iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XI **Month of publication:** November 2022

DOI: https://doi.org/10.22214/ijraset.2022.47562

A CNN and VGG- 16 Approach for Identification System Based on Ear Biometrics

Tehniyat Zahoor[1], Ravinder Pal Singh[2], Dr. Monika Mehra[3]
[1]*M. Tech Scholar, Department of ECE Engineering, RIMT University, Mandi Gobingarh, Punjab, India*
[2]*Technical Head and Professor, Dept of Research, Innovation & Incubation, RIMT University, Mandi Gobingarh, Punjab, India*
[3]*HOD and Professor, Department of ECE Engineering, RIMT University, Mandi Gobingarh, Punjab, India*

Abstract: With the increase in digitization over the entire world, there is a growing concern for security and privacy. Alphonse Bertillon, who created a technique to identify a person by precise body measures, is credited with developing biometrics for the first time in Paris during the 1800s. Since then, biometrics have greatly aided in preserving human security and privacy. The proposed study offers a through investigation of the ear as a biometric system utilising a novel taxonomy. It also provides a detailed analysis of databases, performance evaluation criteria, and cutting-edge methodologies. This research work acquired 99.8% accuracy and 0.2% loss while training and testing the model. Convolution Neural Network (CNN) is used due to its proven accuracy in object detection. To achieve better results in large scale image recognition. The research has been supported with the datasets from Kaggle. As a benchmark for comparison, the database from IIT Delhi. The results show an enhancement over the existing techniques and act as a springboard for new ear biometrics researchers.
Keywords: Skin cancer, CNN, deep learning, kaggle, python.

I. INTRODUCTION

The term "biometrics," derived from the Ancient Greek words "bios" and "metron," refers to the recognition of people based on their inherent physical or behavioural characteristics. The way that biometric systems identify the body in addition to knowledge or possession tokens, like passwords or papers, sets them apart from more conventional kinds of security. The face, fingerprint, or iris serve as common recognition characteristics. However, recent research is creating detection of DNA, hand geometry, vascular pattern, and even body odour[1]. Handwriting, speech, and keyboard patterns are all included in behavioural biometrics, along with developing topics like gait recognition. The use of biometrics for main security access and, in certain cases, as a replacement for more traditional forms of identification has undergone a revolution since the late 1990s. Biometric systems that combine physiological characteristics such the face, fingerprints, eye, palm, knuckles, and ears have been described by researchers. Several examples of these physiological biometric traits are shown in Table 1 along with their advantages and disadvantages. Each biometric characteristic has benefits and drawbacks, and it is believed that no single biometric property consistently functions .

Biometric trait	Features to distinguish	Advantages	Disadvantages
Face	Form of the ears, distance between the eyes, length and shape of the nose, and type of hair, skin texture	Doesn't require any contact, high accuracy with 3d cameras, it can be used to identify fake passwords	Skin texture can be affected by aging, recognition system does not produce accurate results while wearing mask, privacy concerns, obstructions like hats and hair can affect the system, camera angle needs to be properly aligned
Fingerprint	Finger's ridges	Easy to use, cost affective,	Not suitable for older people with a history of manual, it requires contact, can be affected by moisture, sweat, oil
Palmprint	Palm's handlines and the thickness or width of the palm	Low rejection rate, easy to use	Requires contact,
Iris	Distance of eyelids and eyelashes, as well as the iris edge, the centre and edge of the pupil, and eyelids.	Most accurate, contact less	High maintenance, costly, intrusive
Ear	Helix, concha, tragus, lobe	Doesn't require contact, not affected by skin aging, cost affective	obstructions like hair, headphones, accessories

Table 1Biometric Traits Comparison Table

A. Human Ear Anatomy

Figure 1.1 depicts the structure of the human ear and the 11 essential ear part. Helix is the ridge surrounding the top part of the ear. Parallel to helix is the antihelix. A concentric ridge known as the antihelix is located between the helix and the scapha, also called the helix's fossa. The region between the inner helix and the lower branch of the antihelix makes up the concha, which resembles a shell. A robust intertrack notch marks the bottom end of the concha. The intersection of the helix and antihelix is known as the crus of helix. On the right side of the intertrack notch, there is a little hump called the antitragus. The Tragus masks the earhole or canal. The Tragus masks the earhole or canal. The only cartilage-free area of the outer ear is the lobe, which is the fleshy lower portion of the auricle.

