Shibel et al, 2022

Volume 8, p.p. 61-78

Received: 11th May 2022

Revised: 19th June 2022, 01st July 2022, 08th July 2022

Accepted: 08th July 2022

Date of Publication: 15th July 2022

DOI- https://dx.doi.org/10.20319/lijhls.2022.8.6178

This paper can be cited as: Shibel, A. M., Ahmad, S. M. S., Musa, L. H., & Yahya, M. N. (2022). Deep Learning Detection of Facial Biometric Presentation Attack. LIFE: International Journal of Health and Life-Sciences, 8, 61-78.

DEEP LEARNING DETECTION OF FACIAL BIOMETRIC PRESENTATION ATTACK

Ahmed Muthanna Shibel
Department of Computer and Communication Systems Engineering, Faculty of Engineering, University Putra Malaysia (UPM), Malaysia
muthanaa83@gmail.com

Sharifah Mumtazah Syed Ahmad
Department of Computer and Communication Systems Engineering, Faculty of Engineering, University Putra Malaysia (UPM), Malaysia
s_mumtazah@upm.edu.my

Luqman Hakim Musa
Department of Computer and Communication Systems Engineering, Faculty of Engineering, University Putra Malaysia (UPM), Malaysia
sr71luqman@gmail.com

Mohammed Nawfal Yahya
Department of Computer and Communication Systems Engineering, Faculty of Engineering, University Putra Malaysia (UPM), Malaysia
mohamedaljaff785@gmail.com

Abstract

Face recognition systems have gained increasing importance in today's society, which applications range from access controls to secure systems to electronic devices such as mobile *phones and laptops. However, the security of face recognition systems is currently being threatened by the emergence of spoofing attacks that happens when someone tries to unauthorizedly bypass the biometric system by presenting a photo, 3-dimensional mask, or replay video of a legit user. The video attacks are perhaps one of the most frequent, cheapest, and simplest spoofing techniques to cheat face recognition systems. This research paper focuses on face liveness detection in video attacks, intending to determine if the provided input biometric samples came from a live face or spoof attack by extracting frames from the videos and classifying them by using the Resnet-50 deep learning algorithm. The majority voting mechanism is used as a decision fusion to derive a final verdict. The experiment was conducted on the spoof videos of the Replay-attack dataset. The results demonstrated that the optimal number of frames for video liveness detection is 3 with an accuracy of 96.93 %. This result is encouraging since the low number of frames requires minimal time for processing.*

Keywords

Biometric, Deep Learning, Presentation Attack, Face Liveness Detection, CNN

1. Introduction

The term "biometric" refers to the automated identification of people based on their physiological and / or behavioral features [1]. DNA, ear, face, fingerprint, gait, iris, keystroke, smell, palm print, retinal scan, signature, and voice are examples of biometric traits [2]. The biometric characteristic that is chosen is determined by the biometric application's goal. Biometric characteristics are features that are unique to a person and may be used to identify people in a biometric system [3]. Face recognition, for example, is often used in security and immigration verification systems, such as border entry control [4]. Facial recognition technology is among the most practical approach for verifying people's identities. Researchers from a variety of fields, such as image processing, computer vision, and pattern recognition, have combined their efforts to improve the performance of biometric systems [5], allowing biometrics to be used in a wide range of applications, including forensics, border and access control, surveillance, and online commerce. Because of their widespread use, biometric systems are susceptible to a broad variety of increasingly sophisticated assaults. Thus, the creation of strong counter-measures is necessary. In the history of biometric research, the face is the second most widely used biometric in terms of market share, behind only fingerprints [6]. As a result, automated facial authentication systems have been widely deployed. However, the ability of this fast-developing technology to withstand external assaults has become a major concern. Spoofing, in particular, is an assault in which a legitimate user's picture, video, or mask is put in front of a facial recognition system in an attempt to obtain access [7]. The repetition of facial recognition system spoofing assaults has become a major source of worry within the biometric community. In this situation, a legitimate user's facial biometric data may be acquired without physical touch by capturing it with a camera or downloading it over the internet [8]. Thus, it is of paramount importance to protect facial recognition systems against presentation attacks. As a result, developing effective countermeasures is a necessity. The emphasis of this research paper is on detecting video presentation assaults. Because of video attacks are amongst the most frequent, cheapest, and simplest spoofing techniques to cheat face recognition systems [9]. It occurs when someone attempts to impersonate another by producing a false biometric characteristic (replay video) of the user and presenting it to the sensor, thus impersonating the actual user. At the present, the research on using deep learning on frames of video in detecting facial biometric presentation attacks are limited and there are not many studies on using deep learning in the detection of video presentation attacks especially face liveness detection on video frames sequence. The issue when it comes to video is the accuracy vs efficiency, and more frames mean more time needed for face liveness detection [10][11]. So, it is necessary to find out what is the optimal number of frames. To do that we should investigate the optimal number of frames, by classifying the video frames and using a deep learning algorithm as a classifier for face liveness detection of video presentation attacks to find out if we can get fewer frames for detection, why do we need more for analysis! this happens by starting with a very low number of frames and keep increasing and see the optimum number of frames should be used to do the face liveness detection with good accuracy and less time. In terms of the detection method used in this research, there are various approaches have been done for detection in biometrics we presented some of the previous literatures that have done by using deep learning, the use of deep learning which is a type of machine learning to extract features from images, videos and others types of data to classify the real or spoof person led to make an evolution in face liveness detection because it gave us the ability to do detection with high accuracy [12] .In this research we used a deep residual network, the deep residual network [13] was probably the most significant contribution to the field of computer vision and deep learning in recent years. ResNet allows us to train hundreds or even thousands of layers while still achieving excellent results. Many computer vision applications other than picture classification, including object identification and face recognition, have improved as a result of their strong representational capabilities. Numerous in the research community have delved into the secrets of ResNet's success since it blew people's minds in 2015, and many improvements in the design have been made. We used ResNet50 in this research which is a variation of the ResNet model, having 48 Convolution layers, 1 MaxPool layer, and 1 Average Pool layer. ResNets were first employed to improve image recognition, but the framework may now be used to improve accuracy in non-computer vision activities as well. This research paper targeted the face liveness detection in video attacks, intending to determine if the provided characteristic came from a live sample, by extracting frames from the videos and classifying them by using a deep learning algorithm. The primary aim of this paper is to find the optimum number of frames that should be used for face liveness detection in video attacks. To achieve this goal the objectives below have been identified:

