Unveiling Identity with Siamese Network: Earbiometrics for Person Identification

G.Harshavarthini, N.Abishek, E.Karuppathal, and A.N. Nithyaa

G. Harshavarthini*
Department of Biomedical Engineering
Rajalakshmi Engineering College, Chennai, India.
varthiniganesh23@gmail.com
*Corresponding author

N. Abishek
Department of Artificial Intelligence and Machine Learning,
Rajalakshmi Engineering College, Chennai, India.
abisheksudhar@gmail.com

E. Karuppathal, Assistant Professor
Department of Biomedical Engineering
Rajalakshmi Engineering College
Chennai, India.
karuppathal.e@rajalakshmi.edu.in

A.N. Nithyaa, Associate Professor,
Department of Biomedical Engineering
Rajalakshmi Engineering College
Chennai, India.
nithyaa.an@rajalakshmi.edu.in

***Abstract*—Identifying individuals through ear images is a topic of ongoing research in the biometrics field. The ear, similar to other biometric traits such as the face, iris, and fingerprints, possesses many distinct and specific features that make it feasible for identification purposes. With the ongoing COVID-19 pandemic and widespread use of masks, face recognition systems have become unreliable. On the other hand, the ear is a suitable alternative for passive identification as it can be easily acquired without the subject's cooperation and its structure remains relatively unchanged over time. Furthermore, the ear is still visible even when a person is wearing a mask. This ear biometric system can complement other biometric methods in human recognition systems, providing identification information when the other methods are unreliable or unavailable. This study proposes a Siamese neural network for one-shot image recognition and tests its efficiency using the IITD-II dataset.**

Keywords—Ear-biometrics, Siamese neural network, one-shot image recognition.

I.INTRODUCTION

Covid 19 has had an inevitable impact on people's daily life in 2020. Changes in behaviour such as wearing masks have a considerable impact on biometric systems, especially face recognition systems. The project's motivation is the impact of COVID-19 on various biometric systems. In order to overcome the effects of the pandemic on Biometrics, the EAR is a suitable alternative for passive identification as it can be easily acquired without the subject's cooperation. Ear biometrics, on the other hand, provides a contactless and highly accurate method of person identification. By using ear biometrics, individuals can be identified without the need for physical contact or proximity, thereby reducing the risk of viral transmission. The potential applications of contactless person identification using ear biometrics go beyond the current pandemic. The COVID-19 pandemic has highlighted the need for contactless person identification methods. Ear biometrics provides a safe and accurate alternative to physical contact-based systems, promoting social distancing and reducing the risk of viral transmission.

The literature survey conducted by Zhang et al highlights the significance of ear recognition and explores the application of wavelet transform for feature extraction in this domain.[1]. Joshi and Joshi provide a comprehensive survey on ear biometrics, covering detection, feature extraction, and recognition methods[2]. Islam and Bappy present an in-depth survey on ear recognition, discussing various techniques and algorithms employed in this field[3]. Chen and Wu conduct a thorough review of ear recognition techniques, focusing on pattern recognition methods and machine learning algorithms[4]. Chakraborty et al. offer a comprehensive survey on ear biometrics, addressing topics such as the acquisition of images, preprocessing, feature extraction, and classification algorithms[5]. Sharma and Gupta present a survey on ear recognition, exploring the application of convolutional neural networks for feature extraction[6]. In a paper titled “Siamese Neural Networks for One-shot Image Recognition” by Koch et al**.** proposes a Siamese neural network architecture for one-shot image recognition[7]. The network is fed pairs of images and a binary label indicating whether the images are from the same class or different classes. The paper titled “Siamese network features for image matching” by I. Melekhov, J. Kannala, and E. Rahtu**,** demonstrates the effectiveness of using Siamese networks for feature extraction and matching in image retrieval tasks[8]. Su et al. focus on deep learning approaches for ear biometrics, offering valuable guidance for the utilization of the Siamese Neural Network[9]. The paper titled“Ear Biometrics for Human Identification” by Nikose, Shruti, and Hemant Kumar Meena[10], compares different algorithms for ear recognition, including principal component analysis, linear discriminant analysis, and support vector machines. The authors evaluate the performance of these algorithms on a dataset of ear images and provide insights into the strengths and weaknesses of each algorithm**.**

