Journal of Theoretical and Applied Information Technology
31st March 2021. Vol.99. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A COMPARATIVE ANALYSIS PERFORMANCE OF DATA AUGMENTATION ON AGE-INVARIANT FACE RECOGNITION USING PRETRAINED RESIDUAL NEURAL NETWORK

KENNEDY OKOKPUJIE[1], ABERE REUBEN[2], JOYCE C. OFOCHE[1], BASUO J. BIOBELEMOYE[1], IMHADE PRINCESS OKOKPUJIE[3]

[1]Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria
[2]Department of Computer Sciences, Federal University of Petroleum Resources, Effurun
[3]Department of Mechanical Engineering, Covenant University, Ota, Nigeria

[1]kennedy.okokpujie@covenantuniversity.edu.ng, [1]mildredofoche@gmail.com,
[1]beyoofficialmusic@gmail.com, [2]abere.reuben@fupre.edu.ng, [3]isaac.odu-ayo@covenantuniversity.edu.ng,
[4]ihmade.okokpujie@covenantuniversity.edu.ng

ABSTRACT

There has been an immense improvement in face recognition research. Unfortunately, the accuracy of face recognition systems recognizing the same person over time due to ageing is open research. Minor geometric changes in the face that occur due to ageing contribute to face recognition systems' inaccuracy. Researchers, over subsequent years, have come up with methods to improve the performance of Age Invariant Face Recognition (AIFR) systems, the most recent one being the use of Convolutional Neural Network (CNN) to create face recognition models. The pre-trained residual network (ResNet) is trained and tested using a heterogeneous database to actualize this improvement. The heterogeneous database consists of images from 82 Caucasian subjects in the FG-Net database and 11 African subjects. These obtained images were augmented using geometric transformation and Noise to increase the amount of data for training. Afterwards, a model robust is developed. The Sliding Window framework was used to detect the faces fed into CNN for training and testing. After getting the results from our classification model, an analysis was carried out on the classification models of both the original dataset and augmented datasets. It was observed that the model performed remarkably with the noise-injected dataset and performed worst with the geometric transformation database.

Keywords— *Age Invariant Face Recognition (AIFR), Convolutional Neural Network (CNN); Geometric Transformation; Noise injection; Data Augmentation*

1. INTRODUCTION

Biometrics has become the most secure and accurate means of security in all most all spheres of life. It works by capturing unique physical features and patterns from a subject. These data will then be used as identifiers that the software will interpret and use for identification applications. Since it is very hard for people to steal, some parts of your body like your face, for instance, make biometrics people's top choice for top-level security [1]–[3].

A type of biometric authentication is Facial recognition. This works by capturing the subject's facial features. The facial recognition algorithm creates a biometric template by capturing key points (feature) points on the face known as nodal points. The features include the location of the eyes (distance between them), eyebrows, mouth, depth of the eye, ears, cheekbones, jawline, nose(width of the nose) and chin, as well as the measurement of the distance between them. The Two templates will be compared to check if the two images belong to the same person. The nodal points(facial features) are then measured to create numerical codes(a string of numbers) used to represent a person's face [4]–[8].

Having explained how to face recognition works, it is significant to point out that, the major reason behind how accurate a facial recognition system will depend on how you train the deep learning neural network. Research shows that the more data you feed into a deep learning neural network, the better. However, that amount of large data is not easily available; therefore; we need to apply data augmentation techniques to our already existing data. In this paper, we will be showing that data augmentation does not only provide a large amount of data to train with, but it also helps the face recognition model perform better in the classification of faces, even when they have changed due to ageing [9]–[11].

As contradictory as it may sound, noise addition can also be used as a data augmentation method to boost neural networks' accuracy. The addition of Noise such as Gaussian and Salt and Pepper to data will not only help us increase the amount of data that we can train, but it will enable the deep learning model to generalize better on noisy data [12].
Another technique of data augmentation includes Geometric transformation, which is a traditional form of data augmentation; data will be cropped, padded, trimmed, flipped horizontally and so on [13].

ResNet is a type of convolutional neural network. Deep neural networks are difficult to train because of gradually vanishing and degrading gradients to avoid ResNet being created. It uses the idea of shortcuts or skips connection, which allows you to take the activation from one layer and feed it to another layer even deeper in the network. We chose ResNet for this experiment because, it has been trained on more than one million images from the ImageNet database and has the ability to classify images into 1,000 object categories, such as pencil, keyboard, mouse and other types of animals. This makes it easier for us to pre-train and fine-tune the network to create a model that can recognize our images [14]–[16].

2 RELATED WORKS

Data Augmentation is a way of creating samples by transforming data, for increasing accuracy of classifiers. A technique allows people to enhance the quantity of already existing data for training. Data augmentation transforms data, which in this case are images, into diverse forms to increase data acquisition. In the process of traditional augmentation, data will be cropped, padded, trimmed and flipped horizontally. Although there have been new methods of augmenting data with the use of software or algorithms, an example is the GANs [17]–[20].

We augment data because it helps to trains large neural networks to increase accuracy in image classification. When dealing with neural networks, you need to provide a large, robust amount of dataset for training[21]. The latest improvements in deep learning models because of the increase and variety of data gathered over time. In a lot of classification problems, the data available is usually not enough to train robust and accurate classifiers, which leads to inaccurate classification, showing that without enough data for training, image classifiers are prone to lots of errors [22]–[25].