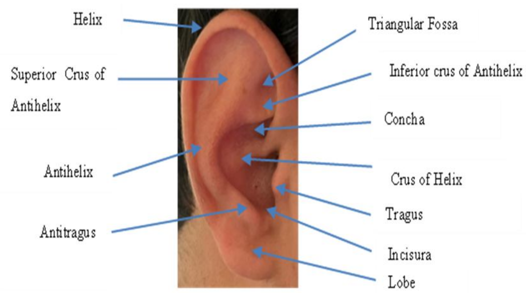

Figure 1: Human Ear Anatomy

B. Advantages of Ear biometrics

A French criminologist Alphonso Bertilon was the first person to recognize the uniqueness of the ear anatomy and propose its application as a biometric in 1890. Later in 1989, Iannarelli A [7] conducted a practical investigation on the matter by photographing 10,000 pairs of ears, proving that each pair is distinct. He said the twins' ears are likewise different from one another. This investigation supports the notion that the human ear has a distinctive shape. Police made a match based on ear patterns as evidence [8]. Additionally, it has been presented in Polish courts as scientific proof [9]. According to Amazon's patent for the ear, it would be possible to accept phone calls instantaneously without unlocking and to operate numerous capabilities remotely. After age 70, the form of the ear remains constant, unlike the face, which alters with aging. Additionally, unlike photographs of the face, images of the ears don't show mood or cosmetics. Significant user engagement was required for the non-intrusive iris and fingerprint scans. Before the covert ear photographs are taken, the target is not told or asked for permission. As a result, ear images are helpful in forensic analysis and monitoring. The built-in cameras of the mobile device are applied to extract images of the user's ears, but a separate sensor is also required to collect information of the user's fingerprints and irises. It is also helpful when a person can only be seen from their side. Multi-modal biometrics users must supply several characteristics to enhance their ability to detect liveness and defend themselves from spoofing attempts.

II. LITERATURE REVIEW

An in-depth analysis of existing ear detection and recognition techniques that rely on individual traits is provided here

Hurley et al. (2002) [11] in the paper titled “Force field energy functionals for image feature extraction” proposed a technique that transforms the force field's ear image. Image is recognised by viewing the image as an array of particles that are drawn near one another and act as the source of a Gaussian force field. The structure of the force field consists of potential energy wells, each of which is coupled to a potential channel. The model that serves as the foundation for ear description. It includes both the force field transform and potential well and channel extraction steps. The technique was verified with database of ears chosen from the XM2VTS face database. At a threshold value of 165, equal FRR and FAR error rates of 13.5% were forecasted. For manual registration (111*73 sized image) this model achieved a recognition rate of 91.5% and decidability index of 2.36. For automatic registration (150*120 sized image) this model achieved a recognition rate of 87.3% and decidability index of 2.71. With 20% occlusion from top, this model achieved a recognition rate of 80.4% and decidability index of 1.

Chen, H et al. (2005) [12] in the paper tiled “Contour matching for 3D ear recognition” proposed an iterative closest point (ICP) technique performed in two steps for matching 3D ears. The helix of the ear is found in 3D images in the first stage. To match a model ear helix with the test ear helix, the ICP algorithm is used to determine in the initial rigid transformation. Next stage is to align the model ears and test ear as closely as possible; the initial transformation is applied to a few points on the model ears. ICP algorithm iteratively improves the transformation.

The matching error criterion is based on the root mean square (RMS) registration error. The recognised ear is the model ear with the smallest RMS inaccuracy. Results were verified on a dataset of 30 subjects For true-match and false-match, the minimum, maximum, mean, and standard deviation are [0.56 1.96 0.79 0.26] and [0.78 3.81 1.57 0.53], respectively. The approach gave an EER (equal-error rate) is 6.7%.