II.EAR ANATOMY

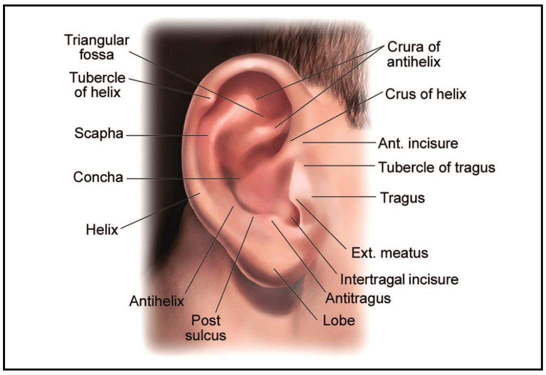

Fig.1.Anatomy of Human Ear

Human ears are unique for everyone, just like fingerprints or iris patterns. Ear biometrics relies on the fact that the shape, size, and contours of the ear are unique for each individual, even between identical twins. For person identification using ear biometrics, we need to consider the anatomy of the external ear.

1. Pinna: The pinna is the external part of the ear that is made of cartilage covered by skin. It is shaped like a funnel and has several ridges and grooves that help to collect and funnel sound waves into the ear canal. The pinna also helps to locate the direction of sound sources.

2. Ear canal: The ear canal, sometimes referred to as the tympanic membrane, is a tube-like structure that runs from the pinna to the eardrum. Earwax, which serves to shield and clean the ear, is produced by glands and fine hairs that line the ear canal. Individual differences in ear canal length and shape can have an impact on how well sound waves resonate and are amplified.
3. Helix: Starting from the top, the image shows the helix, which is the prominent curved rim that forms the upper and back part of the pinna. The antihelix is the inner ridge that runs parallel to the helix.

4. Concha: The concha is the hollow, bowl-shaped depression located just inside the opening of the ear canal. It serves to collect and direct sound waves into the ear canal.

5. Tragus: The tragus is the small, triangular projection of cartilage that partially covers the entrance of the ear canal. It helps protect the ear canal and assists in directing sound waves

6. Lobule: The lobule is the soft, fleshy, lower part of the auricle that does not contain cartilage. It is made up of connective tissue and fat.

III.SIAMESE NEURAL NETWORK

Neural networks are a type of machine learning algorithm inspired by the structure and function of the human brain. Neural networks consist of interrelated layers of artificial neurons that process and learn from input data. Each neuron receives input from other neurons or directly from the input data and produces an output that is propagated to the next layer of neurons until a final output is generated. Neural networks are used in various applications, including computer vision, automatic speech recognition, natural language processing, and predictive modeling. In the context of person identification using ear biometrics, neural networks can be used to learn and classify the features extracted from ear images to identify individuals with high accuracy.

a) Siamese Neural Network: Siamese neural networks are a specific type of neural network architecture utilized for tasks such as one-shot learning and similarity-based comparisons. They consist of two identical subnetworks with the same weights. These subnetworks take pairs of input images, encode them into feature vectors, and then compare these vectors to determine a similarity score. In the field of ear biometrics, siamese neural networks can be employed to analyze ear images from different individuals and as certain whether they belong to the same person. By learning the distinct characteristics of each individual's ear, the network can generate a similarity score that reflects the degree of similarity between the two ear images.

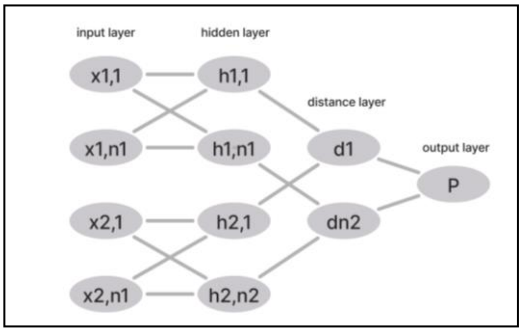

Fig.2. A 2 hidden layers Siamese network is used for binary classification.

b) Architecture of SNN: The chosen convolutional architecture for the verification task deviates from the standard Siamese network design. Instead, it incorporates a Siamese twin immediately after a fully-connected layer with 4096 units. At this stage, the Siamese architecture calculates the L1 component-wise distance between feature vectors for comparison in the verification process. This modified architecture is aimed at improving the performance of the audit task by integrating the Siamese network at a specific point in the network architecture.

