International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 11 | Nov 2022 www.irjet.net p-ISSN: 2395-0072

A Comparative Study on Identical Face Classification using Machine Learning

Joshi Jaya[1], Saket Swarndeep[2]

[1]*Student of Masters of Engineering, Ahmedabad, Dept of Computer Engineering, L.J. Institute of Engineering & Technology, Gujarat, India*
[2]*Assistant Professor, Ahmedabad, Dept of Computer Engineering & Technology, Gujarat, India*

Abstract - *This paper presents a new identical face classification using Support Vector machine (SVM) classification tree. Automatic face classification from facial image attains good accuracy with large set of training data. While face attribute classification from facial image still remains challengeable. In this paper, we improve the same faces features. Each SVM acts an independent membership/non-membership classifier and several SVM are combined in a plurality voting scheme that chooses the classification made by more than the half of SVMs. The global systems classify faces by classifying a single feature vector consisting of the gray values of the whole face image. For a good encoding of face efficient reduction of data dimension and strong separation of different faces, respectively. Next, the SVM ensemble is applied to authenticate an input face image whether it is included in the membership group or not. We compare the SVMs based classification with the standard identical approach.*

Key Words: **Face Classification, Haarcasecade, Wavelet Transform, Support vector Machine, Random Forest, Logistic Regression**

1. Introduction

It is well classification that the use of face pictures for personal surveillance, identification and verification is a rapidly expanding field of study in many Computers vision application. Some examples in this area are face recognition face action, classification, faces recognition, skin color classification, age estimation, gender recognition and ethnicity recognition. Face classification has achieved better results according to the research done for nearly three decades. However, similar accuracy of classification could not be gained from facial attribute recognition.

The most effective and powerful classifier for pattern recognition is found in the human brain. Its amazing power comes from the fact that it is a dynamic organ engaged in training and learning for a redetermined amount of time. In order to achieve the same or better results, researcher's key goals is to replicate this biological and behavioral trait of the human brain in artificial neurons. The classification of human characteristic like identical faces, age gender and ethnicity using facial photographs is one of the main applications of this work. In this study, we aim to develop a precise technique for identifying the same faces from facial photographs.

Identical face classification is done according to geometric difference of primary feature in male and female. This algorithm can classify the facial images into faces features. Identical classification is based on the texture variation of eye lids, wrinkle density in the forehead, lips and cheek area. Classification is done using SVM and two separate neural networks for identical face classification.

Fig-1: Similar faces in male[1]

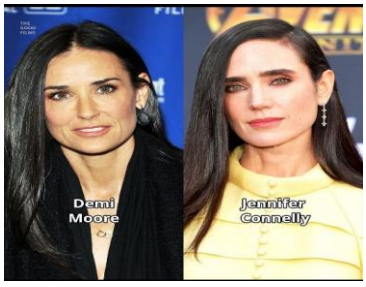

Fig-2: Similar faces in female[1]

1.1 Machine Learning:

Machine learning is a growing technology which enables computers to learn automatically from past data. Machine learning uses various algorithms for building mathematical models and making predictions using historical data or information. Currently, it is being used for various tasks such as image recognition, speech recognition, email filtering, Facebook auto-tagging, recommender system, and many more.[2]

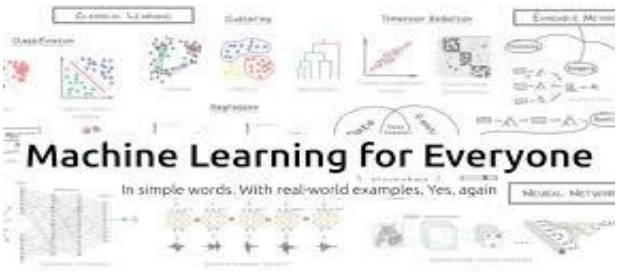

Fig-3 Machine Learning Model [3]

1.1.2 Machine Learning Techniques:

- **Supervised Learning:**

  It is defined by its use of labelled datasets to train algorithms that to classify data or predict outcomes accurately [4]. Supervised Algorithms are Naïve Bayes, Logistic Regression, Support Vector Machine, Random Forest and many more.