Kumar et al. (2012) [13] in the paper titled “Automated human identification using ear imaging” The ear was segmented using a morphological and Fourier descriptor-based technique. Then, to extract local information, Gabor, log-Gabor, and sophisticated Gabor filters were applied. 125 and 221 people from two separate private databases were used to evaluate the approach. According to the experimental findings, shape features, force field transforms, and Eigen's ear are all outperformed by log-Gabor-based features. On 125 subjects, this approach achieved an accuracy of 96.27, and on 221 subjects, it had an accuracy of 95.92. The photos have a restricted orientation and scale, and their work has not been tested against any benchmark ear databases

Resmi et al. (2019) [14] in an article called “A novel approach to automatic ear detection using banana wavelets and circular Hough transform” developed an ear recognition method based on banana wavelets andthe Hough transformation. The accuracy is increased by using the Hough transformation to identify circular regions and the Banana wavelet to identify curvilinear ear characteristics. The image is pre-processed using the top hat technique and adaptive histogram equalisation. The technique has outperformed the template- and morphological-based operations in tests using typical datasets. As a classifier to identify whether or not the observed region is an ear, SNM and KNN were utilised while LBP and Gabor features were used for ear localization verification. The method was verified on UND database, CMU – PIE database and RR database and acquires detection rates of 85.56%, 92.78% and 91.24% respectively. This method achieved an accuracy of 89.86%. Their method for automatically determining the correct ear has a major drawback in that it has demonstrated subpar performance

Hassaballah, et al. (2020) [15] in the paper titled “Robust local oriented patterns for ear recognition” proposed a local-oriented patterns method for identifying ears. Edge directional information is used in the method to learn local structure information. The descriptor's robust characteristics are rotation- and illumination-invariant. The suggested descriptor encrypts every image pixel, producing a multi-region RLOP-encoded image. To estimate the distribution of characteristics, histograms of the regions are created. The final descriptor is created by concatenating each histogram to create it. The AMI, IITD-II, and AWE databases were used to evaluate the approach. When compared to other descriptor-based techniques, the strategy has performed better. This method acquired an average accuracy of 98%. However, the unconstrained database shows poor performance.

III. PROBLEM FORMULATION AND METHODOLOGY

A. Research Gap

Manual methods were employed to pre-process images of the ears in some algorithms which sometimes failed to consider the challenges posed by hair and earrings. Lighting and posture were not taken into consideration in some of the research methodologies. Ear localization property was not taken into consideration in some of the methods. Skin and fat content changes, light reflection differs from ear to ear, indirectly influencing identification accuracy. Some methodologies used ultra-images and those with barrel distorting, by utilizing this method, information is lost. In the proposed methodology, results are produced while taking care of all the above factors.

Some methodologies involved continuous detection system that regularly checks the examinee while conducting the examination was still required. This is to ensure that the examinee is never replaced by another user throughout the course of the test. In the proposed methodology, there is no continuous detection involved. Once the model is trained it will always recognise the ear of the individual Object detection and image recognition using AI algorithms is becoming a common practise and is essential for accurate ear image detection and recognition Even though ear biometrics had greater three-dimensional detection rates, special equipment was needed to take three-dimensional photographs. As a result, Python 3.110 and Anaconda Navigator 2.2.0 are used in the suggested methodology for detection and recognition of ear images. The number of training images, number of epochs, width and height of training images, number of hidden layers of CNN employed, matrix, etc. are some of the variables that affect accuracy. The accuracy is improved, nevertheless, as the number of training images increase.

Extensive research has been done on other biometric technologies like face and fingerprint. While fingerprint and face recognition have proven to be accurate methods for individual detection, there is still some gap for error, on the other hand ear biometrics when combined with other biometric technologies like fingerprint, iris and face can be used as an additional layer of security will help to decrease the error away.

B. Objectives

1) To study and analyse existing cutting-edge object detection and image recognition technologies and provide a secure ear biometric system.
2) To apply the proposed ear biometric system on commonly available dataset and increase the accuracy and decrease the number of epochs required to run the system.

C. Methodology

Preprocessing phase

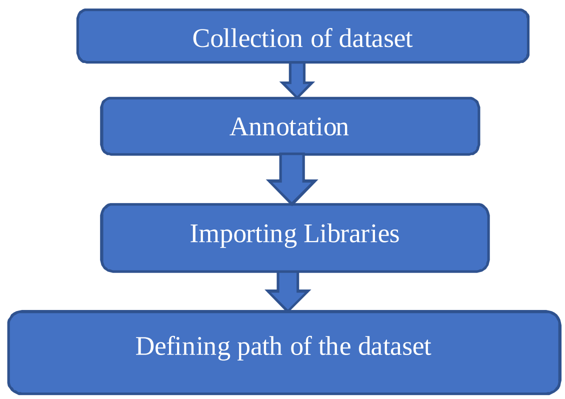

Figure 2 Flow Diagram of preprocessing phase of the model.