- To design an effective classifier by using deep learning.
- To investigate the optimal number of frames that should be used from videos to get the best result.
- To evaluate the performance of the new system of face liveness detection in terms of its accuracy and the time for detection.
- To improve the effectiveness of the face liveness detection system and achieve high accuracy in the result.

The scope of this research paper is to design an effective classifier by using ResNet50 which is a deep learning algorithm to detect the liveness of the face in video presentation attacks by extracting some frames firstly from videos and then classifying frames, also to investigate the optimal number of frames to get high accuracy in detection. A Replay Attack database used in this research paper is a public facial database, the samples utilized from the database are video attacks by mobile and frames extracted from these videos.

2. Literature Review

For facial liveness detection, numerous systems have been created in different fields, including police investigations, airport and homeland security, clinical testing, and human resource offices in organizations and businesses. This section of the paper presented the face recognition technology and the use of deep learning algorithms for face liveness detection.

2.1. Face Recognition

Face recognition is one of the physiological biometric techniques that has gotten a lot of interest in each academic and industrial field [14]. Face recognition is most frequently utilized as most biometric identification techniques. It is a mature technology that provides a quick, reliable, convenient, and low-cost method of identifying people. It has a lot of applications, ranging from security applications like passport check-in at airports also in consumer levels like logs in to smartphones and laptops. In different categories from systems based on three-dimensional face scans to a system that deals with videos or even images [15], the established methods range in complexity as well as hardware and software requirements. Face recognition, unlike fingerprint and iris recognition, which rely on high-resolution pictures, is based on data that is readily available in the public domain. It is a top contender method for biometrics identifications that needs reliable, real- time, and also unobtrusive user authentications without the use of specialist hardware in terms of performance and acceptance. Face recognition is susceptible to various spoofing attacks in terms of circumvention [16]. Because of this flaw, face recognition authentication is currently mainly limited to applications with minimal security needs or applications in well-controlled settings. Generally, there are three types of face spoofing attacks: mask, images, and video attacks. In this research paper, the static approach examines a 2D face picture from video frames and extracts key elements that may be utilized to determine if the picture is genuine or not.

2.2. Deep Learning for Face Liveness Detection

Deep learning is a type of machine learning that extracts features from data such as images, texts, and audio [17]. This technique extracts traits that are learned and can be employed in subsequent tasks. Figure 1 illustrates the key distinction between machine learning and deep learning in classification tasks. While the approach to generating good estimation in machine learning should be taught to the algorithm by providing more information, the algorithm in deep learning may learn this by processing data. In simple terms, feature extraction in machine learning is done by humans, but deep learning models figure it out on their own [18].

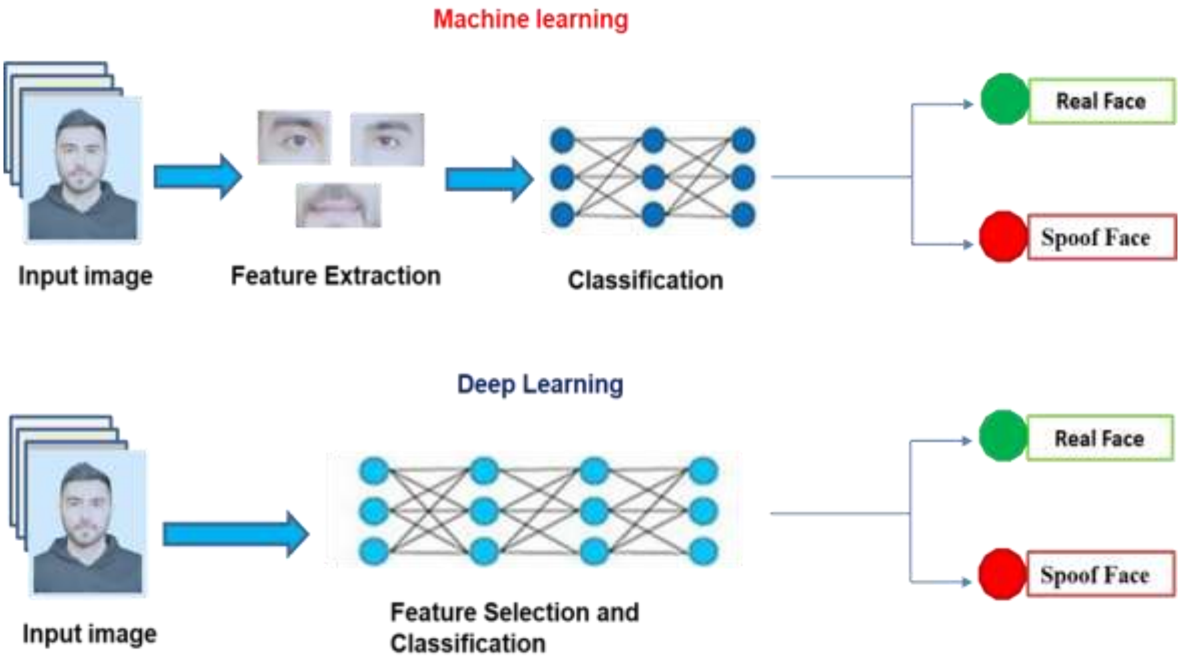


Figure 1: *Difference between Machine Learning and Deep Learning in classification.*