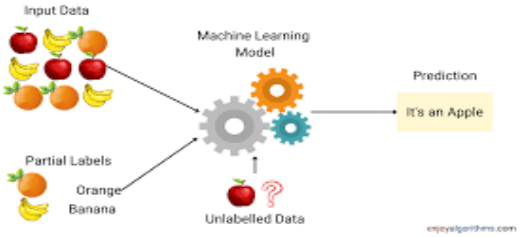

Fig-4: Supervised Learning Model [4]

- **Unsupervised Learning:**

  These algorithms discover hidden patterns or data groupings without the need for human intervention. Unsupervised Algorithms are Association and Clustering**.**

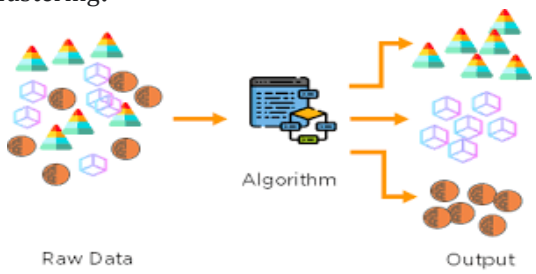

Fig-5: Unsupervised Learning Model [5]

- **Reinforcement Learning**:

  It is an area of machine learning concerned with how intelligent agents ought to take actions in an environment in order to maximize the notion of cumulative reward. Machine makes observation of rewards and acts accordingly.

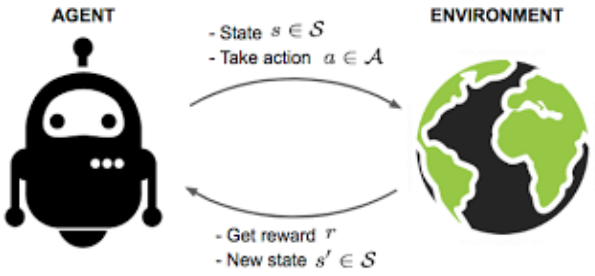

Fig-6: Reinforcement Learning Model [6]

1.1.3 Machine Learning Algorithm:

1. **Logistic Regression:** It is often used for classification and predictive analytics. Logistic regression estimates the probability of an event occurring, such as voted or didn't vote, based on a given dataset of independent variables. Since the outcome is a probability, the dependent variable is bounded between 0 and 1. It uses the activation function called Sigmoid.

$$log\left[\frac{y}{1-y}\right] = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$$

Fig-7: Logistic Regression Equation [7]

Sigmoid Curve:

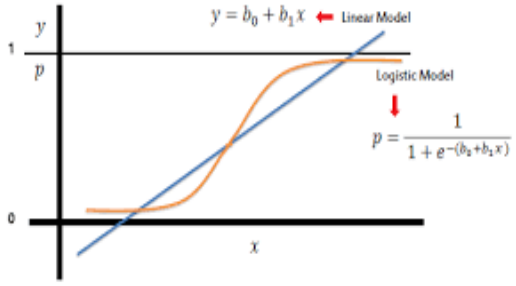

Fig-6: S Curve for Sigmoid [8]

2. **SVM Algorithm:**

   Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane. SVM chooses the extreme points/vectors that help in creating the

hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine. Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane [9]:

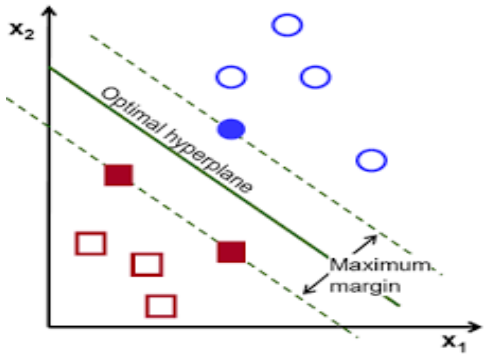

Fig-9: SVM Working [10]

$$R(\alpha) = R_{train}(\alpha) + \sqrt{\frac{f(h)}{N}}$$

$$f(h) = h + h\log(2N) - h\log(h) - c$$

$$Margin = \rho$$

$$Relative\ Margin = \frac{\rho}{D}$$

$$h \leq \min\left(\left\{d, \left\lceil\frac{D^2}{\rho^2}\right\rceil\right\}\right) + 1$$

Fig-10: SVM Equations [11]

3. Random Forest Classifier:

It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset. Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.[12]