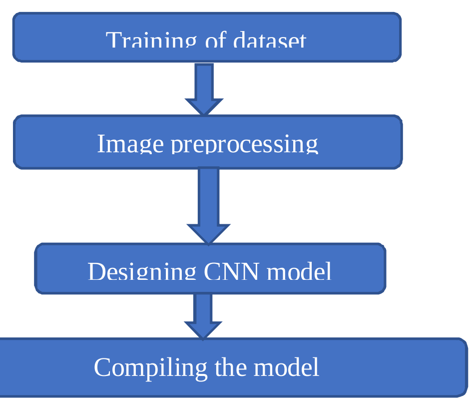

Figure 3: Flow Diagram of training phase of the model

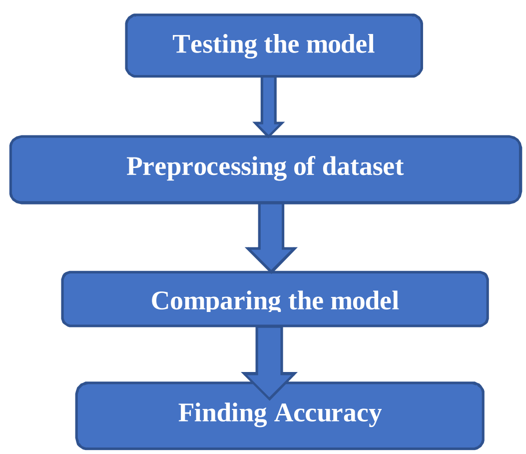

Figure 4: Flow Diagram of training phase of the model

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

1) *Collection of Dataset:* Dataset is downloaded from Kaggle The database was collected in an open environment. Position, scale, lighting, gender, ear ornaments, and cluttered backdrops are among the challenging elements that the photos must contend with. It has 10,200 images with 520 distinct themes. Participants ranging in age from 15 to 62 contributed the data, which has been collated. IITD designed the layout of another database. It has images covering 121 distinct topics. Three images, all in grayscale, are included for each subject. The photographs' angles occasionally vary.

Figure 5: Sample of images from IITD database

Figure 6*:* Sample images of ear databases in the unconstrained environment downloaded from Kaggle

2) *Annotation:* Four distinct points are used to represent the bounding rectangle which are x origin, y origin, w and h. In this instance, the starting points are w, which stands for width, and h, which stands for height. This bounding rectangle is known as the ground truth (GT) box. We may assess the performance of the ear recognition module by measuring how much of the expected bounding box of the ear inside face picture overlaps with. As an illustration of annotation, the location of the ear is indicated by the bounding box in red in Figure 3.3

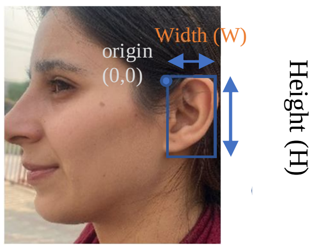

Figure 7: Annotation sample

3) *Importing the Libraries:* A collection of modules that are frequently used in different programmes without having to be written from scratch is referred to as a library. The required libraries are downloaded, then imported into the code to establish the environment. Same library can be imported several times to run different codes once it has been downloaded. The core principle of developing and designing models is made easier by libraries. In the proposed system, libraries that are used are tensor flow, numpy, cv2 and random.
4) *Defining the path of Dataset:* The dataset is imported into the model after collection. The path of the dataset is defined in the code to import it into the model. Once the path is specified in the code, further processing of the dataset can be carried out.
5) *Training the Model:* First step of training is image pre-processing. It is done for error-free execution of the code. Images are then converted into array form. Then the model of the system is designed and visualised for training. Different architectures can be used to model the system. In the given proposal, the module employs convolution layers, and kernel size. Dataset is repeatedly given to the model. Giving dataset repeatedly to the model is called running the epochs. More the number of epochs, more is the accuracy until it reaches a saturation point where accuracy decreases by running more epochs.

