



# Face Classification from Ethnicity Identification using Face Images

S. Priyanka<sup>1</sup>, C. Christy<sup>2</sup>, Dr. S. Arivalagan<sup>3</sup>  
Research Scholar<sup>1</sup>, Assistant Professor<sup>2,3</sup>

PG & Research, Department of Computer Science<sup>1,2</sup>, Department of Computer Science & Engineering<sup>3</sup>  
St. Joseph's College of Arts & Science (Autonomous), Tamilnadu, Cuddalore, India<sup>1,2</sup>  
Annamalai University, TamilNadu, Chidambaram, India<sup>3</sup>

## Abstract:

Data mining is the developing area of studies that have been attracting many styles of research to a chunk of information from the dataset. The photography category is to categories the numerous instructions of their visual traits. To classify the human face and to discover which foundation they belong to. The particular ethnicity of a separate can be vastly useful in face images and personal identification. A representation of the ethnic classification of a person from face images. Anthropoid is able to facial images demographic feature involve ethnicity or race and gender. The process used for the classifier human facial image namely Naive Bayesian, Random Forests, and Decision Tree into different ethnic groups. Identification is used to identify the face from the face images. Rather such images are affected by noise, shadow, surrounding element like a light that times correct identification is difficult. In this reason support to individual fields such that anthropology as have taken identification of humans and pictures if humans. The most critical characteristics in particular ethnicity are an idea to be the facial organs such as eyes, nose, forehead, and mouth. In a various location of features in well-defined images, the corner points of the features are extracted because corner points are the same in all images of a person. In which adequate filter is applied to minimize noise as possible is noise reduction.

**Keywords:** Ethnicity classification, Facial features, Picture identification, Decision Tree.

## I. INTRODUCTION

Data mining also called knowledge to find in databases, in computer science, the process of finding interesting and useful patterns and relationships in a huge volume of data. [1] A procedure which uses a combination of information evaluation gear to find styles and relationships in facts that may be used to nicely-described predictions. To analyze records from numerous specific dimensions, classify it, and the repository and databases followed into system enthusiastically. The objective of building a computer system is adaptive to environments and learning from researchers in many areas, such as engineering, computer science, physics, and neuro science mathematics. [2] This research has come to a variety learning method that transforms scientific fields. Image type hopes to discover a document the photos in a single magnificence and to decide these snapshots from all the different classes. The maximum regularly used strategies for extracting facts from the image. The technique of sorting and categorizing data into numerous types, paperwork or some other distinct class. [5] Data class permits the separation and class of information in keeping with statistics set requirements for diverse business or personal targets. [3] A human face is easier to understand that provides demographic information such as racial features ethnicity identification of a person. [8] Image classification is one of the important area in image mining it growth the demand for developing real-world vision system. Image classification is one of the basic tasks of image mining. Image classification has two leading stages, the first stage defines an effective representation of an image. It includes required information of the image for classification. The second stage classifies the new image minus error.

## II. CLASSIFICATION OF RACE

To classify human on the basis of the material attribute is difficult, also the coexistence of races through defeat, capture,

defection and mass deportations has produced a heterogeneous world population. Limiting the criteria to such features as skin pigmentation, color and form of hair, eye form, shape of head, stature and nose form, most anthropologists historically agreed on the fact of three relatively decided variety the Hispanic, Asian Americans, and white and European Americans. [6] The Hispanic, found in Cuban, Mexican and South or Central American is indicated by pure white to brownish skin, short and cute blue eye, short and broad nose, medium lips, medium height, a broad head form. The hair is to dark shiny black in color. [4] Asian Americans race basically in East Asia, South Asia, or Southeast Asia is indicated by a fair white skin, high-bridged noses, a small V-shaped face and double eyelids. Whiteface, medium nose and black hair. The most indicate of white and European, inhabitants of white racial groups of Europe, the Middle East and North Africa. Indicate by the brownish skin, short and broad nose and black hair.

