RoTFace: A Framework for Robust and Time Efficient Face Recognition System

Shilpa Garg[1], Dr. Sumit Mittal[2], Dr. Pardeep Kumar[3]

[1]MMICT & BM, MM(DU), Mullana, Ambala, Haryana, India, E-mail: shilpa111987@gmail.com

[2]MMICT & BM, MM(DU), Mullana, Ambala, Haryana, India, E-mail: sumit.mittal@mmumullana.org

[3]Department of Computer Science & Applications, Kurukshetra University, Kurukshetra, Haryana, India, E-mail: mittalkuk@gmail.com

ABSTRACT

Now a days face recognition is commonly used applications which need security. Many authentication systems have been developed like fingerprint, palm, iris, face and many more but face recognition system is widely used as authentication system to verify the person's identity as the face is most natural way to identify the person, its uniqueness property and as it requires no touch to the screen that's why its use is rising day by day but due to spoofing attack like photo attack, video replay attack, facial mask attack and many more, face recognition requires more attention. Also due to large date set, it takes a long time to search the client id. This paper proposed a framework RotFace for robust and time efficient face recognition system. At first, deep featured are extracted of the face images using deep network ResNeT50 then classification is divided into two parts. In first part, real attack predictor is applied by using Gaussian Naïve Bayes classifier. In second part, a client id predictor is applied by using KNN. Experiment is executed on Replay Attack Dataset and achieves very good results in terms of accuracy and time. This paper also analyses the proposed framework with the existing techniques of face recognition system and performs better than the existing techniques.

Keywords – Face Recognition, Liveness Detection, Deep Network, Gaussian NaïveBayes, KNN

1. INTRODUCTION

For the security purpose, many authentication systems have been developed but face recognition system is widely used as authentication system to verify the person's identity [1] as the face is most natural way to identify the person [2], its uniqueness property [3] and as it requires no touch to the screen that's why its use is rising day by day. But due to spoofing attack like photo, 2D & 3D mask, video replay and many more [4], it is misused by the attacker.Attackers used the recorded video of the authorized person in front of sensor device to authenticate the user and have gain access to the others device. So,for robust face recognition, it is necessary to check liveness of the authentic person [5].

A robust and time efficient framework for face recognition is proposed in this paper. Robustness of the face recognition system is to check whether the person is live or not.

For the experiment Replay Attack dataset of the face videos is used. This dataset contains both real and fake videos of the persons. At first, videos of the face are converted to the frames. Then the class balancing is done to remove the imbalance problem of the real and fake frames so that it will not affect the accuracy results. After balancing real and fake image, feature extraction technique is applied to extract deep features of the image and then the system is trained using real attack predictor.

If the image is detected as fake, it will skip the image for further processing which will decrease the run time and if the image is detected as real then the image is forwarded to the next step i.e., to the client id predictor which will recognize the face and gives the client id.

As seen in the Figure 1, after checking liveness of the image, image is discarded if the image is fake which will reduce the run time. Run time is vary from system to system. This

experiment is done in python 3.7 installed on i7 8th generation, 240 SSD, 2TB HDD, 8gb ram, 4gb Graphics system.