6) *Compiling the Model:* Model is saved and compiled after the training is completed. Epochs, the pattern of losses and accuracy in training and validation is saved. The training process is ended, and the model is saved if the training and validation parameters, such as training loss, training accuracy, validation loss, and validation accuracy, do not improve as the number of epochs increases.
7) *Testing the Model:* First images are processed, and the parameters are made same as the training parameters. Random images are taken from the dataset and the model is tested on the images. Test is performed to find if the model can identify the correct person from the image given.
8) *Comparing Accuracy:* Accuracy and time executed to run an epoch is calculated and graph is plotted between the different values of epoch and the respective accuracy and time executed.

IV. IMPLEMENTATION AND RESULTS

A. Implementation

First library installed is Keras, which makes it pretty simple to build the VGG-16 model. Tensor flow is used to design and train the model. Numpy is used for performing arithmetic operations and transform data into arrays. Glob is used to give the path of the files and folders. Matplot lib is used to plot graphs and charts. Matplotlib is used to plot confusion matrix. Sklearn metrics is used to import confusion matrix. Itertools is used if there are any iterations in the model. OS (Operating system) is used to define the path of the dataset and import from the specified path. Shutil tools is used to pre-process the dataset and perform manipulations on the input. Random library is imported to randomly select the images. After importing the libraries, the function of each block is explained below.

1) Step 1: Initiate Generator

The first block "Initiate Generator" is used to provide the path for dataset. It also performs splitting the data into Training and Validation sets, reshaping the images in training and validation dataset into 224x224, setting the batch size as 32. It can be changed while optimising the model, checking the number of classes in dataset. In this research, there are 4 classes of individuals whose ear images are to be differentiated, displaying an image from each class of the dataset. The function returns the number of classes, class names, training dataset and validation dataset

2) Step 2: Initiate Normalize Function

Data is normalised to increase the speed of the system.

3) Step 3: Initiate Model Function

This function takes no. of classes as input and applies VGG-16 algorithm on the input classes and combines imagenet model with the classes.

4) Step 4: Model Fit function

This block saves the model's name and monitors validation accuracy with each epoch. Feature of early stopping is added

5) Step 5: Plot Function

Plots the Training and Validation accuracy with respect to epochs. Plots the training and validation loss with respect to epochs.

6) Step 6: Save and Evaluate Model Function

The function shows the percentage with which the model is accurately recognising the image and then saves the model.

7) Step 7: Plot Confusion Matrix Function

This function is used to plot the confusion matrix.

8) Step 8: Call Plot Function

All the above functions run when the functions are called. Call plot function is used the call the functions in a sequential manner. Output of first function is the input of second function and accordingly the model runs.

B. Results

Results of Unconstrained dataset at epoch 5 shows training accuracy of 100%, validation accuracy of 98.64% in 184s.

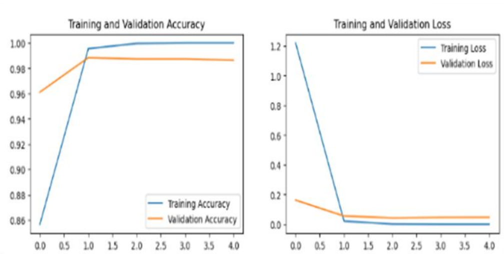

Figure 8: Accuracy and Loss Graph of unconstrained dataset at epoch 5

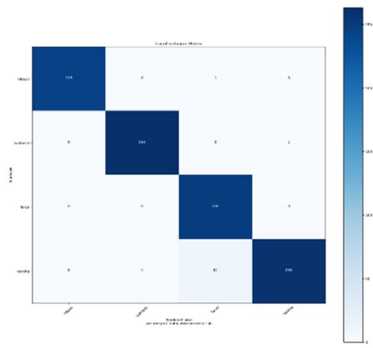

Figure 9: Confusion matrix of unconstrained data at epoch 5

Results of Unconstrained dataset at epoch 10 shows training accuracy of 100%, validation accuracy of 98.64% in 184s.