The below diagram explains the working of the Random Forest algorithm:

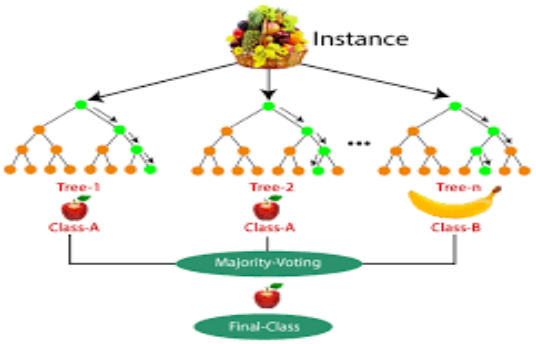

Fig-11: Random Forest Working [12]

1.2 Literature Survey:

Reference [13] has proposed a novel method using support vector machine, Multilayer Perception and KNN as representation of facial expression classification is implemented as a classifier for expression SVM deals with the original image by preserving the useful information and reducing the feature vector's dimension of data. The authors proposed the method of SVM for extracting the quantity of the face feature that has grate ability in improving the generalization performance and the accuracy of the feature vector. According to this proposed method is the improved method of SVM that work through changing the feature vectors for producing the transformation of high dimensional data in to the low dimensional data. The authors found that the classifier of SVM, KNN method is very effective for identifying the other expression. The classification rate gained 93.89% on database which is applied on eight expressions.

Reference [14] present a CFA system to classify face emotions. It utilized the method for feature extraction and used machine learning algorithm for classification which are PCA, LFA, ICA, Naïve and compared their results. The present method applied on image sequences of subjects. The classification accuracy gained 79.3%, while using PCA, 81.1% while using LFA and 95.5% ICA.

Reference [15] in there's literature work the authors develop system using deep learning methods. It combines the results bases on age and gender Neural Network and tries to predicts the results. In Gender classification, it provides accuracy of 87.06%. In age classification, it gives accuracy of 79.09%.

Reference [16] in there's literature work authors using Machine Learning algorithm i.e. Gaussian Mixture Model (GMM) for face age estimation present promising results. And Hidden Markov Model (HMM) supervector to represent face image patches. The dataset containing two datasets of facial images, one with 4000 images of 800 females and the others with 4000 images of 800 males. Each subset has the age range of 0 to 93. Using

GMM they got the accuracy 72.46%, and HMM the accuracy 63.35% in female. Using GMM they got accuracy 58.53% and HMM the accuracy 53.97% in male.

Reference [17] in there's Literature works the authors develop system using deep learning methods. It combines the results bases on Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. It provides accuracy of 95.4% using Convolutional Neural Network.

1.3 Data Description:

Dataset is being used which is downloaded from internet, consists of 789 images for different 8 persons. The images are having different angle of faces, different expression.

1.4 Proposed System:

A Proposed System uses various algorithms such as Logistic Regression, SVM Classifier, Random Forest Classifier. Our Proposed system has several phases in-order to improve the Accuracy of results.

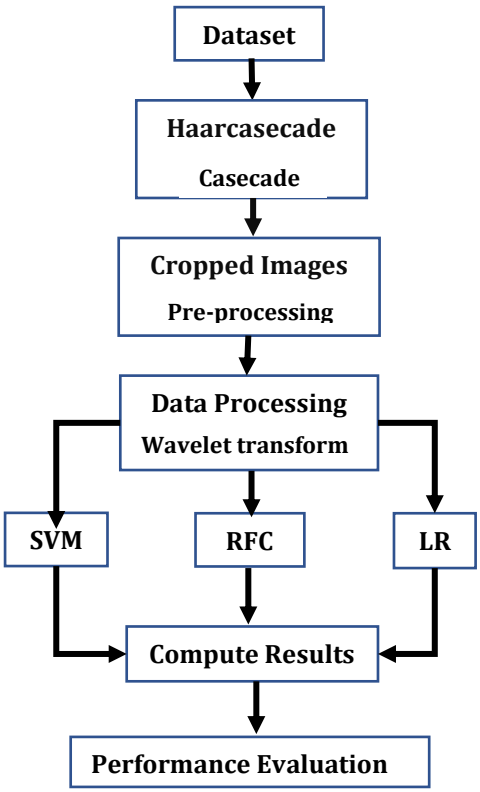

Fig-10: Proposed Model

Step 1: The first step is to load the images and transform the images into the gray color.