(***Source***: *Self compiled)*

Although deep learning is a relatively new discipline of machine learning, it has grown rapidly and has become a hot issue in the artificial intelligence sector [19]. The growing size of datasets is one of the factors driving this trend, in addition to the Hardware growing increasingly capable as technology advances. Researchers can now work on large datasets to acquire better outcomes from their tests as a result of this. Increased memory size and the appearance of cloud technology are two major reasons for deep learning's success. Following the researchers' experiments on extremely big datasets, another issue developed regarding how to process and store such massive volumes of data. The manufacture of computers with higher characteristics or the use of Cloud properties, however, has solved the issue. Last but not least, growing accuracy, complexity, and real-world influence are all factors in deep learning's success. People did not employ deep learning much when it was introduced since larger datasets were required to achieve better results. A huge dataset should be provided to extract every abstract characteristic to employ the algorithm in the task we need. Deep learning deployment has gotten considerably easier and more effective since the "Big Data" era [20]. In terms of face liveness detection, some techniques take a small sequence of frames instead of traditional neural networks which accept a fixed-sized vector as input, a single frame in this case, and produce an outcome for each. This does not happen with Recurrent Neural Networks (RNN) since it allows to operate over a sequence of frames, or a small video, as input vectors [21]. In this research paper, the ResNet50 was used for presentation attack detection as a static approach to do the face liveness detection for the video frames.

3. Methodology

The proposed system that we created for face liveness detection in video presentation attacks consists of different phases including the system architecture, video segmentation, the data pre-processing, and the distribution of the Replay Attack database, which are explained below in this section.

3.1. System Architecture

The first step in our system is to input video as shown in Figure 2 the system architecture of our methodology then extracts frames from videos by using Video-LAN Client media player program software.

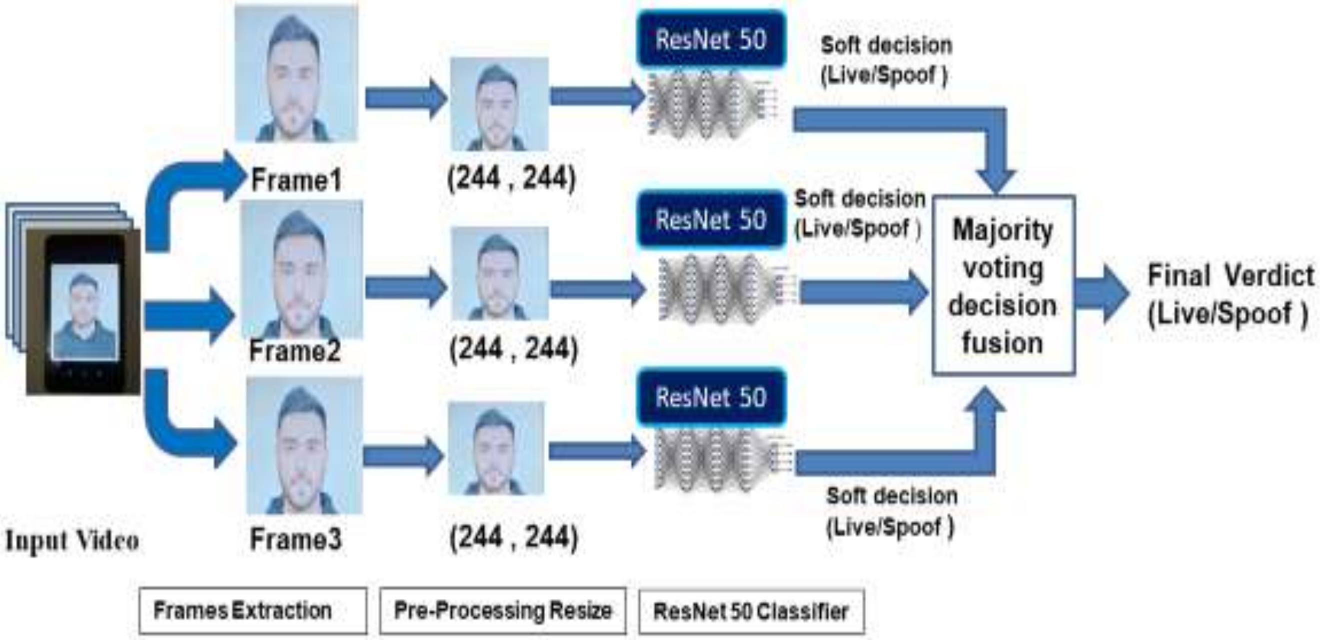


Figure 2: *Example of 3 Frames Architecture for Face Liveness Detection.*

(***Source****: Self compiled)*

Subsequently, we resized all the frames that we used in our methodology for videos we used to (224,244) pixels because in computer vision resizing images is an important preprocessing step the main reason for resizing frames is that the machine learning models are mostly trained faster on smaller images. After that, we used a deep learning algorithm which was