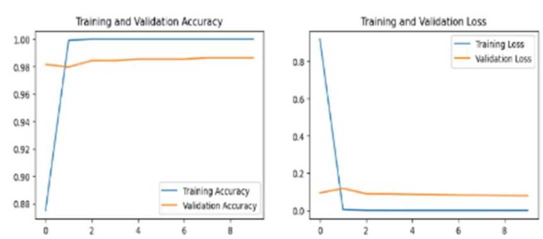

Figure 10 Accuracy and Loss Graph of unconstrained dataset at epoch 10

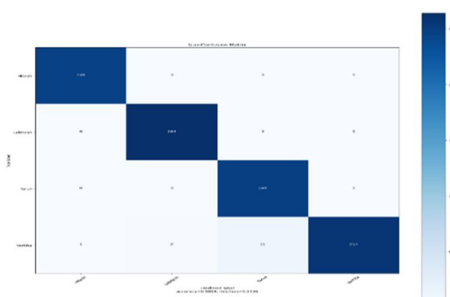

Figure 11: Confusion matrix of unconstrained data at epoch 10

Results of Unconstrained dataset at epoch 20 shows training accuracy of 100%, validation accuracy of 98.93% in 206s.

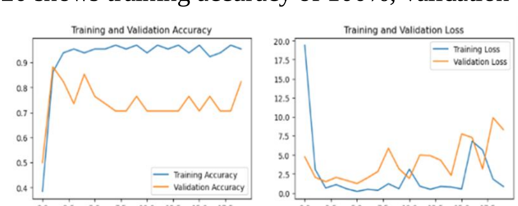

Figure 12: Accuracy and Loss Graph of unconstrained dataset at epoch 20

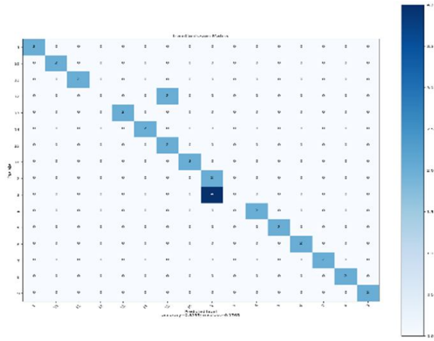

Figure 13: Confusion Matrix of unconstrained data at epoch 20

Results of Unconstrained dataset at epoch 50 shows training accuracy of 100%, validation accuracy of 99.42% in 213s.

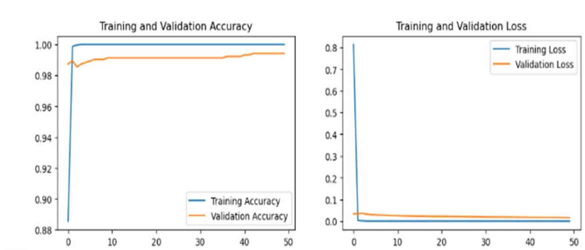

Figure 14: Accuracy and Loss Graph of unconstrained dataset at epoch 50

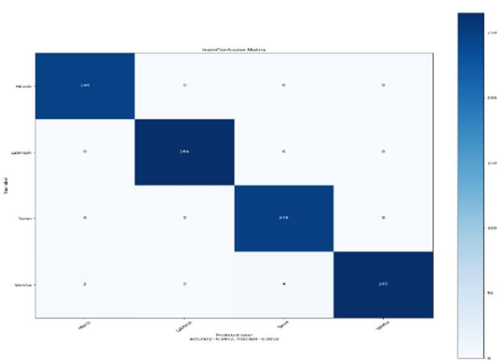

Figure 15: Confusion matrix of unconstrained dataset at epoch 50

Results of dataset at IITD dataset at epoch 5 shows training accuracy of 90.38%, validation accuracy of 62.96% in 206s.

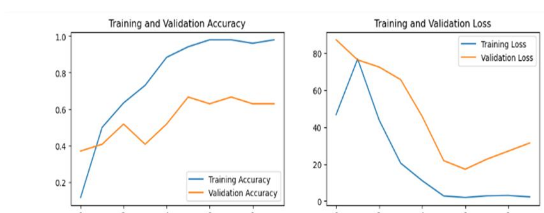

Figure 16: Accuracy and Loss Graph off IITD dataset at epoch 5

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

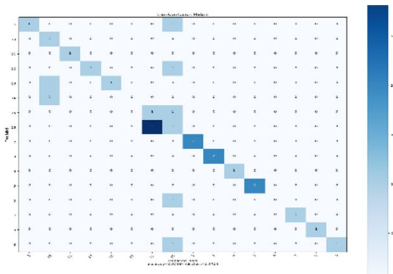

Figure 17: Confusion matrix of IITD dataset at epoch 5

Results of dataset at IITD dataset at epoch 5 shows training accuracy of 90.38%, validation accuracy of 62.96% in 206s.