## III. LITERATURE REVIEW

Hajra Momin and Jules-Raymond Tapamo [1] A Comparative study of a Face Components based Model of Ethnic Classification using Gabor Filters. An essential role in face-related applications is ethnicity and gender. The ethnicity of a separate can be vastly useful in face recognition and person identification. Face images are such components can be used eyes, nose and mouth and then applies Gabor filters to extract the facial features. To classify the human face image into different ethnic groups are applied such as naïve Bayesian (NB), Multilayer Perceptron, support vector machine (SVM). Classification is a task in two classes Asian and Non-Asian. Ayushi [2] A Survey on Feature Extraction Techniques. Face recognition is vastly essential in computer vision. Face recognition has two senses one is the biometric identification and the second is the visual perception of the intimate face. Some issue in face recognition involves that face

identification, feature localization, corner point extraction, and scaling. Qeethara Al-Shayea1 and Muzhir Al-Ani [3] Biometric Face Recognition Based on Enhanced Histogram Approach. Biometric face recognition including image processing and to analyze a facial structure. The extensive attribute such as eyes, nose and mouth, a viewpoint of features such as the jaw and forehead with the various parts of the face. Implemented system is divided into the steps pre-processing, histogram technique, feature extraction and face verification. Harish Kattia, S.P.Aruna [4] Are you from North or South India? A hard race classification task reveals systematic representational differences between humans and machines. A huge variety of judgments on faces differ from gender, race, personality, emotional state etc. race classification problem on Indian face originating from northern or southern India. Weenakorn Ieosanurakand Watch arin Klongdee [5] Face Classification using Adjusted Histogram in Grayscale. An implicit classification method for a face grayscale image constructs on the histogram in grayscale of face images. The methodology used for the grimace, ORL and Jaffe databases. Classification method which is classified by the formula of Euclidean distance and kullback-leibler divergence. S.Md Mansoor Roomi, S.L.Virasundarii, S.Selvamegala, S.Jeevanandham, D.Hariharasudhan [6] Race Classification Based On Facial Features. A human face is easier to understand that provides demographic information such as racial features ethnicity identification of a person. The methodology used for the Gaussian the races. Race classification is divided into the Caucasoid, Negroid and mongoloid. The major features of the race classification are an eye, nose, color hair, forehead etc. Teodora Sechkova Krasimir Tonchev Agata Manolova [7] Action Unit recognition in still images using graph-based feature selection. Facial expression is variable demographic of race, ethnicity, culture, nationality, age, gender, religion. An algorithm used for support vector machine (SVM) and Nearest Neighbor (NN), supervised Gradient Descent(SDM) techniques for classification. Xiaoguang Lu and Anil K. Jain [8] Ethnicity Identification from Face Images. Human face provides demographic information such as ethnicity and gender. The methodology used for linear discriminate analysis (LDA). LDA ensembles at multiple scales. Identification is finished for the 2 magnificence Asian and Non-Asian ethnicity class mission Classification is divided into three racial groups of North American Caucasian, African-American and Chinese. Srinivas Gutta, Jeffrey R. J. Huang, P. Jonathon and, and Harry Wechsler[9] Mixture of Experts for Classification of Gender, Ethnic Origin, and Pose of Human Faces. The mixture of experts in gender and ethnic classification of a human face and pose classification of facial images. The techniques used for radial basis functions (RBFs),decision tree(DT) and support vector machine (SVM). They collected data from FERET databases the labels were male and female and the ethnic labels were Caucasian, South Asian, East Asian, and African.

#### IV. METHODOLOGY

The purpose of this paper is to apply a few strategies to represent the human face for the ethnic type. After the features have been extracted from the face, distinctive classifiers are used to assess the effectiveness of techniques while carried out to the face components.

##### A. DECISION TREE

The number one motive of any idea-mastering induction device is to bring together regulations for classifying gadgets

given an education set of gadgets whose magnificence labels have seemed. In the formalism used right here, all gadgets are defined via a hard and fast collection of attributes, every with its very own set of discrete values and every item ("face") belonging to one in each of numerous education. The decision tree employed is Quinlan's C4.Five and it makes use of a statistics-theoretical method based on entropy. C4. Five builds the selection tree using a top-down, divide-and-conquer method: select out a feature, divide the schooling set into subsets characterized thru the feasible values of the feature, and follow the identical partitioning system recursively with every subset until no subset contains items from multiple elegance. The single-magnificence subsets correspond then to the leaves of the choice tree. The entropy-based criterion that has been used for the choice of the attribute is called the benefit ratio criterion.

##### B. RANDOM FOREST

An ensemble classifier the usage of many selection tree models.Can be used for category or Regression. Accuracy and variable importance information are supplied with the results.

