Detecting Human and Classification of Gender using Facial Images MSIFT Features Based GSVM

G D K Kishore, BabuReddy Mukamalla

Abstract: Classification of gender using face recognition system is an essential concept for different types of applications in human-computer interaction and computer-aided related applications. It defines a wide range of features from human images to detect male, female and others using real-time data. There are different machine learning approaches were implemented to classify gender and also detects other images during the classification phase, which are not humans based on features extracted from human images datasets. All these existing techniques mostly depend on controlled conditions like features and other representations of the human image. Because of significant and uncertain variations of a particular image, it may be a challenging task in gender classification for real-time image processing application, whether it is male, female and others. So that in this document, we propose a Human detection and Face based gender Recognition System (HDFGR); to investigate male or female classification on real life faces using real world face databases. Our proposed approach consists Multi-Scale Invariant Feature Transform (MSIFT) to describes faces and Gaussian distance-based support vector machine (GSVM) classifier is used to classify gender and objects, i.e. male, female and other from features extracted human image datasets. We obtain an experimental performance of 98.7% by applying DSVM with boosted MSIFT features. Our proposed approach gives better classification accuracy and other performance parameters compared to different existing approaches with benchmark and evaluation of possible databases.

Index Terms: Gender Classification, Scale Invariant Feature Transform, Support vector machine, classification accuracy, Adaboost, Feature extraction.

I. INTRODUCTION

Gender classification [1][19][21][24] is the attention to identify the gender and classifying male and female for real time applications. Main aims of the classification of the gender of a particular person based on their characteristics, which are differentiates femininity and masculinity. In artificial intelligence related applications, classification of gender is one of the most pattern recognition methods. In many potential related application types of research, gender classification progressed; for instance gender classification with computer systems can be used in different human-computer interaction (HCI), industry and security related applications[4] [5], business advancement [6], and portable and computer related gaming applications[15] [7].

Moreover, multi-systems are proposed to upgrade the execution of gender acknowledgement as far as both Precision and recall are concerned. To date, nonetheless, few related reviews exist concerning human general orientation acknowledgement and characterization. In [9] author completed a near investigation of the gender classification techniques triggered by detecting the face location and ongoing face location.

To comprehend the flow, look into difficulties, advancement, and chances of gender classification. The underlying and typical issue to investigate to explore face images with different conditions which are present in the database (FERET) [14] to be considered to face detection based gender classification either it is genuine or not genuine for different face images at different situations. As can be viewed, there is significant appearance minor departure from genuine faces, which incorporate outward appearances, brightening changes, head present varieties, impediment or make-up, poor picture quality, etc. Consequently, general orientation acknowledgement, in actuality, faces is significantly more difficult contrasted with the case for countenances caught in obliged conditions. Barely any investigation in the writings has tended to this issue. Different authors revealed the execution of 81.22% on the classification of gender applications; in any case, as they, for the most part, centred it on face verification, they did not wholly examine general orientation acknowledgement on reality faces. So that, in this document, we propose a Human detection and Face based gender Recognition System (HDFGR); it is an efficient non-parametric approach which is summarized the basic structure of an image with MSIFT features[31] for face analysis. Experimental results of the proposed approach give efficient gender classification results with accuracy, time etc.

II. MODIFIED SCALE INVARIANT FEATURE TRANSFORM

MODIFIED Scale Invariant Feature Transform (MSIFT) is an algorithm to employ a system to extract special and specified features of input images for different matching related applications and identifying objects [10] present in the input image. They obtain different features with invariant scalable rotations and partially change invariant lighting schemas to extract features with four steps

In initial step [17], Deference of Gaussian (DOG) filtering approach at every factor of pixel present in the image with explored to find invariant relations in images based on the selected region. In the following phase, find appropriate pixel points based on different durability requirements of image.

Revised Manuscript Received on September 15, 2019

G D K Kishore, Research Scholar, Computer Science, Krishna University, India
BabuReddy Mukamalla, Assistant Professor, Computer Science, Krishna University, India

DOI:10.35940/ijrte.B3782.098319

1466
Lastly, data found based on features with different factors which are relevant to a pixel with different features. Converting picture data into scale-invariant dimensions concerning the selected region with different functions. It is vector representation with 128-pixel dimensions identified in images.