ResNet50 as a classifier and the output from the Resnet50 classifier was a soft decision (live / spoof) from an individual frame, in our methodology we used three classifiers in total. Then, the output of each classifier was going to the next step which was the majority voting decision fusion, the majority voting decision fusion simply can clarify it which means from the odd numbers of frames we used for making the final decision the output depends on the more than half input frames, For example, if we have 3 frames for two scenarios as a hypothesis the first one if two of frames real and the third one is spoof for the same video that was used in testing the final decision will be a live face, and for the second scenario if we have 2 frames spoof and the third one real so the results will be spoofing face. The final verdict was depending on which more half-frames we have real or fake to get the final result which is live or spoofing face. The use of a deep learning algorithm works better when we have a huge amount of samples so we used more than one thousand frames to achieve our goal of high accuracy in the results, and the number of videos we used (100 videos) covered all the 50 clients from the Replay attack database to achieve high accuracy in face liveness detection method and make our model able to predict the real person or spoof attack even though for any data that did not see before, furthermore, the classifying of video frames in our methodology depended on the odd number of frames which was starting from 3 frames and then increasing until 7 to investigate the sufficient number of frames that should be used to achieve high accuracy and less time in face liveness detection. The sampling happened in the beginning, middle and the end of clips, and several frames were taken to cover different sequences per video, as a result, that led to giving good results in terms of face liveness detection. The video samples were taken from the Replay attack database which is a popular database that consists of 1300 videos in different scenarios of attack (photo and video attacks) [22], so we chose just videos attacks scenario because our research focused mainly on video replay attacks using mobile or any digital screen. Also, many samples of frames have been taken for testing to check the system performance in terms of detection to give a good result when it comes to prediction, after that we did the test accuracy and our analysis for the results.

3.2. Video Segmentation Phase

The methodology of this research focuses mainly on mobile video attack scenarios so we extracted the mobile videos attacks for all the 50 clients in the Replay attack database which total

100 videos. Figure 3 shows real and spoofing attack examples from the replay attack database by using a mobile screen.

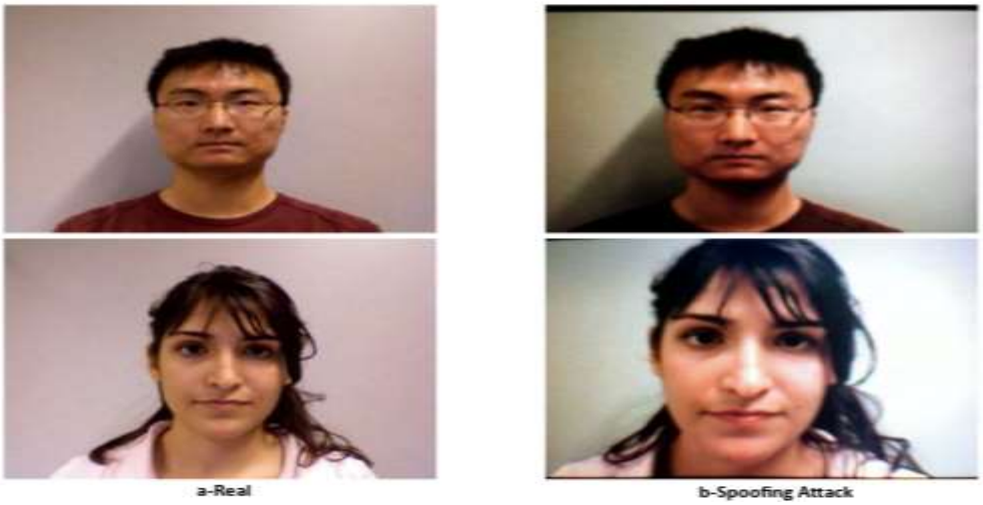


Figure 3: *The real and spoofing attack examples from the replay attack database.*

(***Source***: *Self compiled)*

To extract many frames from videos that were used from the database we covered all the 50 clients in the Replay-attack database to make the system of detection more accurate in detection. To get many frames from videos we chose 100 Videos for clients and that led to getting more than one thousand images for the method because we applied the static method of detection on videos. The program VideoLAN Client media player program was used for video segmentation. The calculation that was used to get frames from the videos was by extracting one frame for every 10 frames per second from videos, for example, if the video contains 30 FPS the total number of frames will be 3, but in the Replay database each video as minimum long 9 seconds, so as the average the number of frames will be from 23 to 38 frames for each video sample used for all the clients targeted from the Replay attack database. The frames were extracted from the start, middle, and end of clips, and many frames were taken to achieve a better result in face liveness detection accuracy.

3.3. The Data Pre-processing Phase

In this part of our research, because the implementation of our research required a high computer specification, we used Google Colaboratory for this task which is a cloud service for machine learning instruction and research based on Jupiter Notebooks. It includes a fully configured deep learning runtime as well as free access to a powerful GPU [23]. Google collab is a so-called Jupiter notebook in the cloud which is a web tool that allows us to create and share documents including live code, equations, and visualizations. It requires no setup to get started. We gain access to powerful hardware through Google Colaboratory, which reduces the training time for our deep learning model. So the classifier model was created utilizing Google Colab as well as the Keras API. Firstly, we imported all libraries that we need to create our system and then we organized our data into three categories which are training, validation, and testing sets, after uploading them to our Google drive we imported them. Secondly, as we mentioned previously, we resized all the frames that we used in our methodology to (224,244) pixels, in computer vision resizing images is an important preprocessing step because the machine learning models are mostly trained faster on smaller images [24]. Furthermore, many deep learning models architectures demand that the photos should have the same size, although the raw gathered images may differ in size but the photos must be adjusted to a predetermined size before being fed into the CNN [25], the images must be resized to a fixed size for less deformation of the features and the patterns inside the images. Thirdly, the data was loaded to the classifier and the batch size was set to 3. The batch size is a hyperparameter that specifies how many samples must be processed before the internal model parameters are updated. The batch is implemented as a for-loop that iterates across one or more samples and generates predictions. The predictions are compared to the expected output variables after the batch, and an error is calculated. The update procedure is used to enhance the model based on this inaccuracy [26]. In terms of the number of epochs which is a hyperparameter that controls how many times the learning algorithm runs over the whole training dataset. The number of epochs can be set to anything from one to infinity. For a learning algorithm, the batch size and number of epochs must be specified. There are no hard and fast rules for configuring these parameters. We must experiment with various values to see which one works best for solving the problem with high accuracy [27]. In this research, we tested a different number of epochs starting from 10,11,16,18,20,25,32 respectively and then we fixed it to 25 because was the best value for training our model to get a high accuracy result. Lastly, the ResNet50 algorithm which was used in our research as a classifier trained on the data that we used from the Replay attack database to make two classifications which are represented by a real person or spoofing attack to achieve our goal for face liveness detection and make our model able to predict the real person or spoofing attack even though through the data that have not seen.