A Random Forest is a classifier along with a collection of tree-based classifiers  $h(x, \Theta_k)$ , okay = 1...In which the  $\Theta_k$  are independent, identically allotted random bushes and every tree casts a unit vote for the final category of input  $x$ . Like CART, Random Forest makes use of the Gini index for determining the very last elegances in every tree. The very last elegance of every tree is aggregated and voted through weighted values to assemble the very last classifier. Gini Index Random Forest uses the Gini index taken from the CART gaining knowledge of gadget to construct selection trees. The Gini index of node impurity is the degree maximum usually selected for type-type problems. If a dataset  $T$  carries examples from  $n$  lessons, Giniindex,  $Gini(T)$  is defined as in which  $p_j$  is the relative frequency of sophistication  $j$  in  $T$ . $Gini(T)=1-\sum_{j=1}^n (P_j)^2$  If a dataset  $T$  is cut up into two subsets  $T_1$  and  $T_2$  with sizes  $N_1$  and  $N_2$  respectively, the Gini index of the split records includes examples from  $n$  lessons, theGini index ( $T$ ) is described as. The attribute fee that provides the smallest SPLIT  $Gini(T)$  is chosen to split the node. $Ginisplit=$

$$\frac{N_1}{N} gini(T_1) + \frac{N_2}{N} gini(T_2)$$

##### C. NAIVE BAYES

Naïve Bayes Called Naïve Bayes because it's based on "Bayes rule" and "naively" assumes independence given the label it is most effective valid to multiply chances whilst the occasions are an independent Simplistic assumption in real existence despite the name, Naïve works nicely on real datasets Naïve Bayes Classifier Simple probabilistic classifier primarily based on applying Baye's theorem (from Bayesian statistics) sturdy (naive) independence assumptions. A greater descriptive time period for the underlying possibility version could be "unbiased characteristic version". Naïve Bayes Classifier Assumes that the presence (or absence) of a specific characteristic of a category is unrelated to the presence (or absence) of a few different features. Example: a fruit may be taken into consideration to be an apple if it's miles red, spherical, and approximately 4" in diameter. Even if those capabilities depend on each exceptional or upon the existence of the opposite capabilities, a naive Bayes classifier considers all of those residences to independently contribute to the hazard that this fruit is an apple.

#### V. CONCLUSION

Data mining is the growing area of research that had attracted many sorts of research to a bit of record from the dataset. To

classify the human face and to locate which starting place they may belong. For evaluation, the image regions are taken from a photo. The primary concept of the image kind is to categorize the diverse schooling of their visible traits. To become aware of the starting place of a however picture and to classes that photograph. Color identifies primarily based strategies are used to select out the coloration of the pores and pores and skin. To analyze the photo that to pre-processing the data after that function extraction is offering higher class usual overall performance.

## VI. REFERENCES

- [1]. HajraMomin and Jules-Raymond Tapamo, (2016). "A Comparative study of a Face Components based Model of Ethnic Classification using Gabor Filters", Applied Mathematics & Information Sciences An International Journal 10, No. 6, 2255-2265.
- [2]. Ayushi,(2013). "A Survey on Feature Extraction Techniques", International Journal of Computer Applications, Volume 66, No.11, 0975 – 8887.
- [3].Qeethara Al-Shayea1 and Muzhir Al-Ani, (2018). "Biometric Face Recognition Based on Enhanced Histogram Approach", International Journal of Communication Networks and Information Security (IJCNIS) Vol. 10, No. 1, 148-154.
- [4].HarishKattia, S.P. Aruna, (2018). "Are you from North or South India? A hard race classification task reveals systematic representational differences between humans and machines", arXiv: 1703.07595v2 [cs.CV], 1-12.
- [5]. Weenakornleosanurak, and WatcharinKlongdee, (2018). "Face Classification using Adjusted Histogram in Grayscale", Proceedings of the International MultiConference of Engineers and Computer Scientists, Vol IIMECS, 1-3.
- [6].S.MdMansoorRoomi, S.L.Virasundarii, S.Selvamegala, S. Jeevanandham, D.Hariharasudhan, (2011). "Race Classification Based On Facial Features", Third National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics, Vol.19, 54-57.
- [7]. TeodoraSechkova, KrasimirTonchev, AgataManolova, (2016). "Action Unit recognition in still images using graph-based feature selection", IEEE 8th International Conference on Intelligent Systems, Vo.10, 646-650.
- [8]. Xiaoguang Lu and Anil K. Jain,(2010). "Ethnicity Identification from Face Images",1-10.
- [9]. SrinivasGutta, Jeffrey R. J. Huang, P. Jonathon and, and Harry Wechsler,(2000). "Mixture of Experts for Classification of Gender, Ethnic Origin, and Pose of Human Faces", IEEE Transactions on Neural Networks, Vol. 11, No. 4,948-959.