According to [26], when using deep neural networks such as CNN for facial recognition, the network goes through two operations: training and testing phases. The training of the CNN is a global optimization process, which learns by observing large dataset of a subject, this large dataset is produced from data augmentation while the testing stage is a process that uses the CNN to classify observed data.

[27] used five different methods of data augmentation to expand the amount of dataset for training. The methods used were landmark perturbation created by Shan hairstyle synthesis, glasses synthesis, pose synthesis, illumination synthesis. In the landmark perturbation, Jiang improved previous work done using Gaussian distribution to model the range of perturbation, instead of using the eight-neighbour deviation [28] to avoid random face landmark perturbation. These methods of data augmentation made the system robust to hairstyles, glasses, poses, and lighting.

[29] proposed that more does not make it effective in increasing image classification accuracy; they also proposed a Pixel Block Pairing (PBP) method to be more effective in image classification. At the end of the experiment, Wei discovered that more actually increased image classification accuracy in traditional augmentation and noted that the PBP method is just as effective as the traditional methods.

[33] researched on "A Discriminative Model for Age Invariant Face Recognition". The authors acknowledged that designing a suitable feature image and an operative matching structure for age invariant face recognition remains an open challenge. Li et al. proposed a discriminative prototype to handle face matching in the existence of age variation. Firstly, denote each face by designing a compactly sampled local feature description system, which the Scale Invariant Feature Transform (SIFT) and Multi-Scale Local Binary Patterns (MLBP) function as the local descriptors. By compactly sampling the two kinds of local descriptors from the whole face image, adequate biased information, including the dissemination of the edge direction in the face image was mined for further analysis. Both SIFT-based local features and MLBP-based local characteristics distance formed an enormous dimensional feature galaxy. To evade the overfitting difficulty, the authors employed an advanced algorithm named Multi-Feature Discriminant Analysis (MFDA) to practice these

two local feature spaces in a joined structure. Investigational outcomes indicate that the method outperforms a viable face recognition system on the FG-NET dataset. The proposed discriminative model used discriminative analysis local descriptors with densely sampled FGNET (82, 82). The recognition accuracy of 47.50% was obtained.

[34] developed "Age Invariant Face Recognition based on Texture Embedded Discriminative Graph Model". The authors acknowledged that an automatic face recognition system remains a challenge. The authors presented a new method to deal with age invariant face recognition by articulating a graph-matching problem. This technique engenders a graph from a set of fiducial landmarks of each face. The texture traces incarcerations that tend to be unchanging in a time-lapse as well as the conjoint face geometry structure. The nodes of the graph the texture of a face region around a landmark and the edges resemble the geometry pattern of the face. For each region, the age invariant texture data is mined by a discriminative and compressed feature prearranged in the Local Gabor Binary Pattern Histogram Sequence (LGBPHS) projected in an LDA subspace. An impartial utility is then calculated to match graphs for the purpose of registration and identification. Experimentations were done on the FG-NET Ageing dataset, and a recognition Accuracy of 64.47% was achieved.

[30] found ResNet to be very successful in face recognition. According to Ivan, with more augmentation in a dataset and with some minor modifications to ResNet, the deep neural network has the potential to produce stellar results.

At the top left of Figure 1, the residual block is the basic ResNet architecture building block for learning the residual function of $F\ (\beta)$ that is related to the standard function as shown in equation 1

$$H\ (\beta) = F\ (\beta) + \beta (1)$$

Any model that is closer to defining function β than random learns about the ideal H (β). Then the residual $F\ (\beta)$ is learned instead of making a network that learns $H\ (\beta)$ from randomly initialized weights. This concept saves both the training time and solves the disappearing gradient issue by letting gradients pass unchanged through the skip-connections. The stacked residual blocks are in the ResNet architecture, as seen in the figure above, along with two layers of 3×3 convolution layers. The convolutional layers carry out down sampling directly with a stride of 2. Additionally, the ResNet model has an additional global average-pooling layer after the last convolution layer at the end and a convolution layer at the start. It has a softmax function and a 1000-neuron fully connected layer. The layered residual blocks in the bottleneck architectures have three convolution layers, resulting in more models that are effective; this is in regards to the deeper ResNet 50/101/152.

2.2 Problem Statement

Despite the fact that facial recognition has become one of the best modes of biometric authentication over the past few years, there are still various drawbacks with this type of biometrics such as the likely hood of face spoofing, but the major challenge here is the inaccurate matching. Inaccurate matching can be caused by low lighting, pose, facial hair, but most importantly ageing.

We augmented our dataset using geometric transformation and noise injection to increase the amount of data for training and to also make the model robust. We used the Sliding window framework on the images to capture just the faces, which we fed into the convolutional neural network for training and testing. After getting the results from our classification model, we analyzed the classification model's performance for both the original and augmented datasets. The comparative analysis of the performance of both augmented and non-augmented is discussed in Section 3.

3 METHODOLOGY

All data was collected from FG-Net, consisting of 82 Caucasian subjects and 11 African subjects. This gave rise to our original dataset, which consists of 93 subjects and 1,145 images. This dataset is privately obtained from a research cluster group from Covenant University, Ota, Nigeria.

3.1 Data Augmentation Procedure

We used MATLAB to perform these perturbations. In MATLAB, an augmentedImageDatastore is used to randomly perturb the training images for each epoch, so that each epoch uses a slightly different data set.