3.4. The Distribution of the Replay Attack Database:

For training, validating, and testing purposes for our model we distributed the data used in our research from the Replay attack database into three sets shown in figure 4 which are:

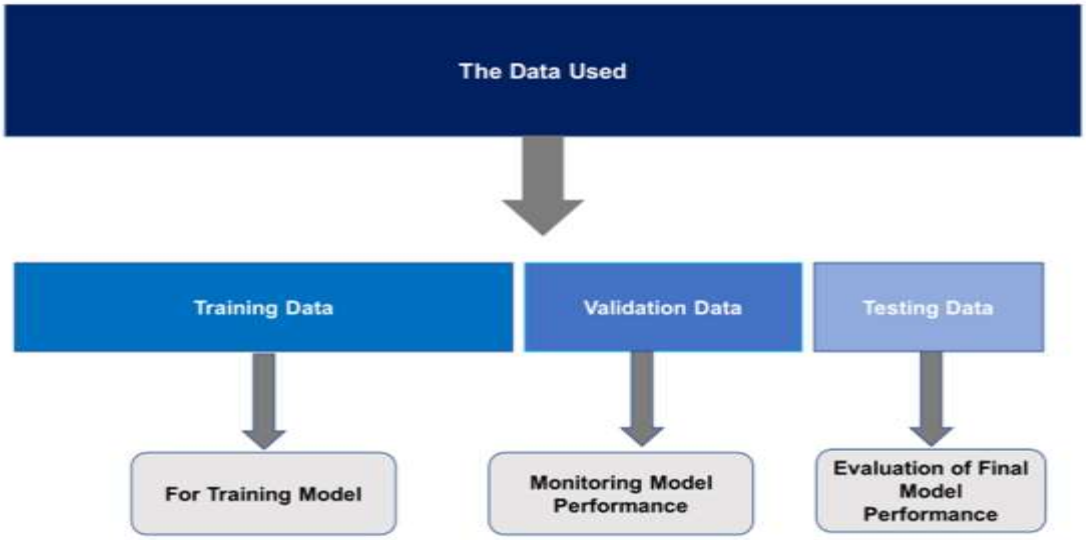


Figure 4: *The data distribution and the purpose of each set.*

(***Source***: *Self compiled)*

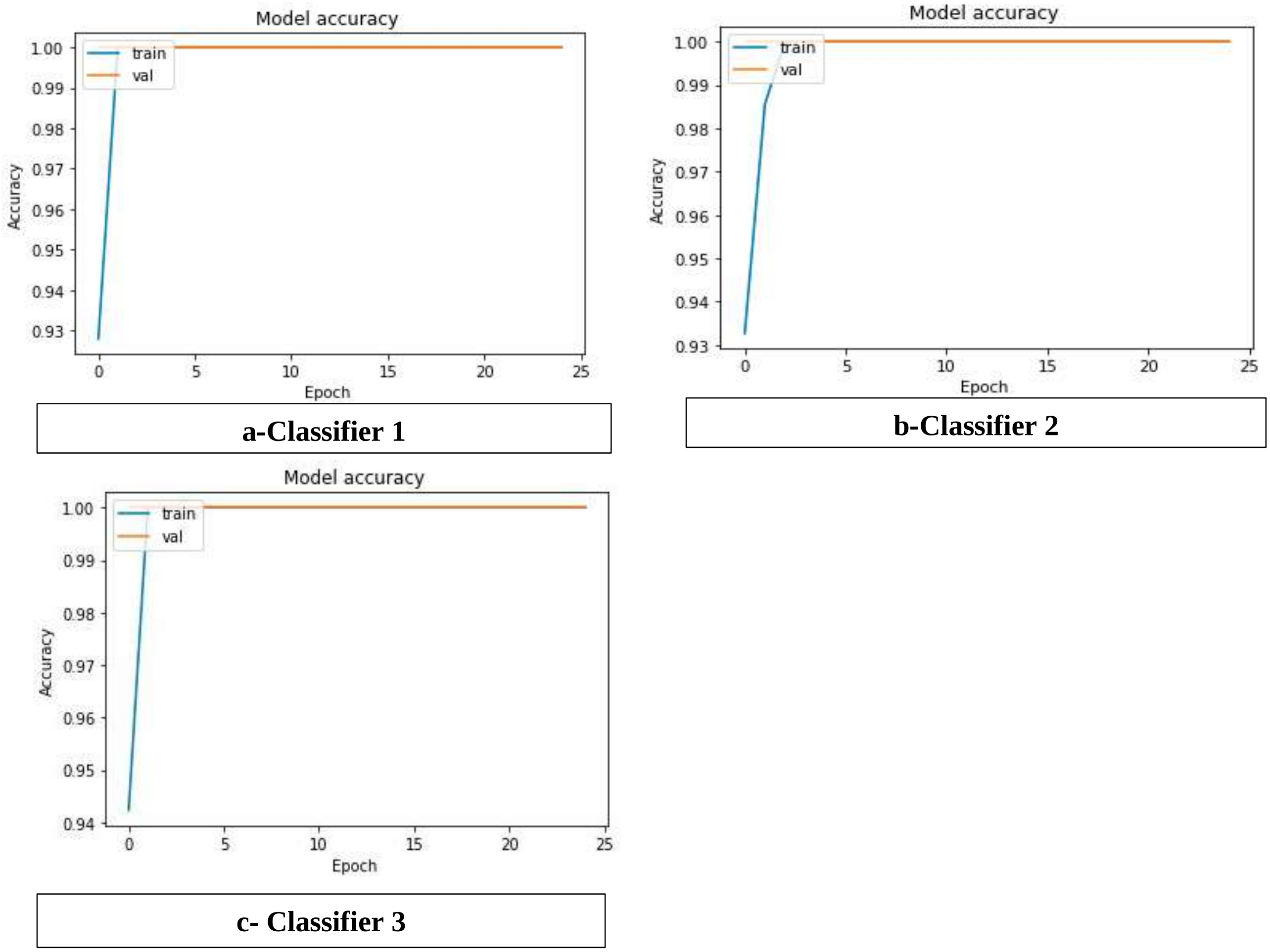
The primary goal of having three independent datasets is to ensure that the model can generalize by properly predicting unknown variables. In each classifier we created, the distribution of data was 70% for the training set, 17%for validating set, and 13% for the testing set.