Step 2: The next step is casecade the images using opencv.

Step 3: In this step crop the all images by length and width of face using eyes lids and lip.

Step 4: The next step to pre-processing the data and wavelet transformation of images.

Step 5: In this step compute the results and performance of each algorithm is evaluated to know the best model for classification.

1.5 Conclusion:

The machine learning algorithm for classification procedure has discussed here. Thus or proposed algorithm has concluded that the overall performance of SVM is high to get accurate classification results for IFC. The appropriate evaluation is needed to determine the classification results of proposed architecture. In our proposed architecture the most relevant feature has been used and tested for the classification, besides the dataset is very clear due to noise remove the filters. The irrelevant dataset should be introduced and tested for the overall performance of the proposed architecture. To obtain better efficiency, the SVM method will be validated for the huge amount of dataset [18]. Machine learning models can be trained to act well as we humans do and learn from nature.

References:

[1] Aanmelden of registreren om te bekijken. (n.d.). https://www.facebook.com/unsupportedbrowser?igshid=YmMyMTA2M2Y=

[2] Machine Learning Tutorial | Machine Learning with Python - Javatpoint. (n.d.). www.javatpoint.com. https://www.javatpoint.com/machine-learning

[3] machine learning - Google Zoeken. (n.d.). https://www.google.com/search?q=machine+learning

[4] Supervised Machine learning - Javatpoint. (n.d.). www.javatpoint.com. https://www.javatpoint.com/supervised-machine-learning

[5] Sharma, M. (2021, December 11). A Beginners Guide to Unsupervised Learning - Analytics Vidhya. Medium. https://medium.com/analytics-vidhya/beginners-guide-to-unsupervised-learning-76a575c4e942

[6] Weng, L. (2018, February 19). A (Long) Peek into Reinforcement Learning. Lil'Log. https://lilianweng.github.io/posts/2018-02-19-rl-overview/

[7] Logistic Regression in Machine Learning - Javatpoint. (n.d.). www.javatpoint.com. https://www.javatpoint.com/logistic-regression-in-machine-learning

[8] Logistic Regression. (n.d.). https://www.saedsayad.com/logistic_regression.htm

[9] Support Vector Machine (SVM) Algorithm - Javatpoint. (n.d.). www.javatpoint.com. https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm

[10] Team, T. A. I. (2020, June 19). Support Vector Machine—Insights. Towards AI. https://towardsai.net/p/machine-learning/support-vector-machine-an-insight-cdae000758dc

[11] Agarwal, A. (2022, June 13). Support Vector Machine — Formulation and Derivation. Medium. https://towardsdatascience.com/support-vector-machine-formulation-and-derivation-b146ce89f28

[12] Machine Learning Random Forest Algorithm - Javatpoint. (n.d.). www.javatpoint.com. https://www.javatpoint.com/machine-learning-random-forest-algorithm

[13] Dino, H. I., & Abdulrazzaq, M. B. (2019, April). Facial expression classification based on SVM, KNN and MLP classifiers. In *2019 International Conference on Advanced Science and Engineering (ICOASE)* (pp. 70-75). IEEE.

[14] Donato, G., Bartlett, M. S., Hager, J. C., Ekman, P., & Sejnowski, T. J. (1999). Classifying facial actions. IEEE Transactions on pattern analysis and machine intelligence, 21(10), 974-989.

[15] Kalansuriya, T. R., & Dharmaratne, A. T. (2013, December). Facial image classification based on age and gender. In *2013 International Conference on Advances in ICT for Emerging Regions (ICTer)* (pp. 44-50). IEEE.

[16] Zhuang, X., Zhou, X., Hasegawa-Johnson, M., & Huang, T. (2008, December). Face age estimation using patch-based hidden markov model supervectors. In *2008 19th International Conference on Pattern Recognition* (pp. 1-4). IEEE.

[17] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE signal processing letters*, *23*(10), 1499-1503.

[18] Chen, J. I. Z., & Hengjinda, P. (2021). Early prediction of coronary artery disease (CAD) by machine learning method-a comparative study. *Journal of Artificial Intelligence*, *3*(01), 17-33.