The sliding window detection algorithm was used for face detection. We passed a grid cell of 2 x 2 across a whole image initially, then later increasing it. After the face have been detected, we cropped out the background and proceeded forms to add Noise and geometrically transform the images.

To create our Dataset 2, we added the different types

Journal of Theoretical and Applied Information Technology
31st March 2021. Vol.99. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of Noise listed below:

- Gaussian Noise
- Salt & Pepper Noise

This increases the original dataset from 1,145 images to 3,435 images.

To create our Dataset 3, we augmented the images using random geometric transformations with the help of the image augmenter to give the following properties:

- Reflection in the left-right direction
- Reflection in the top-bottom direction
- Rotation by an angle in the range 0 to 360°
- Scaling by a factor in the range of 0.5 to 2
- Horizontal translation in the range -10 to 10 pixels
- Vertical translation in the range -10 to 10 pixels

This increases the original dataset from 1,145 images to 3,435 images.

3.2 System Specification

The simulations were carried out using four GPU boards (GTX 1080, Titan X-Pascal, two Titan XP). It was performed in MATLAB R2018b software using the Deep Learning Toolbox. The GPUs were equipped with Intel Core i7-5820 K processors and 32 GB RAM each within two servers. The Ubuntu servers run LTS 14.04.3. Version 0.11.0 and the Python library TensorFlow was used to implement train and visualize the convolutional neural networks.

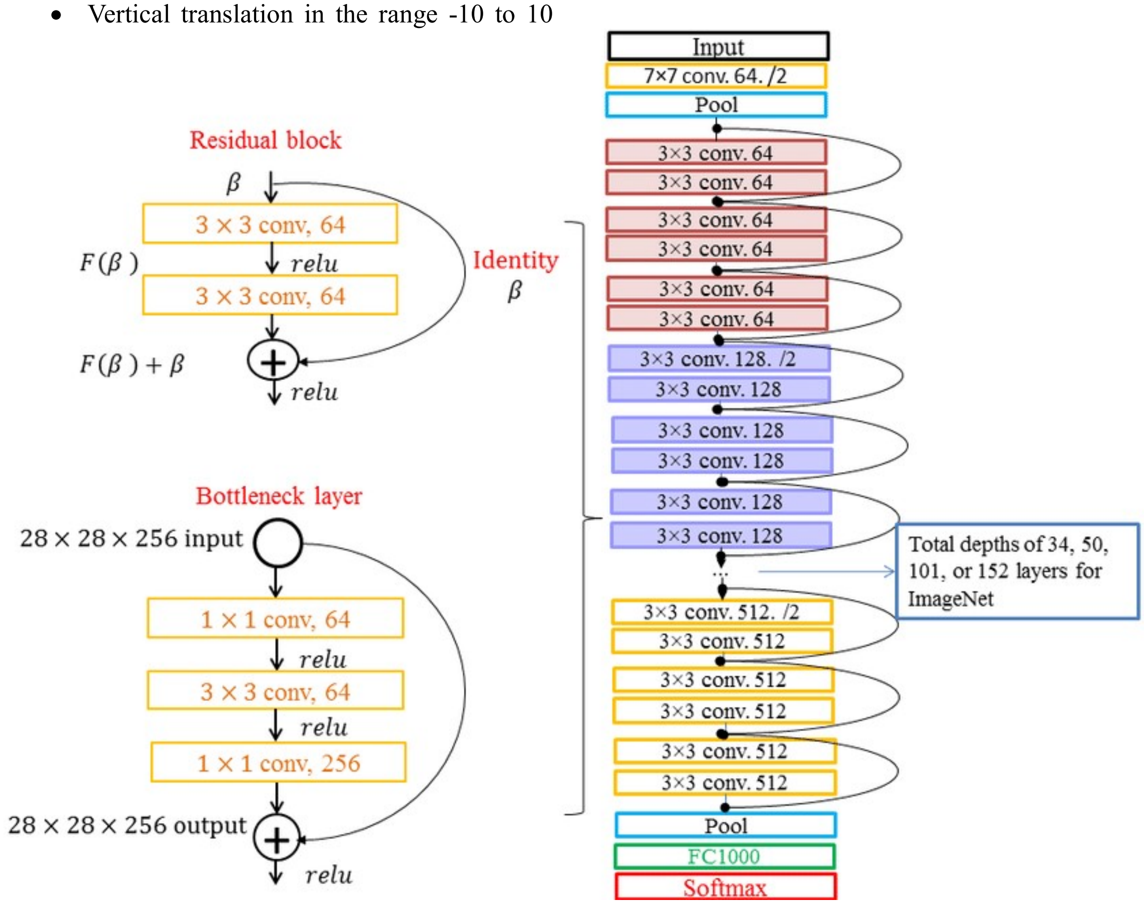

Figure 1: Resnet Architectures (Right) And Residual Block (Top Left), Bottleneck Layer (Bottom Left) [31]