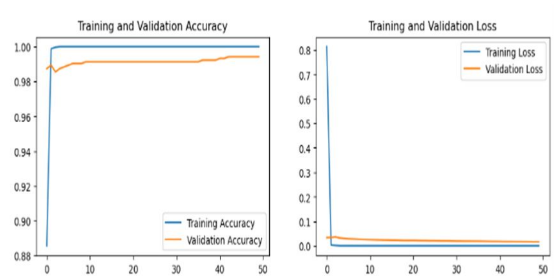

Figure 18: Accuracy and Loss Graph off IITD dataset at epoch 10

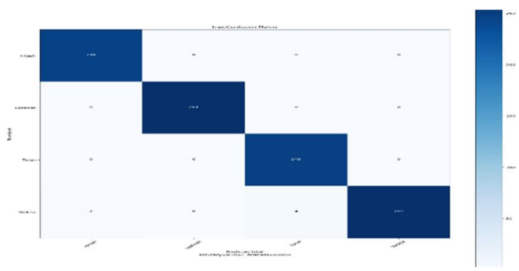

Figure 19: Confusion matrix of IITD dataset at epoch 10

Results of dataset at IITD dataset at epoch 5 shows training accuracy of 96.92%, validation accuracy of 82.35% in 48s

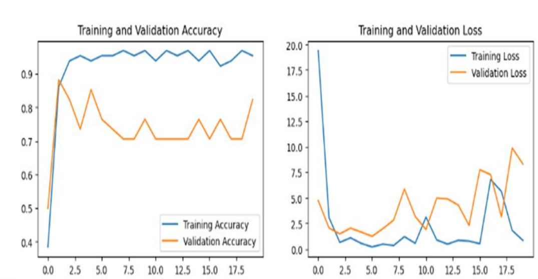

Figure 20: Accuracy and Loss Graph of IITD dataset at epoch 20

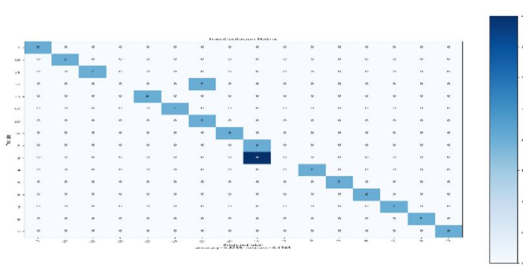

Figure 21: Confusion Matrix of IITD dataset at epoch 20

Results of dataset at IITD dataset at epoch 50 shows training accuracy of 93.25%, validation accuracy of 88.24% in 28s

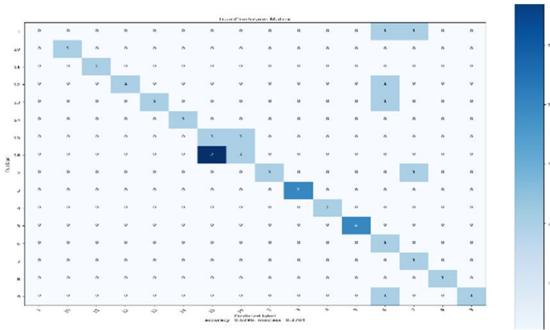

Figure 22: Accuracy and Loss Graph off IITD dataset at epoch 50

Figure 23: Confusion Matrix of IITD dataset at epoch 50

Comparison Table of Datasets

Number of epochs	Unconstrained dataset			IITD Dataset		
	Time taken	Validation Accuracy	Validation Loss	Time taken	Validation Accuracy	Validation Loss
Epoch 5	270 s	98.64%	0.0472	12s	62.96%	3.58
Epoch 10	253 s	98.64%	0.0796	12s	62.96%	3.58
Epoch 20	204 s	98.93%	0.0409	16s	82.35%	0.304
Epoch 50	215 s	99.42%	0.0160	9s	88.24%	0.2744

