning, Guangzhou, China, 2020, 3247-3258.

16. Korshunov, P., Marcel, S. Vulnerability Assessment and detection of Deepfake videos. In The 12th IAPR International Conference on Biometrics (ICB), 2019, 1-6. https://doi.org/10.1109/ICB45273.2019.8987375

17. Korshunova, I., Shi, W., Dambre, J. Fastface-swap Using Convolutional Neural Networks. Procedings IEEE International Conference on Computer Vision, 2017, 3677-3685. https://doi.org/10.1109/ICCV.2017.397

18. Korshunova, I., Shi, W., Dambre, J. Fastface-swap Using Convolutional Neural Networks. Proceedings of.IEEE International Conference on Computer Vision,Seoul, Korea, 2017, 3677-3685. https://doi.org/10.1109/ICCV.2017.397

19. Li, L., Bao, J., Zhang, T., Yang, H., Chen, D. et al. Face X-ray for More General Face Forgery Detection. Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, 5001-5010. https://doi.org/10.1109/CVPR42600.2020.00505

20. Li, Y., Lyu, S. Exposing Deepfake Videos Bydetecting Face Warping Artifacts. Proceedings IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019, 46-52.

21. Luo, Y., Ye, F., Weng, B., Du, S., Huang, T. A Novel Defensive Strategy for Facial Manipulation Detection Combining Bilateral Filtering and Joint Adversarial Training. Security and Communication Networks, 2021, 21. https://doi.org/10.1155/2021/4280328

22. Mirsky, Y., Lee, W. The Creation and Detection of Deepfakes: A Survey. ACM Computer Survey, 2021, 54, 1-41. https://doi.org/10.1145/3425780

23. Mittal, T., Bhattacharya, U., Chandra, R., Beraand, A., Manocha, D. Emotions Don't Lie: A Deepfake Detection Method Using Audio-Visual Affective Cues. arXiv preprintarXiv:2003.06711, 2020. https://doi.org/10.1145/3394171.3413570

24. Miyato, T., Kataoka, T., Koyama, M., Yoshida, Y. Spectral Normalization for Generative Adversarial Networks. arXiv preprint arXiv:1802.05957,2018.

25. Parthasarathi, P., Shankar, S. Decision Tree Based Key Management for Secure Group Communication. Computer Systems Science and Engineering, 2022, 42(2), 561-575. https://doi.org/10.32604/csse.2022.019561

26. Ramachandran, S., Varma Nadimpalli, A., Rattani, A. An Experimental Evaluation on Deepfake Detection Using Deep Face Recognition. arXiv:2110.01640v1, 2021. https://doi.org/10.1109/ICCST49569.2021.9717407

27. Shelke, N. A., Kasana, S. S. A Comprehensive Survey on Passive Techniques for Digital Video Forgery Detection. Multimed. Tools Application, 2021, 80, 6247-6310. https://doi.org/10.1007/s11042-020-09974-4

28. Wang, X., Thome, N., Cord, M. Gazelatentsupport Vector Machine for Image Classification Improved by Weakly Supervised Region Selection. Pattern Recognition, 2017, 72, 59-71. https://doi.org/10.1016/j.patcog.2017.07.001

29. Wei, W., Ho, E. S. L., McCay, K. D., Damaševičius, R., Maskeliūnas, R. et al. Assessing Facial Symmetry and Attractiveness Using Augmented Reality. Pattern Analysis and Applications, 2021, 1-17. https://doi.org/10.1007/s10044-021-00975-z

30. Yadavand, D., Salmani, S. Deepfake: A Survey on Facial Forgery Technique Using the Generative Adversarial Network. Proceedings of 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, 852-857. https://doi.org/10.1109/ICCS45141.2019.9065881

31. Yang, X., Li, Y., Lyu, S. Exposing Deep Fakes Using Inconsistent Head Poses. Proceedings of IEEE Inter-

national Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 2019. https://doi.org/10.1109/ICASSP.2019.8683164

32. Zhang, H., Goodfellow, I., Metaxas, D., Odena, A. Self-attention Generative Adversarial Networks. arXiv preprintarXiv:1805.08318,2018.

33. Zhangand, Y., Zheng, L. Automated Face Swapping and Its Detection. Proceedings of the 2nd International Conference on Signal and Image Processing (ICSIP), Suzhou, China, 2017, 15-19. https://doi.org/10.1109/SIPROCESS.2017.8124497

34. Zheng, L., Duffner, S., Idrissi, K., Garcia, C., Baskurt, A. Siamese Multi-layer Perceptrons for Dimensionality Reduction and Face Identification. Multimedia Tools and Applications, 2016, 75(9), 5055-5073. https://doi.org/10.1007/s11042-015-2847-3