Picture coordinating is a fundamental procedure in picture handling and assumes a critical job in photograph grammetry and remote detecting forms. A filter is a useful coordinating calculation which has effectively been utilized to coordinate the remote detecting pictures naturally. In any case, highlight extraction administrator is as yet a noteworthy shortcoming of the calculation. Regardless of the high computational intricacy, it tends to be mostly matched features from related features. To ease the computational features and to enhance detection rate this investigation constrains the element to explore services in MSIFT and additionally key-directs extraction administrators toward a specific number to build up a productive method to coordinate pictures. The primary element of this technique is taking the upsides of a particular methodology to choose the best key focuses. Neighbour pixel features extracted as image descriptors and compared with Euclidean distance between image descriptors to use numerous one-to-one features in a trained database. At that point, utilizing slightest squares display, the exactness of the underlying coordinating sets is checked to locate the good coordinating. At long last, the two pictures coordinated by deciding the parameters of a section change work. Cognitive outcomes acquired by applying the strategy on various face pictures show the incredible productivity of the proposed technique in examination with the standard MSIFT calculation.

In the proposed technique, a steady component viewed as contrasted and another element for which the separation is not as much as a specific conclusion (in this undertaking 0.79) of separation to the following closest element. This technique diminishes the number of erroneous matches. For mistaken matches, there are various other close highlights because of the high-dimensional real highlights. Then again, if there is a right match, it is probably not going to discover different highlights because of the particular idea of the MSIFT descriptor.

The main key point discussed in the initial step, an image with different scale of image pixels. Structure of image with different distinctive pixel values filtered with different pixel features using Defereence of Gaussian architecture, which has shown in figure 1 represented in layers. Gaussian filter approach for a particular image with different image regions L(x, y, σ) is made by convolving the picture I(x, y) with Gaussian part G(x, y, σ) through a condition (1):

\[ L(x, y, σ) = G(x, y, σ) ⊗ I(x, y) \]

Where \( ⊗ \) means the convolution administrator. In Eq. 1, \( σ \) speaks to the scale with the underlying estimation of 1.6 rehashed progressively utilizing K at various levels of the octaves. Scale variant factors based on the above equations as follows:

\[ SC_{ol} = σ \cdot 2^{(0-1)i \cdot \frac{1}{O}} = σ \cdot K^{L(N(0-1)i)} \]

\( l=1,2,…..LN; k = 2 \cdot \frac{LN}{k} \)

Where SCol is the range aspect for stage 1 of octaveo. Finally, the Euclidean distance metric used in two different images.

Now scale space representation for the opted filter size this provides the smoothness portion of the selected rectangular regions with even weights.

**Figure 1. Gaussian and DOG filtered images with different features.**

In the proposed technique, a steady component viewed as contrasted and another element for which the separation is not as much as a specific conclusion (in this undertaking 0.79) of separation to the following closest element. This technique diminishes the number of erroneous matches. For mistaken matches, there are various other close highlights because of the high-dimensional real highlights. Then again, if there is a right match, it is probably not going to discover different highlights because of the particular idea of the MSIFT descriptor.

The main key point discussed in the initial step, an image with different scale of image pixels. Structure of image with different distinctive pixel values filtered with different pixel features using Defereence of Gaussian architecture, which has shown in figure 1 represented in layers. Gaussian filter approach for a particular image with different image regions L(x, y, σ) is made by convolving the picture I(x, y) with Gaussian part G(x, y, σ) through a condition (1):

\[ L(x, y, σ) = G(x, y, σ) ⊗ I(x, y) \]

Where \( ⊗ \) means the convolution administrator. In Eq. 1, \( σ \) speaks to the scale with the underlying estimation of 1.6 rehashed progressively utilizing K at various levels of the octaves. Scale variant factors based on the above equations as follows:

\[ SC_{ol} = σ \cdot 2^{(0-1)i \cdot \frac{1}{O}} = σ \cdot K^{L(N(0-1)i)} \]

\( l=1,2,…..LN; k = 2 \cdot \frac{LN}{k} \)

Where SCol is the range aspect for stage 1 of octaveo. Finally, the Euclidean distance metric used in two different images.

Now scale space representation for the opted filter size this provides the smoothness portion of the selected rectangular regions with even weights.