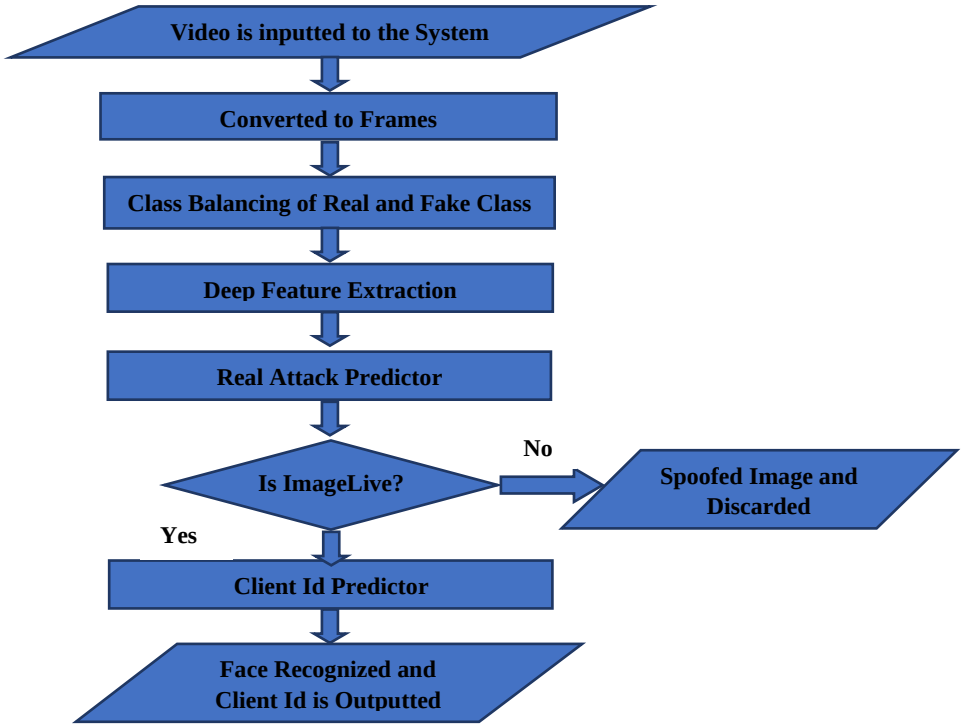

Figure 1: Proposed Framework of RotFace

2. LITERATURE SURVEY

Face recognition is widely used for the authenticity of a person these days and that is the reason, face recognition is a popular for research purpose. Many deep learning techniques for face recognition have been implemented.Mugalu et. al. proposed a web-based face recognition system in which MTCNN is used for detection and LinearSVC is used for classification and achieved 95% accuracy [6]. Sajjad et. al. suggested a hybrid approach for anti-spoofed face recognition system. This hybrid approach is two tier approaches. In tier-I,palm prints, face and finger print are integrated and in second tier, CNN is used to detect spoofing attack. This approach is applied on five dataset and achieved very promising results [7].

Linn et. al. proposed an anti-spoofing face recognition by detecting movement of eye CNN is used to extract features and classification. Experiment is done on three different dataset OWN Replay, NUAAand Replay Attack and achieved 96.5%, 98% and 95.8% accuracy respectively which is better than the exiting techniques [8]. Increase in the use of computing device and mobile needs checking of authentic user and many deep learning approached have been developed for face recognition [9]. Many deep feature extraction techniques like LBP, SIFT, DoG, HoG and SURF [10,11] etc. have been developed for feature extraction of the images similarly different deep learning classifier have been developed for detection of spoof attack.

3. CLASS IMBALANCE

For the experiment REPLAY ATTACK dataset is used from which video clips of real user and hand attacked video clips are used means video of real client in mobile or tablet is placed in front of sensor to access the device. These video clips are first converted to the frames and number of frames obtained from real and hand attacked video clips are different in number means class imbalance in real and hand attacked classes and class imbalance lead to deteriorate the accuracy. Different class imbalance ration gives different accuracy means if the ratio of real images are more than the spoofed images, it will lead to more accuracy but if the ratio real images are less than the spoofed images then the accuracy is less so it is compulsory to balance the class before classification [12].

4. FEATURE EXTRACTION

For deep feature extraction of the images five different deep features extraction techniques are used- ReSNet50, VGG16, VGG19, InceptionV3 and DenseNet121.These deep networks are the pretrained weighted networks. ReSNeT is referred to Residual Neural Network given by Kaiming et. al. in 2015 and perform better than VGG for image processing [13].

VGG16 is type of CNN model given by Simonyan et. al. VGG16 isbetter than AlexNet suggested by the authors and gives better accuracy for image processing[14].VGG19 is improvement over VGG16 and comprises of some more layers than VGG 16 and deeper than VGG16 [15]. InceptionV3 is also a deep CNN architecture used for image processing given by Google and developed as a challenge for imagenet recognition [16]. DenseNet is proposed in 2018 by Huang et al., suggested that this network is also an improvement over some other state of the art and worked in feed forward direction [17].

5. REAL ATTACK PREDICTOR

Real attack predictor is trained and used to predict the liveness of the person by using the deep features extracted from different techniques and machine learning algorithm - Gauss Naïve Bayes and Random Forest. Gauss Naïve Bayes is commonly used for image classification [18] and Random Forest performs very well to predict liveness of recognition system [19]. This trained Real Attack Predictor is used to predict the imputed image as real or spoofed during testing phase.