V. CONCLUSION AND FUTURE SCOPE

In this proposal, a comprehensive overview of the body of ear biometrics including benchmark datasets, performance assessment criteria, and current methodology has been provided. The database was downloaded from Kaggle. It contains 3500 images from 4 subjects taken in an unrestricted setting and shows technological shortcomings. The database is substantial and appropriate for deep learning. The database can be used to evaluate ear detection and recognition techniques. The results are also verified on a sample of 70 images of 16 subjects from IITD database. CNN and VGG-16 are utilised to identify and recognise ears. Results on IITD database shows an increase in accuracy and decrease in loss till epoch 5. In the proposed model an accuracy of 99.42% is achieved at the epoch 50 after increasing the epoch there is no increase in accuracy. On the sample of dataset from IITD database model achieved an accuracy of 88.24% at epoch 50. Analysis has demonstrated that these models perform admirably with large scale datasets.

Due to the challenging environmental constraints, there is a discernible performance difference compared to the unconstrained datasets.

VGG-16 can be combined with faster algorithms like YOLO algorithm to produce faster results.

Model shows less accuracy on the sample of dataset from iitd database, which shows that the model is less efficient on small scale dataset. Combination of other models can be used to increase the accuracy on smaller datasets.

New biometrics can be created which combine the ear biometric model with other biometrics like face, iris and fingerprint to make a more efficient biometric system.

REFERENCES

[1] G C Chandran. J, Dr. Rajesh. R. S., “Performance Analysis of Multimodal Biometric System Authentication”. IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.3, March 2009, pp 290-294

[2] Phillips et al., “An Introduction to Evaluating Biometric Systems, Guide to Biometrics”, IEEE Computer, February 2000, pp 56-63.

[3] A. K. Jain, A. Ross, and S. Pankanti, "A Prototype Hand Geometry-Based Verification System", 2nd Int'l Conference on Audio- and Video-based Biometric Person Authentication, Washington D.C., March 22-24, 1999, pp. 166-171

[4] Sa Prabhakar, "Fingerprint classification and matching with filterbank", Ph.D Thesis, University of Michigan State, November 2001 pp 33-34.

[5] M.A. Mohammed: "Automated algorithm for iris detection and code generation". Computer Engineering & Systems, International Conference on Computational & Experimental Engineering and Sciences, December 2009 pages 475–781, 201

[6] S Trabelsi, D Samai, A Meraoumia, K Bensid, Azeddine, Benlamoudi, Fadi Dornaika, Abdelmalik taleb-ahmed: "Finger-Knuckle-Print Recognition Using Deep Convolutional Neural Network" 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP)", May 2020 page 167

[7] Iannarelli A.: "Ear Identification, Forensic Identification Series" Paramount Publishing Company, Fremont, California, ISBN 0962317802, 9780962317804 pp 86-89

[8] Meijerman, Lynn; Thean, Andrew; Maat, George (2005). "Earprints in Forensic Investigations". Forensic Science, Medicine, and Pathology. 1 (4): 247–256.

[9] J. Kasprzak, "Forensic Otoscopy- new method of human identification" University of Warmia and Mazury Press, Jurisprudencija t. 66(58); 106–109 2005. (in Polish)

[10] S. Khobragade, D. D. Mor, and A. Chhabra, "A method of ear feature extraction for ear biometrics using MATLAB," Annual IEEE India Conference (INDICON), 2015.

[11] Hurley, D.J., Nixon, M.S., Carter, J.N.: "Force field energy functionals for image feature extraction. Image and Vision Computing 20 (April 2002) 314-316

[12] Chen, H. and Bhanu, B. : Contour matching for 3D ear recognition. In Proceedings of the IEEE Workshops on Application of Computer Vision (WACV). 123-128. January 2005

[13] Ajay Kumar, Chenye Wu : Automated human identification using ear imaging. Pattern Recognition 45(3), 956–968 (March 2012)

[14] Resmi, Raju G. : "A novel approach to automatic ear detection using banana wavelets and circular Hough transform" March 2019 International Conference on Data Science and Communication (IconDSC)

[15] Hassaballah, M., Alshazly, H., Ali, A.A.: Robust local oriented patterns for ear recognition. Multimedia Tools and Applications 79(41), 31183–204 August 2020

[16] Chidananda, P., Srinivas, P., Manikantan, K., Ramachandran, S.: Entropy-cum-hough-transform-based ear detection using ellipsoid particle swarm optimization. Mach. Vis. Appl. 26(2), 185–203 (February 2015)