**Figure 2: filters for successive scales**

**INPUT:** Considering the processed face images and other sample images, I(x, y).

**STEP 1:** Given relational data D by fixing the centres c, 2 ≤ c < n, m > 1. Initializing β = 0, U(0) ∈ M_{Ecn}

**STEP 2:** Calculating Centroid mean Vectors

\[ v_i = (U_{1n}^m, U_{2n}^m, U_{3n}^m, ..., U_{m_n}^m)/(U_{1n}^m + U_{2n}^m + U_{3n}^m + ... + U_{m_n}^m) \]

**STEP 3:** Calculate

\[ d_{ik} = (D_p v_i) - \frac{(v_i^T D_p v_i)}{2}, \text{for } 1 ≤ i ≤ c \text{ and } 1 ≤ k ≤ n \]

if \( d_{ik} ≤ 0 \) for any i and k

**calculate the change**

\[ \Delta β = \max \left\{ -2 \cdot \frac{d_{ik}}{\|v_i - e_k\|^2} \right\} \]

\[ d_{ik} ← d_{ik} + \frac{\Delta β}{2} \|v_i - e_k\|^2 \]

**update** \( β ← β + \Delta β \)

**STEP 4:** Update \( U^r to U = U^{r+1} \)

**STEP 5:** Normalizing updated matrix

\[ \|U^{r+1} - U^r\| ≤ ε \text{ then} \]

**Retrieval Number:** B3782078219/19©BEIESP

**DOI:** 10.35940/ijrte.B3782.098319

Published By: Blue Eyes Intelligence Engineering & Sciences Publication
stop normalizing else set \( r = r + 1 \) and got step 2.

** OUTPUT: Updated features of SIFT acquired.**

### III. PROPOSED APPROACH

In this section, we define detail description of Human detection and Face based gender Recognition System (HDFGR); to explore the classification of gender [11][12][13][29] in facial images. Generally, classification of gender in trained machine learning settings which require exploration of different features from human facial images, classifier train features and the classify the features obtained from faces. The general methodology took in general orientation characterization illustrated in figure 2. First, pre-preparing undertakings on the data go up against picture. These assignments go up against institutionalization. Picture compels institutionalization and balancing of histogram. Here face institutionalization, step is human faces are cut as well as balanced with the objective that pieces of the face (for example eyebrows, ears, eyes, mouth, nose etc) reduce into predefined zones. Then picture drive institutionalization and histogram evening out are completed to speak to changing lighting conditions. Finally using MSIFT for a course of action of reason vectors in a lower-dimensional space is settled and the face pictures foreseen onto the subspace crossed by those reason vectors. These reason vectors encode the biased features of a human face. A readiness set is then encircled by taking a game plan of named faces and evacuating the features using the above technique. By then, a classifier is set up with the named data and rundown of capacities consolidate. For an inquiry picture, the features are isolated also. The classifier uses these features to choose the gender of the image given from the datasets of PRIVATE FGNET DATA SET, FGNET DATASET and from local college real time database.

**INPUT:** Considering the processed face images and other sample images, \((x, y)\).

**STEP 1:** Convert RGB into grey related image \((G(x, y))\).

**STEP 2:** MSIFT with multi feature extraction on IMAGE \((G(x, y))\).

**STEP 3:** RESIZING features values into a column image with extracted MMSIFT features

**STEP 4:** Features gets Multi-class SVM [2][3][23] classified by using a Gaussian kernel for distance measure between test and trained samples.

**STEP 5:** FACE recognition and other samples classified outcome

**Algorithm 1. Procedure to extraction features and objects classification for different facial images.**

Initially, a set of rectangular components extracted as a region of interest formed by the origin and the location of origin represented by the summation of components regions and is given by:

\[
H(x, y) = \begin{bmatrix}
L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\
L_{xy}(x, \sigma) & L_{yy}(x, \sigma)
\end{bmatrix}
\]

This process gathers information from Rectangular region of interest and approved by the summation of locations in later stages. Then the representation of the image with different sizes as follows

\[
D_{xx}D_{yy} - (wD_{xy})^2
\]

**A. Extraction of Feature**

As was depicted before feature extraction is done by envisioning the face picture onto a cut down dimensional subspace. Dimensionality decline systems MSIFT is used for this philosophy. One of the objectives of this endeavour was to see how the two methodologies perform for sexual introduction gathering. The motivation driving doing dimensionality decline is to work with the most accommodating pieces of an image. This approach improves both the estimation time and execution of the technique used. In going with, the two-dimensionality decline procedures examined. Procedure for representation objects for classification shown in algorithm 1.