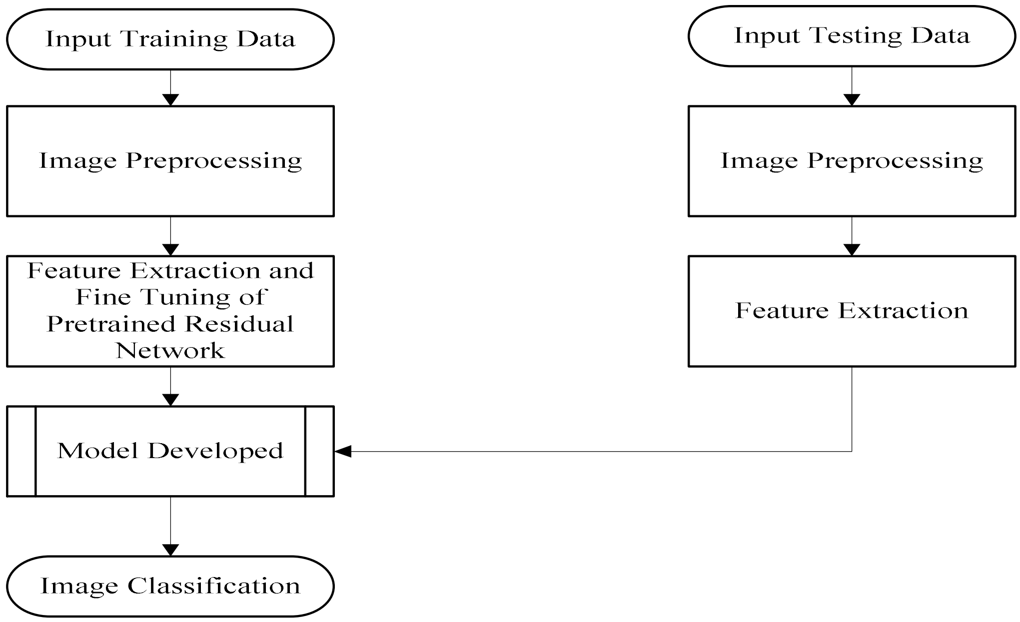

Figure 2. Flow Chart Of The System Architecture

3.3 System Architecture

Our model's system architecture has been designed to be able to accurately recognize subjects, irrespective of how their faces have changed due to ageing. The model goes through two stages; the training stage and the testing stage, we did this for both the non-augmented and augmented databases. Figure 2 shows the flow chart of the system architecture.

3.3.1 Training stage

Each database consists of 93 subjects, and these images went through some steps to train our model.
The training stage goes through the following procedures:

a. Input Training Data

In this process of our architecture, we put into the system, the images, which were in size 224 x 224 x 3, because that is the input image size ResNet can take. We input the dedicated 70% of each dataset during this step.

b. Image Preprocessing

Some of these pictures have been taking under low lighting or shot with low-quality cameras, and because of these occurrences, these pictures turn out dark, unclear or with a lot of Noise. This process aims to improve the quality of these images to increase the performance of the model. During this process, any grey-scale image is changed to RGB, because the pre-trained model reads images with three channels. The background noise is also cropped out leaving behind the just the face, with the aid of Viola-Jones, a framework known for its sliding-window face detection algorithm.

c. Feature Extraction and Fine Tuning of Pretrained Residual Network

During this process, the ResNet extracts the features of the processed images. This is done by reducing the data dimension into smaller sizes for easier processing and reducing the number of computing resources needed.

The fine-tuning process is referred to as transfer learning. Here we take a pre-trained model with its existing weights, which in our case is ResNet, and add our 93 fully connected softmax layer (output layer) along with its activated weights. Also in this process, we re-trained and updated the weights of all the parameters in ResNet with the extracted features from each input training data, we did not fix nor freeze the other existing weights, while training this model.

Adding our softmax classifier and re-training allows us to adjust the pre-trained model in such a way that, it can classify the images we feed into the system. We chose ResNet because, the network has been trained

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

on more than one million images from the ImageNet database and can classify images into 1,000 object classes, such as pencil, keyboard, mouse and other types of animals. For face recognition. ResNet already solves a problem similar to ours, and that is why we were able to work easily with the existing data learnt by the model. This not only saves time and resources in building a new model from scratch, but it also increases the performance of our model.

d. Model Developed

Now, we have created our desired model. The model has now learnt all the features of each data and can classify any image we feed into it. The model consists of the ResNet architecture and the acquired data from training. In the training phase, this is where it ends.

3.3.2 Testing stage

This testing stage will allow us to know for certain, whether our model can accurately classify images. The testing stage goes through the following procedures;

a. Input Testing Data

During this process of our architecture, we put into the system, the augmented images, which were in size 224 x 224 x 3, because that is the input image size ResNet can take. We input the dedicated 30% of each dataset during this process.

b. Image preprocessing

Just like in the training phase, the testing phase also carries out this process to improve the quality of these images, to increase the performance of the model.

c. Feature extraction

During this process, the ResNet extracts the features of the processed images. This is carried out by reducing the data's dimension into a smaller size for easier processing and reducing the number of computing resources needed.

d. Testing the developed model

After the features have been extracted, the model acts like the human brain by using what it has learnt from the training data, to classify the new input data brought forward. It will compare the extracted features learnt during training, with the new input data, to distinguish the subject. This process is where the classification takes place. The Softmax classifier that we spoke about previously will then predict the image in a probabilistic manner. For example, suppose we input image data into the system. In that case, the softmax will result, based on the information extracted from the input data, saying that this image is 0.9 (90%) subject 1 and 0.1 (10%) subject 2.

e. Image Classification

This is the output of the developed model. After the model has been able to classify the image, the result is the recognized subject. The result analysis criteria are justifying because evaluating and assessing machine learning models' performance is essential to use standardized measures [32]. These standard criteria are usually derived from the confusion matrix; in this case, it is a classification problem of 93 by 93 matric. The derived measures used in the analysis of the results are shown in equation 2 to equation 4.

