Image Processing For Face Recognition Using HAAR, HOG, and SVM Algorithms

M Maria Dominic Savio\textsuperscript{1*}, T Deepa\textsuperscript{1}, Anudeep Bonasu\textsuperscript{2}, and Talluru Sai Anurag\textsuperscript{1}
\textsuperscript{1}Department of Electronics and Communication Engineering, SRM Institute Of Science and Technology, Kattankulathur, Chennai, Tamil Nadu, India
\textsuperscript{2}California State University, Fresno, California
Email: *mariadom@srmist.edu.in

Abstract - Face recognition is one of the most active areas of research from the past two decades. Attempts are being made to understand how a human recognizes another human face. It is widely accepted that facial recognition can be based on structural information and non-structural / spatial details. In the present study, he is applying differential observations using Eigen / docking characteristics of many built-in facial features and artificial neural networks. The proposed method aims to obtain a facial feature by reducing facial features such as eyes, nose, mouth, and face depending on the importance of facial features. The face recognition system developed in this paper will inform the human face and assess the current percentage of accuracy. Therefore, this work is for human facial recognition and includes a percentage of facial expressions. The implementation of this function also offers many applications such as photography, bio-metric in bank Lockers, etc.

Keywords – Face Recognition, Image Processing, HOG, SVM, Haar Cascade

1. Introduction
The interaction of humans and robots is gaining increasing attention during research in various areas of interest over the years. In the present work to look at one of the computer observation functions involved in the design of such an efficient system \cite{1}. It have a desire to discover the human presence of the person and to follow his or her attention. Possible applications may include robot receipts, which need to perform a specific action when a person approaches or leaves the system, interactive games, sales, etc \cite{2,3}. Discovery has been a major challenge for the last few years. In computer vision, the discovery of human faces is an important research topic. It is required for many computer programs such as Human-computer interface (HCI), surveillance, human-robot interactions, etc. At this stage, the facial trajectory is gaining additional momentum for security applications and security applications to detect various situations \cite{4,5}.

Deep learning is a system design application that imitates an intersecting arrangement of a specialized cell transmitting nerve impulses, a nerve cell in an electronic device with functions comparable to those of the human brain. Interconnected surfaces accompany the deep learning algorithm \cite{6}.

- The 1st layer is named as Input Layer.
- The final layer is named as Output layer.
- There are hidden layers are present among them. The term depth implies that the network joins neurons in more than twolayers.
Neurons are present in each hidden layer. Everything is connected through Neurons. Figure 1 shows the Deep Learning Layers.

The input signals are processed and broadcasted with neurons that are above the receiver layer. The weight plays a crucial role in the signal strength given to neurons. A huge amount of input data is consumed by the network, and it is used in many layers. At each level, the complex aspects of data can be learned by the network [7-9].

A. Importance of Deep Learning
Deep learning is a commanding tool for making predictions a viable outcome. In-depth studies work well on immoderate learning and information-based forecast. Deep learning runs on Big data. When these two are merged, the coordination can gather unprecedented outcomes. Effects of time to produce, sell, manage and innovate. Deep learning can override the traditional approach. For example, when compared to machine learning, deep learning is more accurate by 41% for editing images and facial expression by 27%, and word detection by 25% [10].

B. Deep Learning Process
The deep neural network gives the positional precision of common functions, from identifying objects to voice. They can be read by themselves. Program programmers encode the specific information coded directly [11].

Each layer shows a deep amount of information, e.g., information management. The neural network is more complex with four layers than that which is having two layers. Learning takes place in several stages. Application of the input of the linear regression of the application is in the beginning stages and constructing the arithmetic representation as output. The succeeding stages target improving the representation in an arithmetic way known as an imitative. These two stages have been repeated several times till it gets to a permitted stage of perfection. The multiplication of these two stages is known as iteration [12,13].

C. Recurrent Neural Networks
The workstation has the Recurrent Neural Networks (RNN), a part of a neural network that keeps data in it, and the data sequence is allowed to extract numbers and another sequence. In common terms, an Artificial neural network is its communication between neurons, including logs [14-15]. Recurrent Neural Networks (RNN) is well-suited for processing sequences of entries:

- Signals are received at the beginning of the sentence of RNN neurons.
- "Do" is received as the input to the network and number veteran are produced. The memory to the network is given to neurons when they are transmitted back to the vector. The first instance received "Do" is remembered by the memory section and helps the network.
- The network does move side by side with the following word if the words "you" and "you
won't," then to receive each word, the nervous state is stimulated.

- The end most point comes after finding the word "a" the neural network will give the opportunities to finish the phrase. A complete RNN may have shared the possibility of a "cafe," "drink," "burger," etc.

D. Convolutional Neural Networks

To extract complex data elements at each stage unique composition designed network is used, which is a part of the neural network. CNN is also one of its parts for a good visual task Convolutional Neural Networks (CNN) are used when there is a set of random data and other operations information is needed to extract from it. For a sample, let us consider if a function is to guess a caption:

- CNN gets the photo to show that the cat, this photo, is a system, a pixel group. Usually, one layer of the greyscale image and three layers of a color image.