6. CLIENT ID PREDICTOR

Client id predictor is trained using deep feature of the face images and two machine learning algorithm – k-Nearest Neighbors (KNN) and Decision Tree. k-Nearest Neighbor and Decision Tree are good classifier for image processing [20]. After detecting liveness of the image, only lived person images is send to the client id predictor for face recognition which predict the client id in testing phase.

7. METHODOLOGY

Methodology discussed the dataset used, performance evaluation metrics and experimental setup and results.

7.1. DATASET

REPLAY-ATTACK dataset [21] contains recorded video clips of 50 persons in two varied conditions - adverse and controlled. In controlled condition, background is illuminated with fluorescent lamp light and uniform whereas in adverse condition, background is non uniform and illuminated with day light. This dataset contains two type of attack frame - Fixed and Hand. In fixed attack, a hard copy of the image of a person is placed at fixed position in front of camera whereas in Hand attack, videoclips from Mobile and tablet is hold by the operator's hand.

Figure 2: Sample Images of the Replay Attack Dataset. Top Row shows the Controlled Condition and bottom row shows the adverse condition. From left to right- Real Image, mobile phone and Tablet Attack

For this experiment, Real and Hand attack recorded video clips of 20 persons are used and converted to image frames. From these frames, 50 real frames of each client are chosen and 50 hand attacked frames of each client are chosen so total 2000 images are used for experimental work of proposed framework in which 1000 images are real images and 1000 images are attacked images of 20 clients.

7.2. PERFORMANCE EVALUATION METRICS

For measure and compare RoTFace performance, two metrics Accuracy and F1 score are used [22].Accuracy is used to calculate true prediction and represented by equation 1.F1 Score is also an accuracy measure and defined as weighted mean of recall and precision represented in equation 2.

$$Acc = \frac{t_{pv}+t_{nv}}{t_{pv}+t_{nv}+f_{pv}+f_{nv}} \quad (1)$$

$$F1 = \frac{2t_{pv}}{2t_{pv}+f_{pv}+f_{nv}} \quad (2)$$

Where t_{pv} predicts real images as real, t_{nv}predicts attacked images as attacked, f_{pv}predicts attacked images as real, f_{nv} predicts real images as attacked.

7.3. EXPERIMENT SETUP AND RESULTS

For experiment, 60:40 ration is followed for training and testing. From 2000 images of the total images, 600 real and 600 attacked i.e., total1200 images are used for training the predictor and 400 real and 400 attacked i.e., total 800 images are used for testing.
In experiment, five different feature extraction techniques VGG16, VGG19, inceptionV3, DenseNet121, ReSNet50 are used to extract deep features of the image. Then Gauss Naïve Bayes classifier is used as real attack predictor and then K Nearest Neighbour(KNN) is used as client id predictor similarly after deep feature extraction, Random Forest is used as real attack predictor and Decision Tree is used as client id predictor. These approaches are represented as VGG16_GK, VGG16_RFDT, VGG19_GK, VGG19_RFDT, InceptionV3_GK, InceptionV3_RFDT,DenseNet_GK, DenseNet_RFDT, ReSNeT_RFDT and RoTFace (ReSNet_GK) and gives 97.64%, 94.132%, 98.58%, 96.25%, 94.63%, 90.201%, 98.37%, 91.939%, 90.523%, 99.7% accuracy respectively.
Table 1 shows the accuracy, F1 score and Run time excluding training time (in sec) results of different approached used in experiment.
As shown in table, run time of proposed approach is too less than the other except InceptionV3_GK. Run time of inceptionV3_GK and RoTFace is 1.11 seconds but accuracy of proposed approach is much better than the InceptionV3_GK so the proposed approach is performing better than the other existing techniques.