3.4 Evaluation Metrics

We had to get our classification metrics derived from the confusion matrix to analyze our model, which are as shown in equations 2-4.

- Positive Predicted Value (Precision)

$$PPV = \frac{TP}{FP + TP} \quad (2)$$

- True Positive Rate (or Sensitivity)

$$TPR = \frac{TP}{TP + FN} \quad (3)$$

- F-Measure (F1 score)

$$F - Measure = \frac{2TP}{2TP + FP + FN} \quad (4)$$

The Sensitivity evaluates our model's ability to predict the true positives (TP) in each class. The Precision is the correctly classified positive sample (TP) ratio to the total positive predicted samples. The F-measure will represent high classification performance of our model. F-Measure also reflects the harmonic mean between Sensitivity and Precision.

4 RESULT AND ANALYSIS

With our classification assessment metrics table's aid, we created a graph that visually displays the results from our classification model. From the information represented on the graph, we will be able to analyze the classification model easily.

Based on Figure 3 which is the Comparative analysis of AIFR Sensitivity produced by the pre-trained RESNET model, we can see that the system's Original Sensitivity ranges between 0 – 1, although most of the classes have values ranging from 0.167- 0.667. While the Geometric Transformation Sensitivity rate fluctuates from 0 – 1, most classes have values ranging from 0.167- 0.4. However, the Noise Sensitivity rate (Orange) of most classes has values ranging from 0.760 – 1.

Based on Figure 4, which is the Comparative analysis of AIFR Precision produced by pre-trained RESNET model, we can see that the Original Precision(Blue) of the classes fluctuates between the values of 0 – 1, but the majority of the classes range from 0 - 0.533. On the other hand, the Geometric Transformation Precision (Grey) of the classes fluctuates from 0 – 0.917, but the majority of the classes have values range from 0.09 – 0.538. While the Noise Precision (Orange), fluctuates from 0 – 1, we can see that the majority of the values are within the range of 0.785 – 1.

Based on the results from Figure 5, which is the Comparative analysis of AIFR F-Measure produced by pre-trained RESNET model, we can see that the Original F-Measure(Blue) of the classes fluctuates between the values of 0 – 1. Majority of the classes possess values that range from 0 – 0.37. The Geometric Transformation F-Measure of the classes has values that vary within the range of 0 – 1, and we can also see that majority of classes have values from 0.118 – 0.428. The F-Measure of the classes when Noise was injected has values that vary within the range of 0 – 1, and we can also see that majority of classes range from 0.7 – 0.9.

In addition, it is observed that the graphs are shown there is a level of variance in the values, whereas some classes perform far better than some other classes, this occurs because of the number of images a class had for the experiment. For example, class 50 has an accuracy of 0.994 because it has eight images and class 86 has a value of 1.000 because it has twelve images, so the number of images a class used for experimentation will determine high or low a class' value might be on a graph. This affirms the already established theory that deep learning models accuracy largely depends on the dataset's quantity. That is, the larger the data use in training the model, the better the Accuracy and Precision of the model.
Using the data augmentation techniques with pre-trained ResNet model for the enhancement of Age-invariant face recognition system. Some observations were made. Firstly the AIFR dataset was able to be increased by the factors of the types of noise injection used in this case by two and the geometric transformation used in this case by six.
Secondly, significant improvements were observed with the image noise data augmentation using the Sensitivity, Precision and F-measure metrics. These three metrics are a positive set of measurement; meaning the higher these values to one the better the quantities. Thirdly, the image geometric transformation augmentation has very little or no effect on the AIFR dataset using the Sensitivity, Precision and F-measure metrics.

Journal of Theoretical and Applied Information Technology
31st March 2021. Vol.99. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