- During the learning feature, the network can recognize different characteristics, for example, cat's tail, ears, etc.

- The system can give possible results if the network has learned how to acknowledge a photo know a tag containing great probability will be a network forecast.

2. Existing System

The existing system highlights the significance of studying the human condition of the complication of preprogrammed sensory validation. This work submitted the EMOTIC datum, a datum of twenty-three thousand five hundred and seventy uninformed nature photos and thirty-four thousand three hundred and twenty-one people written as per their visual impressions. Photographs in a dataset are defined according to two non-identical emotion representations: 26 desk categories, three everlasting dimensions, Valence, Stimulate, and Dominance. Here it is critically reported the adjective process and checked the consistency of the adjectives of non-identical adjectives. The authors have also contributed various stats and algorithmic inspections to the data, demonstrating the EMOTIC data's features. Additionally, we have developed a basic CNN model to recognize the contextual content that includes personal information (body-boxing) and contextual information (overall photographs). The work present here also differentiates various types of coding considering position information. The results obtained indicate the importance of using situational data to monitor reactions and, in combination with EMOTIC data, stimulate more investigation in this regard.

A. Advantages

- It only uses surface area material.
- Semantic paring is used in a more complex environment.
- EMOTIC emotion recognition is suggested, but for face recognition and the application of local and global symbols is described.
- No faces can be seen in real-time.

3. Proposed System

Face identification is the common operative area of exploration for the past twenty years. Numerous experiments have been made to recognize the process of how people discern people. This is universally acquired that personal identification may be based on binary data and inconsistent/integrated data (dimensional relationships allying these features). However, particular patterns that should be amalgamated remain unclear. The fundamental idea of the submitted method is to build a facial characteristic with a face inspecting techniques elements such as eyes, ears, nose, and mouth are all different choices that establish the importance of facial features; for this work, the development of a facial recognition system that will recognize the human face and assess the percentage of accuracy. Therefore, this work helps to recognize people's faces and emphasizes the percentage of face accuracy.
A. **Advantages**

- Automatic face recognition system.
- Saves time.
- Displays the percentage of accuracy.

4. **SYSTEM ARCHITECTURE**

The necessary libraries needed for the work are installed using the terminal command window, which is used to run all the commands of the work and run the system. The data sets are for the face recognition have been collected by taking the images of the individuals. Figure 2 shows the System Architecture. Now the data sets are trained by the PC using the object detection machine learning algorithm such as HAAR to identify the correct face automatically.

![System Architecture](image)

**Figure 2: System Architecture**

Now, as the PC is ready to detect the faces trained by it, it is switched on. Whichever image file of a face or multiple faces is given as an input, it is detected, and the features are extracted. Histogram of Oriented Gradients (HOG), in addition to Support Vector Machines (SVM) algorithms, is used for face recognition and classification. As the human face matches with any of the faces in the database, then it finds a correct person with accuracy. Thus, this work helps in the effective execution of automatic face recognition systems, utilizes the latest face technology as well as saves time. There are mainly three modules in the proposed system.

a. **Face Detection Module**

b. **Face Recognition Module**

c. **Comparison Module**

A. **Face Detection Module**

For detecting the face, the HAAR Cascade algorithm is used in this paper. HAAR cascade is one of the deep learning algorithms which can associate things. It is found on the theory presented by Paul Viola also Micheal Jones in their paper "Rapid Object Detection with a Boosted Cascade of Simple Features." Thus, popularly known as the Viola-Jones algorithm. It is based on a resemblance where a cascade function is instructed using to a great extent positive images also negative images. It also associates the things in many images. It has four stages: The algorithm has four stages:

1. Selection of HAAR features
2. Creation of Integral Images
3. Adaboost Training
4. Cascading Classifiers

Let us take the example of Image identification. Starts the design requires more productive photos with proper features, also a set of negative images which do not contain any faces. After gathering all
the required images, we need to extract some features from the collected face images.

It all depends on the identification; an opening for a particular chosen dimension is then shifted on top of the given photo. At every section for a given photo, HAAR characteristics are then studied. Then depending on the threshold, objects are separated from the non-objects. Each HAAR characteristic is considered as a weak classifier individually, and a set of HAAR characteristics will be needed for classifying the article along with particular efficiency, so they are combined to get a strong classifier.

Figure 3: Example of Haar Classifier

Figure 3 shows the Example of the Haar Classifier. In each stage, the classifiers classify the image from the sliding window as either a positive image or a negative image. Positive indicated that the required part of the image is found, and negative indicates that no image has been found.

B. Face Recognition Module

For face recognition Histogram of Oriented Gradients is being used in this paper. Hog works on the descriptor technique by counting the gradient orientation in the given images' localized parts, which are also called Region of Interest.

HOG Implementation:

- By separating every cell into a V-shaped bucket by the orientation of the gradient.
- Each cell's pixel contributes a weighted gradient to its matching angular bin.
- The descriptor consists of a block of histograms; a standardized group of histograms characterizes a block of histograms.