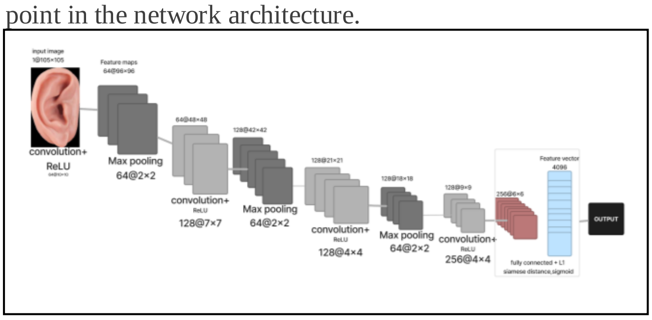

Fig.3.Architecture of Siamese Neural Network

IV.METHODOLOGY

a) Data Collection and Preprocessing: Data collection and pre-processing are essential steps in any machine learning task, including ear biometrics using Siamese neural networks. We have divided our dataset into 3 directories as Anchor, Positive, and Negative as inputs. Anchor is to compute feature representation, positive is similar to anchor, and negative is dissimilar to anchor.

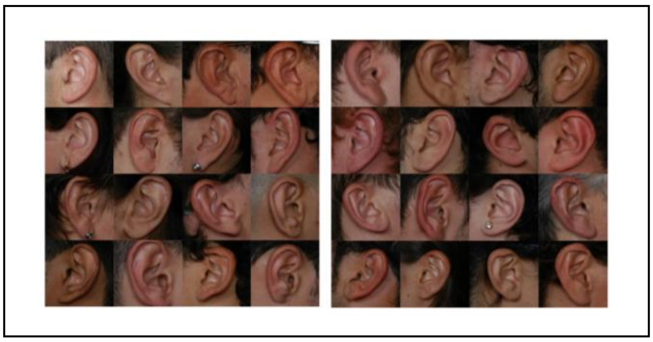

Fig.4.Sample Images of IITD II ear database

We used the IITD2 dataset for the negative. We have taken a subject and collected 2015 ear images from the subject with the help of a Python library OpenCV and resized them for our convenience. The ear image database used in this research study is sourced from IIT Delhi, New Delhi, India. The database comprises ear images collected from students and staff members of IIT Delhi. The data collection spanned from October 2006 to June 2007, and it is an ongoing process. The imaging setup employed for data acquisition was a touchless method, capturing images from a distance. The imaging took place indoors. The images in the database are assigned unique sequential numbers as identification.

Table 2: Data Representation Table

S.No	*Parameters*	*Subject 1*
1	*Number of images*	*2015*
2	*Gender*	*Male*
3	*Age*	*20*
4	*Image size*	*360x513*
5	*Resized image size*	*105x105*
6	*No.of Anchor Images*	*509*
7	*No.of Positive Images*	*506*

b)Image Preprocessing: The Python function takes a file path as input, reads the image file, decodes it using JPEG format, resizes the images to (105,105) dimensions, and normalizes the pixel value to the range [0,1].

c) Splitting train and test data: The length of the dataset will be calculated and takes 70% of the data for training and batches into groups of 16, prefetches the next 8 batches of the data to help improve training performance by reducing the time spent waiting for data to be loaded from disk.

d) Embedding Layer: This function returns a Keras model that consists of four convolutional blocks followed by a fully connected layer. The output size of a convolutional layer in a CNN can be calculated using the following mathematical formula (1).

Output size = [input size-filter size+2*padding)/stride+1 **---**(1)

Table 2: Parameters for designed network

Layer	*Output Shape*	*Learnable Parameters*
Input Layer	*(105,105,3)*	*0*
Conv layer 1	*(96,96,64)*	*19264*
Max pool layer 1	*(48,48,64)*	*0*
Conv layer 2	*(42,42,128)*	*401536*
Max pool layer 2	*(21,21,128)*	*0*
Conv layer 3	*(18,18,128)*	*262272*
Max pool layer 3	*(9,9,128)*	*0*
Conv layer 4	*(6,6,256)*	*524544*
Flatten	*(9216)*	*0*

The model summary indicates that the total number of parameters in the model is 38,960,448, and all of these parameters are trainable. This means that during the training process, the values of these parameters will be adjusted to optimize the model's performance. Overall, this model consists of multiple convolutional and pooling layers, which are commonly used in CNN architectures for image processing tasks. It learns hierarchical features from the input image and eventually produces a compact representation that can be used for classification or other downstream tasks.

e) Loss Function: The loss function that is used here is the Binary cross-entropy loss function. It is commonly used for binary classification problems and calculates the difference between the predicted probability and the actual target labels.

f) Optimizer: The optimizer used is Adam with a learning rate of 1e-4. Adam is a popular optimization algorithm used for stochastic gradient descent (SGD) which helps to improve the convergence rate and accuracy of the model. The binary cross entropy loss can be obtained using the following code snippet.