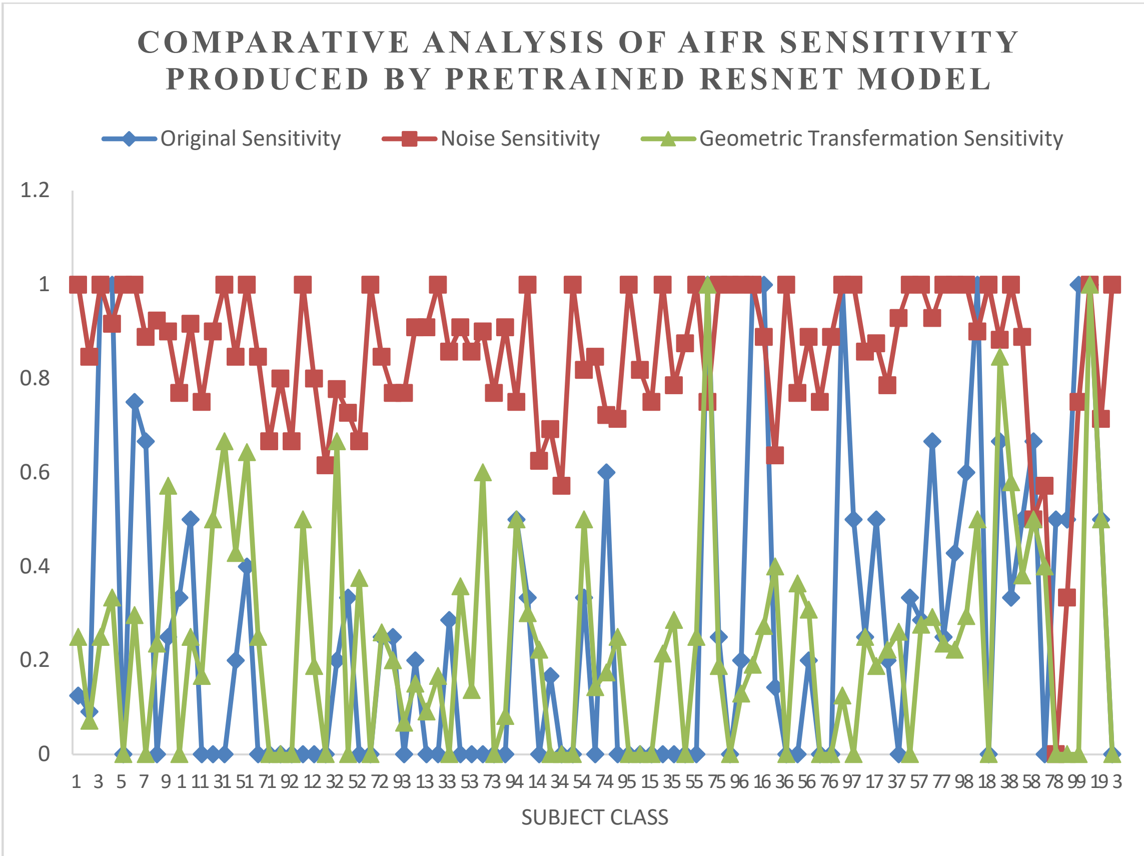

Figure 3. Comparative Analysis Of AIFR Sensitivity Produced By Pre-Trained RESNET Model

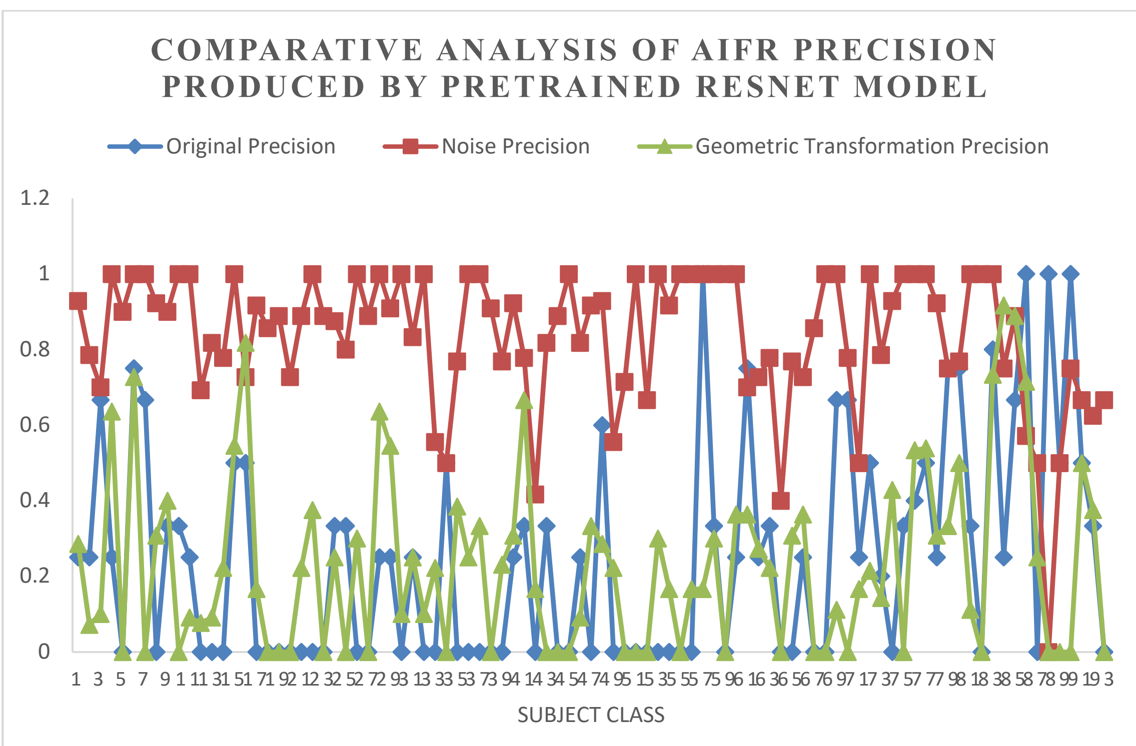

Figure 4. Comparative Analysis Of AIFR Precision Produced By Pretrained RESNET Model