Figure 4: Example of HOG
C. Comparison Module
Support Vector Machine (SVM) is applied for the comparison module in this paper. It is a supervised machine learning algorithm that can be used in resolving classification challenges. Figure 4 shows the Example of HOG. In machine learning, support vector machines are applied as categorizing and reverting exploration. If whether the position of instruction representative containing labeled data is given, then SVM can be used to add new examples to this training data by classifying the gives images into the various available categories. So, it is also a non-probabilistic binary linear classifier. In addition to performing classifications mentioned above, SVMs can also be applied within a nonlinear classification. Supervised learning is not possible without having labeled data. In such a case unsupervised learning approach is used. It searches for clustering of data in groups and then maps the data to the groups it identified. Figure 5 shows an example of SVM. The clustering algorithm of support-vector applies the analysis from support vectors by classification of unlabeled datum, and it is the existing well-liked design known for the manufacturing implementation.

Figure 5: Example of SVM

5. Results
As seen from the figures, figure 6a shows Single Face Recognition Output, and figure 6b shows Two-Face Recognition Output, figure 6c shows Three Face Recognition Output, Figure 6d shows Four Face Recognition Output, Figure 6e shows Real-Time Face Recognition Output. The proposed method can recognize faces from the given input images and can be used for real-time facial recognition. It can also detect faces even if there is more than one face in the given input images.

Figure 6a: Single Face Recognition Output

Figure 6b: Two-Face Recognition Output
6. Conclusion
In this paper, the design of a face recognition system can be used to recognize the human's face and check the presence of the face with high accuracy. This work reduces the human effort of person identification in many commercial environments and avoids manual work. Thus, this system helps to recognize the face of the people and shows the face with high accuracy with its name defined below.

References
[1]. Deep, E. B. E. R. U. (2021). Consumer Emotional State Evaluation Using EEG Based Emotion Recognition Using Deep Learning Approach. In Advanced Computing: 10th International Conference, IACC 2020, Panaji, Goa, India, December 5-6, 2020, Revised Selected Papers, Part I (p. 113). Springer Nature.
[2]. Farzaneh, A. H., & Qi, X. (2021). Facial Expression Recognition in the Wild via Deep Attentive Center Loss. In Proceedings of the IEEE/CVF Winter Conference on Computer Vision Applications (pp. 2402-2411).
[3]. Zhang, J. (2021, February). Children’s Face Recognition Based on Convolutional Neural Network. In Journal of Physics: Conference Series (Vol. 1744, No. 3, p. 032013). IOP Publishing.
[4]. Tamilkodi, R. (2021). Automation System Software Assisting Educational Institutes for Attendance, Fee Dues, Report Generation Through Email and Mobile Phone Using Face Recognition. Wireless Personal Communications, 1-18.
[5]. Asadi Amiri, S., & Rajabinasab, M. (2021). Face recognition using Color and Edge Orientation Difference Histogram. Journal of AI and Data Mining.
[6]. Larik, S. Face Recognition for Automated Attendance using HOG & Machine Learning.
[7]. Kumar Shukla, R., & Kumar Tiwari, A. (2021). Comparative Analysis of Machine Learning-Based Approaches for Face Detection and Recognition. Journal of Information Technology Management, 13(1), 1-21.
[8]. Huang, D. Y., Chen, C. H., Chen, T. Y., Hu, W. C., Guo, Z. B., & Wen, C. K. (2021). High-efficiency face detection and tracking method for numerous pedestrians through face candidate generation. Multimedia Tools and Applications, 80(1), 1247-1272.
[9]. Jayasimha, Y., & Reddy, R. V. S. (2021). A facial expression recognition model using hybrid feature selection and support vector machines. International Journal of Information and Computer Security, 14(1), 79-97.
[10]. Larik, S. Face Recognition for Automated Attendance using HOG & Machine Learning.
[11]. Yang, R., Wang, Y., Xu, Y., Qiu, L., & Li, Q. (2021). Pedestrian Detection under Parallel Feature Fusion Based on Choquet Integral. Symmetry, 13(2), 250.
[12]. Hussain, S. A., & Al Balushi, A. S. A. (2020). A real-time face emotion classification and recognition using a deep learning model. In Journal of Physics: Conference Series (Vol. 1432, No. 1, p. 012087). IOP Publishing.
[13]. Ambre, S., Masurekar, M., & Gaikwad, S. (2020). Face recognition using Raspberry Pi. In Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough (pp. 1-11). Springer, Cham.
[14]. Hu, M., Zheng, Y., Yang, C., Wang, X., He, L., & Ren, F. (2019). Facial expression recognition using fusion features based on the center-symmetric local octonary pattern. IEEE Access, 7, 29882-29890.
[15]. Marsot, M., Mei, J., Shan, X., Ye, L., Feng, P., Yan, X., ... & Zhao, Y. (2020). An adaptive pig face recognition approach using Convolutional Neural Networks—computers and Electronics in Agriculture, 173, 105386.