Table 1:comparison results showing the accuracy, F1 score and Run time

Approach Used	Accuracy (in %age)	F1 score (in %age)	Run Time (in Seconds)
VGG16_GK	97.64	96.37	1.353
VGG16_RFDT	94.132	93.986	5.874
VGG19_GK	98.58	98.38	1.356
VGG19_RFDT	96.25	96.138	5.878
InceptionV3_GK	94.63	93.97	1.11
InceptionV3_RFDT	90.201	90.276	6.152
DenseNet_GK	98.37	96.81	3.28
DenseNet_RFDT	91.939	91.956	16.293
ReSNet50_RFDT	90.523	90.976	17.581
RoTFace (ReSNet50_GK) Proposed	**99.7**	**99.16**	**1.11**

Figure 3 Shows the graphical representation of accuracy and F1 score and clearly shows that proposed framework is performed better than the other approach used. Figure 4 shows graphical representation of the run time comparison of proposed approach with other

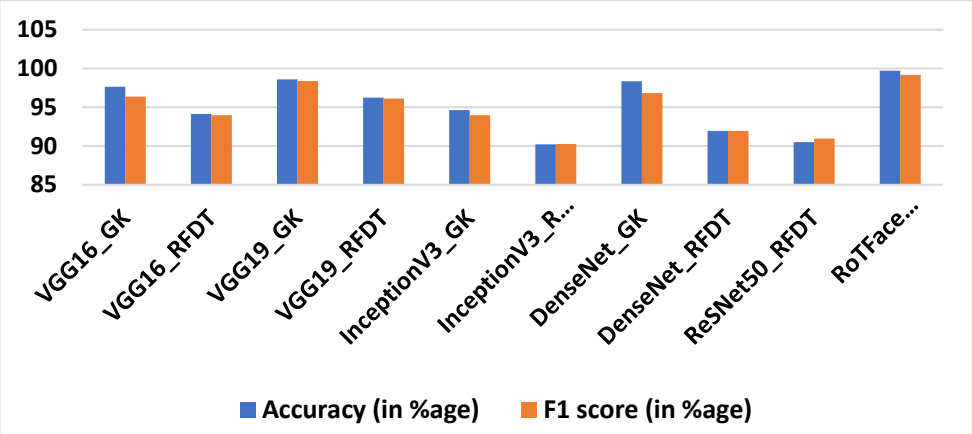

Figure 4:Accuracy and F1 Score comparison Graph

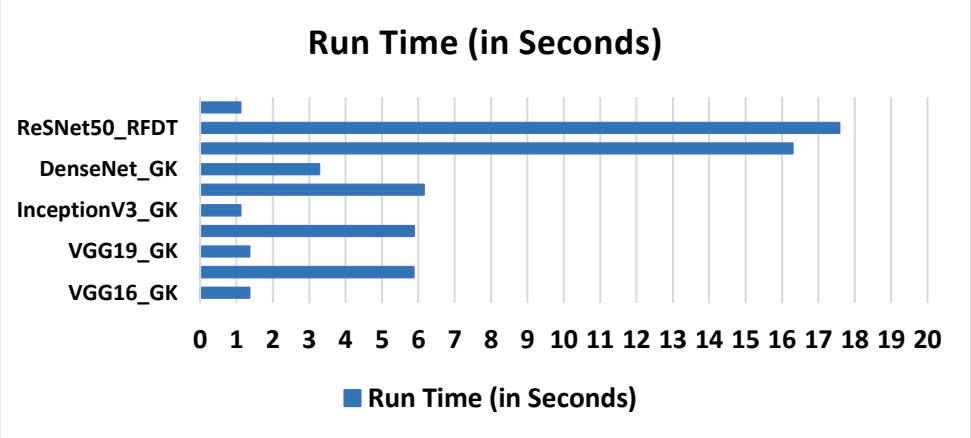

Figure 3:Graph shows the Run time comparison

approach.

8. CONCLUSION

This paper presents a robust and time efficient framework for face recognition system. The proposed approach RoTFace extracts the deep feature using ReSNet50 and Gauss Naïve Bayes and KNN classifiers are used to predict the spoofing attack and client id respectively. Also, to analyses the result with existing techniques, four other different deep feature technique is applied i.e., VGG16, VGG19, InceptionV3, DenseNet121 with Gauss Naïve Bayes and Random Forest as Real Attack predictor and KNN and Decision Tree as Client Id predictor. RoTFace gives 99.7% accuracy, 99.16% F1 score and runs in 1.11 seconds only and performed better than the existing techniques in terms of both run time and accuracy.As a future work try to work on real time face recognition and will try to reduce the run time of real time face recognition system.

REFERENCE

1. Thepade, S., Jagdale, P., Bhingurde, A. and Erandole, S., 2020, February. Novel Face Liveness Detection Using Fusion of Features and Machine Learning Classifiers. In *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)* (pp. 141-145). IEEE.