4. Results and Discussions

In this section of the research paper, we presented the results that have been achieved in our system after testing, and after that, the discussion included comparing our results with other researcher's results.

4.1. The Testing Results

In each neural network, we used 208 frames for training, 52 for validation, and 40 for testing. The two classes were implemented to classify if it is real or spoof attack. For the NN1, NN2, and NN3 the three classifiers that we created, different numbers of epochs were tested to achieve the best accuracy and we found the best one is 25 with 3 batch size and we can see accuracy result for each classifier we created in figure 5 below. The accuracy in training and validation sets was very good.



a-Classifier 1

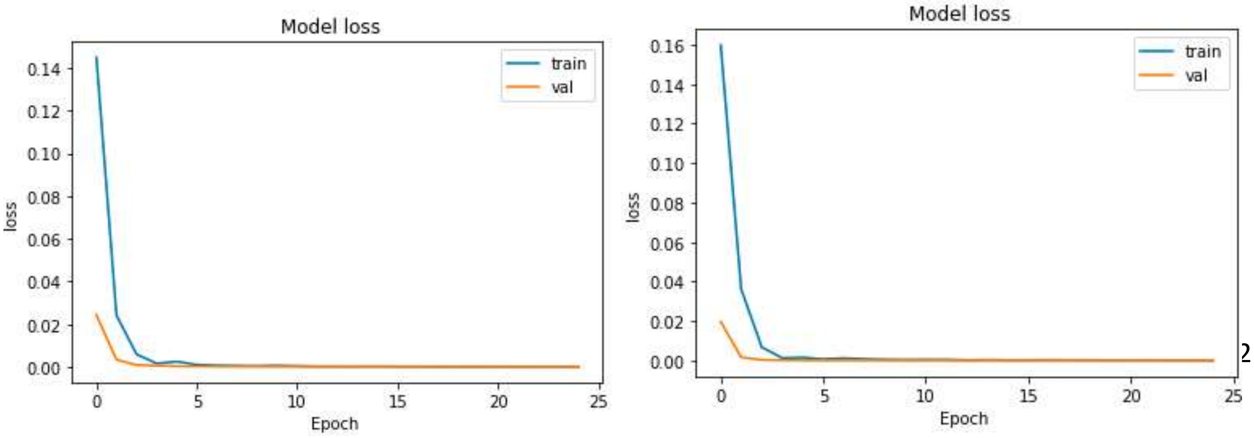
b-Classifier 2

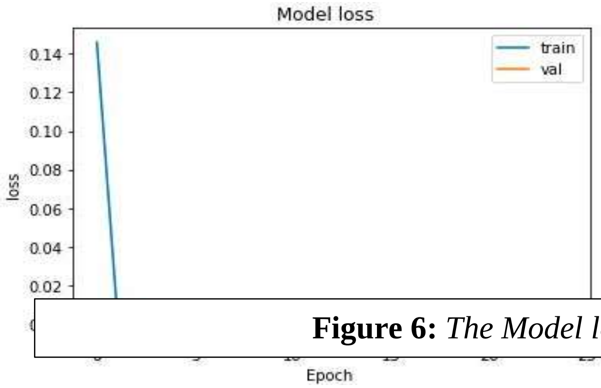
c- Classifier 3

Figure 5: *The Accuracy result for the three classifiers***.**

(***Source****: Self compiled)*

Figure 6 below shows here the training and the validation loss in our models doing good, and there is no underfitting or overfitting, these models are considered a good fit for all the three classifiers we created.





a- The Model loss For NN1.

b- The Model loss For NN2.

Figure 6: *The Model loss For NN1, NN2, and NN3.*

(***Source****: Self compiled)*

4.2. Results and Discussion

For prediction results and what we got, we chose 10 videos for testing from the Replay attack database and we assumed two scenarios for doing the majority voting to the odd number of frames that we used for classifying 3,5, and 7 which are real and spoof attack for the actual input and to each classifier. For example, as we mentioned previously if the first one contained real ,fake ,and real frames for the same video we chose as input, and the second one fake, real, and fake ,and compared with the actual input video we did the majority voting for the correct frames with the value of accuracy and then we calculated the accuracy value for each video we used in the testing set and we found the average accuracy result for each classifier we used in all cases for 3,5, and 7 then we calculated the accuracy for them and made a comparison between them to prove that what is the optimal number of frames should be used to give the accurate results for face liveness detection , as a result, we found that the 3 frames are sufficient instead of going to others because the 5, and 7 frames made our results accuracy of face liveness detection goes down, figure 7 below show the accuracy result for using 3,5, and7 frames for detection.