-1/N*sum (y*log (y_ hat) + (1-y) *log(1-y_hat))

N: Total number of samples in the dataset.
y: True label (0 or 1)
y_ hat: The predicted probability of the positive class

RESULTS AND DISCUSSIONS

In our ear biometrics technique, we have trained a Siamese neural network model using the Kera API in Tensor flow to perform one-shot learning. The model was trained to distinguish between images of the same ear and images of different ears based on the Euclidean distance between the embeddings of the images. Our model was trained for 50 epochs using binary cross-entropy loss function and Adam optimizer.

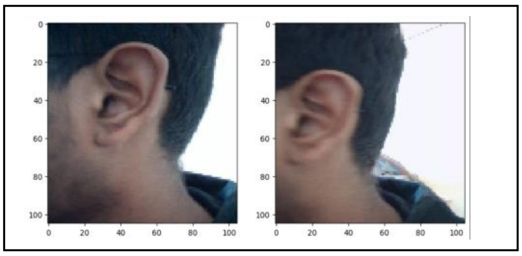

Fig.5.Predicted positive output by the Model

We achieved a precision of 95% and recall of 100% on the test set, which is a promising result for our one-shot learning approach. This indicates that our model is able to successfully distinguish between images of the same ear and images of different ears based on their embeddings.

Table 3: Performance Metrics

PERFORMANCE METRICS	*PERCENTAGE*
RECALL	*100*
PRECISION	*95*

Future work could include testing our model on larger datasets, and multiple subjects as well as exploring different types of neural network architectures and hyperparameters to improve performance. The proposed technique demonstrated high accuracy, robustness, and resistance to spoofing attacks. The proposed approach can be used in various application such as access control, forensic investigation, and healthcare.

CONCLUSIONS

In conclusion, our research project on person identification using ear biometrics with Siamese neural networks has yielded significant findings and implications for various applications. The adoption of Siamese architecture has proven to be a promising approach, showcasing its effectiveness in accurately identifying individuals based on their ear biometrics. The touchless imaging setup utilized during data acquisition holds immense practical value for applications where contactless identification is preferred. This non-invasive approach ensures convenience, hygiene, and user comfort. It opens up opportunities for deployment in various sectors such as access control systems, surveillance, forensics, and border security. Our research contributes to the advancement of biometric identification techniques by showcasing the potential of ear biometrics and Siamese neural networks. The high accuracy and robustness demonstrated in our study provide a solid foundation for the development of practical and reliable person identification systems. This opens up avenues for further research, innovation, and implementation in real-world applications, fostering a safer and more secure society.

CONFLICT OF INTEREST

In accordance with ethical guidelines, the authors declare no conflict of interest in relation to this research paper. The research was carried out with integrity and transparency, ensuring unbiased reporting of the study outcomes. The opinions, findings, and conclusions presented in this paper are solely those of the authors.

REFERENCES

[1] J. Zhang, X. Wu, and C. Zhang, "Ear recognition using wavelet transform," in 2010 International Conference on Image Analysis and Signal Processing, 2010

[2] Joshi and M. V. Joshi, "Ear biometrics: A survey of detection, feature extraction, and recognition methods," International Journal of Advanced Research in Computer Science and Software Engineering, 2013.

[3] M. M. Islam and J. H. Bappy, "A comprehensive survey on ear recognition," International Journal of Computer Applications, vol. 93, no. 10, 2014.

[4] Y. Chen and Y. Wu, "Ear recognition: A comprehensive review," Pattern Recognition, vol. 53, 2016.

[5] R. Chakraborty, N. Das, D. Das, and R. Sarkar, "Ear biometrics: A comprehensive survey," Journal of Ambient Intelligence and Humanized Computing, vol. 9, no. 2, 2018.

[6] N. Sharma and P. Gupta, "Ear recognition using deep learning: A survey," in Proceedings of the 2019 5th International Conference on Computing Sciences (ICCS), 2019.

[7] Koch, Gregory, Richard Zemel, and Ruslan Salakhutdinov. "Siamese neural networks for one-shot image recognition." *ICML deep learning workshop*. Vol. 2. No. 1. 2015.

[8] Melekhov, J. Kannala and E. Rahtu, "Siamese network features for image matching," *2016 23rd International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico, 2016

[9] Z. Su, H. Hu, and B. Zhang, "Deep learning for ear biometrics: A comprehensive survey," Journal of Visual Communication and Image Representation, vol. 71, 2020.

[10] Nikose, Shruti, and Hemant Kumar Meena. "Ear-biometrics for human identification." 2020 Advanced Computing and Communication Technologies for High-Performance Applications (ACCTHPA). IEEE