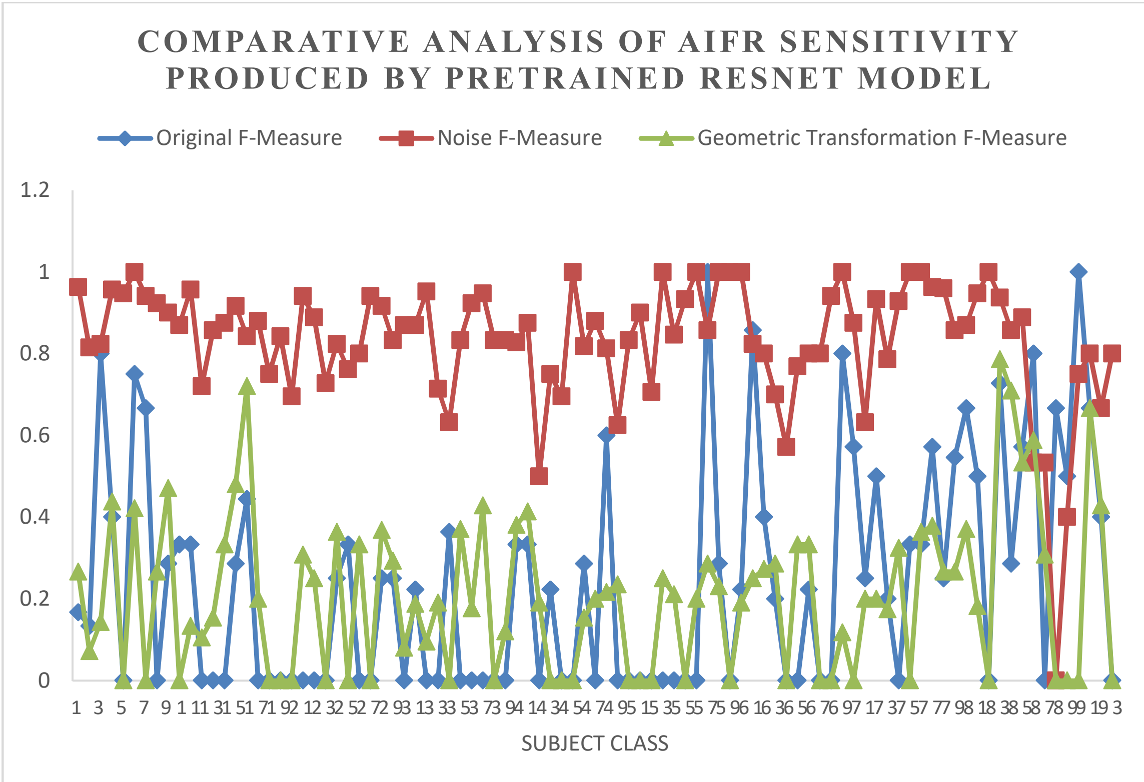

Figure 5. Comparative Analysis Of AIFR F-Measure Produced By Pre-Trained RESNET Model

5. CONCLUSION

The results show that image geometric transformation augmentation of AIFR improved the Precision, Sensitivity and F-measure of facial recognition insignificantly. Still, in other aspects, the system performed worse than when trained and tested without augmenting the AIFR images. And the reason for this might be due to the geometric transformations we performed on the images. It is not possible for a human being to have a face that has been flipped, rotated, nor translated like the images we distorted during this experiment. These distortions change the images' pixel sizes, making it difficult for the system to give an accurate classification.

On the other hand, the results show that Noise injection image augmentation significantly improved the Precision, Precision and F-measures of AIFR system compared with the original dataset and the geometric transformation dataset results. Unlike Geometric Transformation, the addition of Noise in images does not change the pictures' integrity (they remain the same). Using Noise injected images to train and test a face detection model is more realistic since perturbation does not affect an image's facial features.

ACKNOWLEDGEMENT

This paper is sponsor by Covenant University Centre for Innovation Research and Development (CUCIRD), Ota, Ogun State, Nigeria

REFERENCES

[1] K. Okokpujie, O. Modupe, E. Noma-osaghae, O. Abayomi-alli, and E. Oluwawemimo, 'A Bimodal Biometric Bank Vault', *Int. J. Mech. Eng. Technol.*, vol. 9, no. 9, pp. 596–607, 2018.

[2] A. Ross, S. Banerjee, and A. Chowdhury, 'Security in smart cities: A brief review of digital forensic schemes for biometric data', *Pattern Recognit. Lett.*, vol. 138, pp. 346–354, 2020.

[3] T. Habibu and A. E. Sam, 'Assessment of vulnerabilities of the biometric template protection mechanism', *Int. J. Adv. Technol. Eng. Explor.*, vol. 5, no. 45, pp. 243–254,