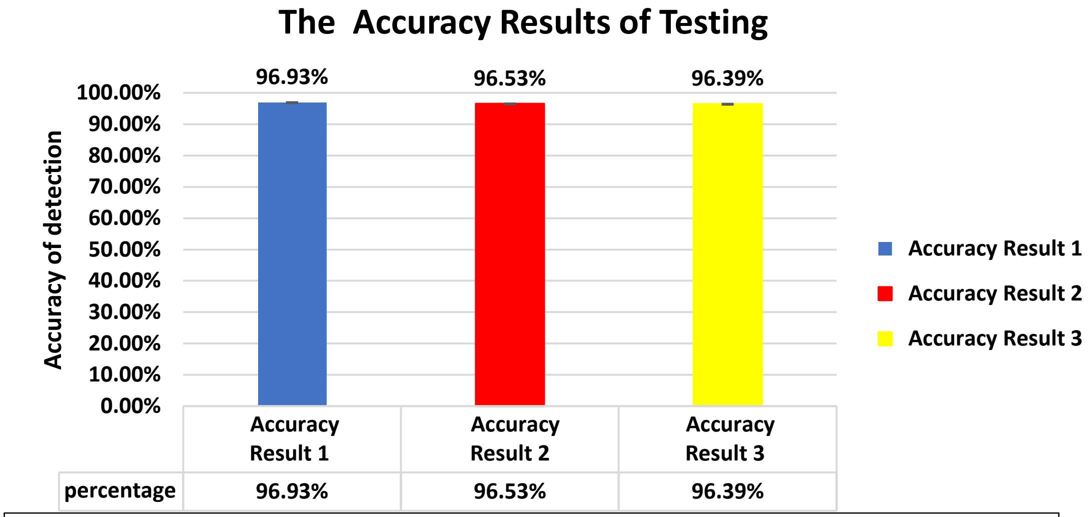


Figure 7: *The accuracy result for using 3,5, and 7 frames for detection.*

(***Source**: Self compiled)*

Results 1 in the paragraph above refer to using 3 frames for detection, result 2 refers to using 5 frames, and results 3 refers to using 7 frames for face liveness detection. What we conclude from here the ideal number of frames that should be used for detection is 3 frames to get a high result and reduce the time of detection because less number of frames is better instead of going to 20 frames [11], also compared with using a single frame in terms of accuracy detection the use of three frames was better [10], figure 8 below shown the comparison with our model results ResNet50 and the other research results we targeted.

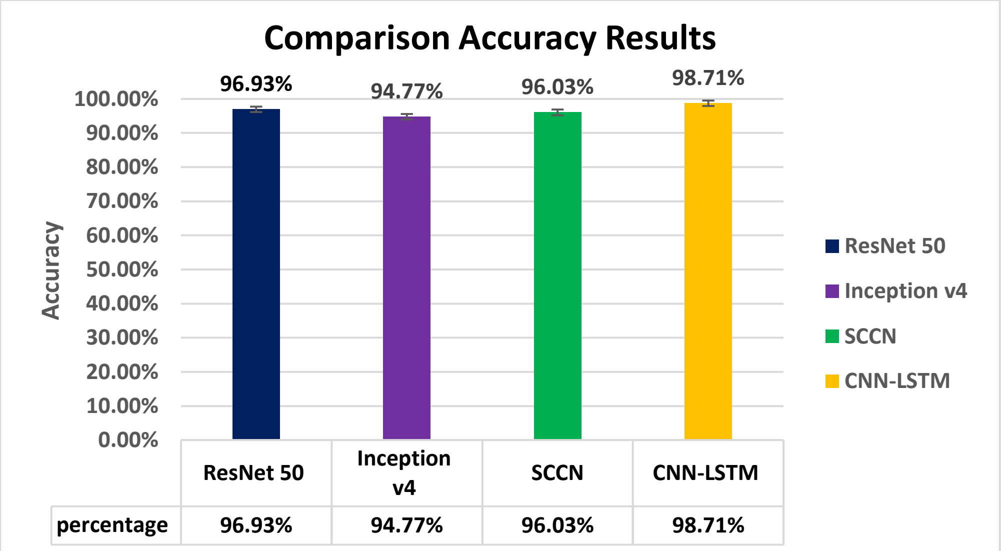


Figure 8: *The comparison with our model results ResNet50 and the other research results.*

(***Source****: Self compiled)*

In comparing with the results of using Inception v4 and SCNN , because the ResNet50 is based on the skip connections technique, we can say that we achieved our goal and make accuracy higher than using the single frame for face liveness detection which is 96.93%, and for the last model which used 20 frames and CNN-LSTM our results are still good because in comparing with them as we mentioned using 20 frames needs more time for detection, so here we can say that we reduced the time of detection by using 3 frames instead of 20 frames.

5. Conclusion, Limitations and Future Work

5.1. Conclusion

In conclusion, the use of ResNet50 which is a deep learning algorithm as a classifier gave us good results in face liveness detection, and the system we created on the Replay attack database achieved our main goal of this research. The model was able to classify the real or spoof person with high accuracy and also from the result obtained we found that the 3 frames are the ideal number of frames that should be used for detection and we proved that by increasing the frames to 5 and 7 frames which caused decreased in the accuracy of face liveness detection in our system and also increased the time of detection which is not good. The ResNet50 classifier gave us 96.93 % by using 3 frames which is good accuracy in detection in less time.

5.2. Limitations and Future Work

Our system approach has some limitations one of them represented by that it was not applied to real-time detection. Although many facial liveness detection problems have been addressed, the static and dynamic approach still has many unresolved problems which need to be considered. As a suggestion for developing instead of using ResNet50 which is 98 MB in size, we can use a different latest deep learning model which has less size such as MobileNetV2 which is 14 MB in size, or NASNet mobile which is 23 MB for detection to investigate in the accuracy that can be achieved and make comparison with our result, because using the lightweight deep neural network can have low latency for mobile and embedded device which will be an interesting topic in biometric field.