2. Garg, S., Mittal, S. and Kumar, P., 2019. Performance Analysis of Face Recognition Techniques for Feature Extraction. Journal of Computational and Theoretical Nanoscience, 16(9), pp.3830-3834.

3. Garg, S., Mittal, S., Kumar, P., &Athavale, V. A. (2020, February). DeBNet: Multilayer Deep Network for Liveness Detection in Face Recognition System. In 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 1136-1141). IEEE.

4. Sajjad, M., Khan, S., Hussain, T., Muhammad, K., Sangaiah, A.K., Castiglione, A., Esposito, C. and Baik, S.W., 2019. CNN-based anti-spoofing two-tier multi-factor authentication system. *Pattern Recognition Letters*, *126*, pp.123-131.

5. Bharadwaj, S., Dhamecha, T.I., Vatsa, M. and Singh, R., 2013. Computationally efficient face spoofing detection with motion magnification. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 105-110).

6. Mugalu, B.W., Wamala, R.C., Serugunda, J. and Katumba, A., 2021. Face Recognition as a Method of Authentication in a Web-Based System. *arXiv preprint arXiv:2103.15144*.

7. Sajjad, M., Khan, S., Hussain, T., Muhammad, K., Sangaiah, A.K., Castiglione, A., Esposito, C. and Baik, S.W., 2019. CNN-based anti-spoofing two-tier multi-factor authentication system. *Pattern Recognition Letters*, *126*, pp.123-131.

8. Linn, P.P.P. and Htoon, E.C., 2019, November. Face Anti-spoofing using Eyes Movement and CNN-based Liveness Detection. In *2019 International Conference on Advanced Information Technologies (ICAIT)* (pp. 149-154). IEEE.

9. Saha, S., Xu, W., Kanakis, M., Georgoulis, S., Chen, Y., Paudel, D.P. and Van Gool, L., 2020. Domain agnostic feature learning for image and video-based face anti-spoofing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 802-803).

10. Sun, W., Song, Y., Chen, C., Huang, J. and Kot, A.C., 2020. Face spoofing detection based on local ternary label supervision in fully convolutional networks. *IEEE Transactions on Information Forensics and Security*, *15*, pp.3181-3196.

11. Patel, K., Han, H., Jain, A.K. and Ott, G., 2015, May. Live face video vs. spoof face video: Use of moiré patterns to detect replay video attacks. In *2015 International Conference on Biometrics (ICB)* (pp. 98-105). IEEE.

12. Vo, T., Nguyen, T. and Le, C.T., 2019. A hybrid framework for smile detection in class imbalance scenarios. *Neural Computing and Applications*, *31*(12), pp.8583-8592.

13. Sharma, N., Jain, V. and Mishra, A., 2018. An analysis of convolutional neural networks for image classification. *Procedia computer science*, *132*, pp.377-384.

14. Fu, Y. and Aldrich, C., 2019. Flotation froth image recognition with convolutional neural networks. *Minerals Engineering*, *132*, pp.183-190.

15. Shaha, M. and Pawar, M., 2018, March. Transfer learning for image classification. In *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 656-660). IEEE.

16. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A., 2015. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).

17. Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q., 2017. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).

18. Rafique, A.A., Jalal, A. and Ahmed, A., 2019, August. Scene Understanding and Recognition: Statistical Segmented Model using Geometrical Features and Gaussian Naïve Bayes. In *IEEE conference on International Conference on Applied and Engineering Mathematics* (Vol. 57).

19. Thepade, S.D., Chaudhari, P., Dindorkar, M., Bang, S. and Bangar, R., Improved Face Spoofing Detection Using Random Forest Classifier with Fusion of Luminance Chroma.

20. Dino, H.I. and Abdulrazzaq, M.B., 2019, April. Facial expression classification based on SVM, KNN and MLP classifiers. In *2019 International Conference on Advanced Science and Engineering (ICOASE)* (pp. 70-75). IEEE.

21. Chingovska, I., Anjos, A. and Marcel, S., 2012, September. On the effectiveness of local binary patterns in face anti-spoofing. In *2012 BIOSIG-proceedings of the international conference of biometrics special interest group (BIOSIG)* (pp. 1-7). IEEE.

22. Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.