2018.
[4] H. Ohmaid, S. Eddarouich, A. Bourouhou, and M. Timouyas, 'Iris segmentation using a new unsupervised neural approach', *IAES Int. J. Artif. Intell.*, vol. 9, no. 1, pp. 58–64, 2020.
[5] E. Noma-osaghae, R. Okonigene, and K. O. Okokpujie, 'Design and Implementation of an Iris Biometric Door Access Control System', in *2017 International Conference on Computational Science and Computational Intelligence Design*, 2017, pp. 590–593.
[6] S. Dargan and M. Kumar, 'A Comprehensive Survey on the Biometric Recognition Systems based on Physiological and Behavioral Modalities', *Expert Syst. Appl.*, p. 113114, 2019.
[7] R. Jafri and H. R. Arabnia, 'A Survey of Face Recognition Techniques', *J. Inf. Process. Syst.*, vol. 5, no. 2, pp. 41–68, 2009.
[8] A. Tolba, A. El-baz, A. A. El-harby, and A. S. Tolba, 'Face Recognition : A Literature Review', *Int. J. Signal Process.*, vol. 2, no. January, p. 2, 2005.
[9] S. Javadi and S. A. Mirroshandel, 'A novel deep learning method for automatic assessment of human sperm images', *Comput. Biol. Med.*, vol. 109, no. April, pp. 182–194, 2019.
[10] Z. An, W. Deng, J. Hu, Y. Zhong, and Y. Zhao, 'APA : Adaptive Pose Alignment for Pose-Invariant Face Recognition', *IEEE Access*, vol. 7, pp. 14653–14670, 2019.
[11] H. Kaiming, X. Zhang, R. Shaoqing, and S. Jian, 'Deep Residual Learning for Image Recognition', in *Proceedings of the IEEE conference on computer vision and pattern recognition, 2016*, 2016, pp. 770–778.
[12] K. Okokpujie and S. Apeh, 'Predictive Modeling of Trait-Aging Invariant Face Recognition System Using Machine Learning', in *" in Information Science and Applications: Springer, 2020*, 2020, pp. 431–440.
[13] K. Okokpujie, S. John, and S. Alex, 'Implementation of a Bimodal Based System for Access Control System in Data Centres', 2019.
[14] Y. Zhou, X. Kang, and F. Ren, 'Employing Inception-Resnet-v2 and Bi-LSTM for Medical Domain Visual Question Answering', *CEUR-WS.org*, vol. 2125, no. 107, 2018.
[15] C. Zhenzhou and D. Pengcheng, 'Face recognition based on improved residual neural network', in *2019 Chinese Control And Decision Conference (CCDC)*, 2019, vol. 139, pp. 4626–4629.
[16] A. C. Carneiro *et al.*, 'Parameter optimization of a multiscale descriptor for shape analysis on healthcare image datasets', *Pattern Recognit. Lett.*, vol. 125, pp. 694–700, 2019.
[17] Y. Toda and F. Okura, 'How Convolutional Neural Networks Diagnose Plant Disease', *Plant Phenomics*, vol. 2019, pp. 1–14, 2019.
[18] P. Wang, F. Su, Z. Zhao, Y. Guo, Y. Zhao, and B. Zhuang, 'Deep class-skewed learning for face recognition', *Neurocomputing*, vol. 363, pp. 35–45, 2019.
[19] B. Yin, L. Tran, H. Li, X. Shen, and X. Liu, 'Towards interpretable face recognition', *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2019-Octob, no. Iccv, pp. 9347–9356, 2019.
[20] S. Zhao, J. Li, and J. Wang, 'Disentangled representation learning and residual GAN for age-invariant face verification', *Pattern Recognit.*, vol. 10, no. 7, pp. 1–18, 2019.
[21] K. Okokpujie, S. John, C. Ndujiuba, and E. Noma-Osaghae, 'Development of an Adaptive Trait-Aging Invariant Face Recognition System Using Convolutional Neural Networks.pdf', in *" in Information Science and Applications: Springer, 2020.*, 2020, pp. 411–420.
[22] Z. Niu, M. Zhou, L. Wang, X. Gao, and G. Hua, 'Ordinal regression with multiple output CNN for age estimation', in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016-December.
[23] M. A. Febriantono, S. H. Pramono, Rahmadwati, and G. Naghdy, 'Classification of multiclass imbalanced data using cost-sensitive decision tree c5.0', *IAES Int. J. Artif. Intell.*, vol. 9, no. 1, pp. 65–72, 2020.
[24] L. Bencke, C. Cechinel, and R. Munoz, 'Automated classification of social network messages into Smart Cities dimensions', *Futur. Gener. Comput. Syst.*, vol. 109, pp. 218–237, 2020.
[25] R. P. W. Duin, T. C. W. Landgrebe, D. M. J. Tax, and P. Paclı, 'The interaction between classification and reject performance for distance-based reject-option classifiers', *Pattern Recognit. Lett. 2*, vol. 27, pp. 908–917, 2006.
[26] I. A. Adeyanju, E. O. Omidiora, and O. F. Oyedokun, 'Performance evaluation of

different support vector machine kernels for face emotion recognition', in *IntelliSys 2015 - Proceedings of 2015 SAI Intelligent Systems Conference*, 2015, pp. 804–806.

[27] H. S. Jiang-Jing, J.-S. Huang, Z. Xiang-Dong, and Z. Xi, 'Data augmentation for face recognition', *Neurocomputing*, vol. 230, pp. 184–196, 2017.

[28] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, 'Local binary patterns and its application to facial image analysis: A survey', *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 41, no. 6, pp. 765–781, 2011.

[29] C. Zhou, L. Wang, Q. Zhang, and X. Wei, 'Face recognition based on PCA and logistic regression analysis', *Optik (Stuttg).*, vol. 125, no. 20, pp. 5916–5919, 2014.

[30] I. Huerta, C. Fernández, C. Segura, J. Hernando, and A. Prati, 'A deep analysis on age estimation', *Pattern Recognit. Lett. J.*, vol. 7, no. 28, pp. 1–11, 2015.

[31] J. Shao and Y. Qian, 'Three convolutional neural network models for facial expression recognition in the wild', *Neurocomputing*, vol. 355, pp. 82–92, 2019.

[32] K. Okokpujie, S. John, C. Ndujiuba, J. A. Badejo, and E. Noma-Osaghae, 'An improved age invariant face recognition using data augmentation', Bulletin of Electrical Engineering and Informatics, vol. 10, no. 1, pp. 179-191, 2021.

[33] Z. Li, U. Park, and A. K. Jain, 'A discriminative model for age invariant face recognition'. *IEEE Transactions on Information Forensics and Security*, *6*(3 PART 2), 1028–1037, 2011. https://doi.org/10.1109/TIFS.2011.2156787

[34] H. Yang, D. Huang, and Y. Wang, 'Age Invariant Face Recognition based on Texture Embedded Discriminative Graph Model'. In *IEEE International Joint Conference on Biometrics (IJCB), Clearwater, 29 September-2 October 2014, 1-8* (pp. 1–8).