REFERENCES

Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. IEEE Transactions on circuits and systems for video technology, 14(1), 4-20. https://doi.org/10.1109/TCSVT.2003.818349

Bhatia, R. (2013). Biometrics and face recognition techniques. International Journal of Advanced Research in Computer Science and Software Engineering, 3(5), 93-99.

Tripathi, K. P. (2011). A comparative study of biometric technologies with reference to human interface. International Journal of Computer Applications, 14(5), 10-15. https://doi.org/10.5120/1842-2493

Hassaballah, M., & Aly, S. (2015). Face recognition: challenges, achievements and future directions. IET Computer Vision, 9(4), 614-626. https://doi.org/10.1049/iet-cvi.2014.0084

Ghiass, R. S., Arandjelović, O., Bendada, A., & Maldague, X. (2014). Infrared face recognition: A comprehensive review of methodologies and databases. Pattern Recognition, 47(9), 2807-2824. https://doi.org/10.1016/j.patcog.2014.03.015

Galbally, J., Marcel, S., & Fierrez, J. (2014). Biometric antispoofing methods: A survey in face recognition. IEEE Access, 2, 1530-1552. https://doi.org/10.1109/ACCESS.2014.2381273

Pinto, A., Schwartz, W. R., Pedrini, H., & de Rezende Rocha, A. (2015). Using visual rhythms for detecting video-based facial spoof attacks. IEEE Transactions on Information Forensics and Security, 10(5), 1025-1038. https://doi.org/10.1109/TIFS.2015.2395139

Bagga, M., & Singh, B. (2016, March). Spoofing detection in face recognition: A review. In 2016 3rd International Conference on Computing for Sustainable Global Development (INDIA Com) (pp. 2037-2042). IEEE.

Chakka, M. M., Anjos, A., Marcel, S., Tronci, R., Muntoni, D., Fadda, G., & Pietikäinen, M. (2011, October). Competition on counter measures to 2-d facial spoofing attacks. In 2011 International Joint Conference on Biometrics (IJCB) (pp. 1-6). IEEE https://doi.org/10.1109/IJCB.2011.6117509

Alotaibi, A., & Mahmood, A. (2016, June). Enhancing computer vision to detect face spoofing attack utilizing a single frame from a replay video attack using deep learning. In 2016 International Conference on Optoelectronics and Image Processing (ICOIP) (pp. 1-5). IEEE. https://doi.org/10.1109/OPTIP.2016.7528488

Koshy, R., & Mahmood, A. (2020). Enhanced Deep Learning Architectures for Face Liveness Detection for Static and Video Sequences. Entropy, 22 (10), 1186. https://doi.org/10.3390/e22101186

Sabaghi, A., Oghbaie, M., Hashemifard, K., & Akbari, M. (2021). Deep Learning meets Liveness Detection: Recent Advancements and Challenges. arXiv preprint arXiv:2112.14796.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). https://doi.org/10.1109/CVPR.2016.90

Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. ACM computing surveys (CSUR), 35(4), 399-458. https://doi.org/10.1145/954339.954342

Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., ... & Worek, W. (2005, June). Overview of the face recognition grand challenge. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) (Vol. 1, pp. 947-954). IEEE.

Chingovska, I., Erdogmus, N., Anjos, A., & Marcel, S. (2016). Face recognition systems under spoofing attacks. In Face Recognition Across the Imaging Spectrum (pp. 165-194). Springer, Cham. https://doi.org/10.1007/978-3-319-28501-6_8

Bashar, A. (2019). Survey on evolving deep learning neural network architectures. Journal of Artificial Intelligence, 1(02), 73-82. https://doi.org/10.36548/jaicn.2019.2.003

Liu, H., & Lang, B. (2019). Machine learning and deep learning methods for intrusion detection systems: A survey. applied sciences, 9(20), 4396. https://doi.org/10.3390/app9204396

Thrall, J. H., Li, X., Li, Q., Cruz, C., Do, S., Dreyer, K., & Brink, J. (2018). Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. Journal of the American College of Radiology, 15(3), 504-508. https://doi.org/10.1016/j.jacr.2017.12.026

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. Journal of big data, 2(1), 1-21. https://doi.org/10.1186/s40537-014-0007-7

Lokhande, B. P., & Gharde, S. S. (2015). A Review on Large-scale Video Classification with Recurrent Neural Network (RNN). International Journal of Computer Science and Information Technologies, Jalgaon, India.

Chingovska, I., Anjos, A., & 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.

Carneiro, T., Da Nóbrega, R. V. M., Nepomuceno, T., Bian, G. B., De Albuquerque, V. H. C., & Reboucas Filho, P. P. (2018). Performance analysis of google colaboratory as a tool for accelerating deep learning applications. IEEE Access, 6, 61677-61685. https://doi.org/10.1109/ACCESS.2018.2874767

Hashemi, M. (2019). Enlarging smaller images before inputting into convolutional neural network: zero-padding vs. interpolation. Journal of Big Data, 6(1), 1-13. https://doi.org/10.1186/s40537-019-0263-7

Hashemi, M. (2020). Web page classification: a survey of perspectives, gaps, and future directions. Multimedia Tools and Applications, 79(17), 11921-11945. https://doi.org/10.1007/s11042-019-08373-8

Brownlee, J. (2018). What is the Difference Between a Batch and an Epoch in a Neural Network. Machine Learning Mastery, 20.

Carney, J. G., & Cunningham, P. (1998). The epoch interpretation of learning. IEEE Transaction on Neural Networks, 8, 111-116.