**Figure 3. Procedural process for object recognition and sampling objects**

![Diagram](image.png)

**Retrieval Number:** B3782078219/19©BEIESP
**DOI:** 10.35940/ijrte.B3782.098319

**Published By:** Blue Eyes Intelligence Engineering & Sciences Publication

**ISSN:** 2277-3878, Volume-8 Issue-3, September 2019
Where weight factor is given by
\[ w = \frac{[L_{\frac{1}{2}}(1.2)][D_{\frac{1}{2}}(9)]}{[D_{\frac{1}{2}}(9)]} = 0.912 \ldots \sim 0.9, \]
the frobenius norm on the weight samples.

Now scale space representation for the opted filter size provides the smoothness portion of the selected rectangular regions with even weights. Then these points were localized and operated for extracting hard features to define the object.

B. MSIFT

Scale Invariant Feature Transform is a well-known approach to decrease the image data dimensionality, for apply MSIFT to facial images, the first image represents in vector presentation with several columns. For implementing the principal component to pictures, an image shows vector representation of different columns, and it is created to store the matrix. Let X be the matrix
\[ X = [x_1, x_2, \ldots, x_n], \]
where \( x_i \) is the \( i \)-th line vector comprising the \( i \)-th coaching picture.

Then variance and mean for the particular matrix is
\[ \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \]
\[ Y = [x_1 - \bar{x}, \ldots, x_n - \bar{x}] \]

Then covariance matrix \( Q = \text{cov}(Y) = YY^T \)

In conclusion, Eigen [18] esteem separating is finished to find the most significant position (given Eigen-values) Eigenvectors. These vectors, known as real components, period the low point of view subspace. Out of these Eigenvectors m most critical vectors are chosen, left these alone be – e1, e2, …, em. The estimation of m chosen by considering the aggregate total of the Eigen esteems. The familiar elements of an image x then determined by foreseeing it onto the spot crossed by the Eigen-vectors as pursues:
\[ g = [e_1, e_2, \ldots, e_m]^T (x - \bar{x}) \]
Where \( g \) is an m perspective vector of functions. This function vector \( g \) used during coaching and category. 

C. SVM Classification

Support vector devices are classifiers that build a maximum splitting hyperplane between two sessions so that the category mistake reduced. For linearly non-separable data, the feedback planned to the high-dimensional function area where a hyper-plane can separate them. This projector screen into a high-dimensional function area efficiently conducted by using popcorn kernels. For an instance-label couple (i) i x y with n i x \( \in \mathbb{R}, \epsilon \in [-1,1] \) i y for 1 \( \leq i \leq n \) where n is the variety of circumstances, the following optimization problem needs to fixed for SVMs –

\[ \min_{w,b,\eta} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \eta_i \]
Subject to \( y_i = (w^T \phi(x_i) + b) \geq 1 - \eta_i, \eta_i \geq 0 \)

Where C is the constant parameter for image data and \( \phi \) charts, a training instance \( x_i \) to enhance the perspective area. The kernel \( K \) determined as
\[ K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \]
\[ K(x_i, x_j) = \exp(-\gamma ||x_i - x_j||^2) \]

The parameter \( \gamma \) manages the distribution of a Gaussian group — the above two sequences classified as known and unknown classifications.

Distance Metric based SVM

Let’s be precise. "Distance" has lots of meanings in data science, i.e. it is Euclidean distance. The Gaussian kernel is a non-linear function of Euclidean distance.
\[ K(x, x') = \exp(-\frac{||x - x'||^2}{2\sigma^2}) \]

The kernel function decreases with distance and ranges between zero and one. In Euclidean distance, the value increases with distance. Thus, the Gaussian kernel function is more useful metrics for weighting observations

IV. EXPERIMENTAL PERFORMANCE EVALUATION

Analyses were completed with the fluctuating size of the preparation dataset, force standardization, histogram evening out, scaling preparing plus test for testing dataset. Picture database [9] as of utilized hence. The database includes complete frontal facial images of 100 images individual (50 feminine and 50 folks). Those Photos cut by emptying hair and establishment. Two getting ready sets were of sizes 40 (20 females and 20 folks) and 60 (30 feminine and 30 folks). Whatever left of the pictures was utilized for testing for same varieties made on MSIFT way to deal with watching the impacts of various decisions.

In the preprocessing stage, the image resizes with 48x48 pixel representations. Histogram evening out is at that point performed (if relevant) on these pictures. Based on the above images, we have to apply our proposed approach on different facial images. Images in figure 3 represent different pixel notations.

Figure 4. Example facial data sets with different dimensions.

Randomly taken some of the images from facial database training some of the human images from selected facial images [27][28].
First, we find the sum of the Eigenvalues using pixel notation present in the facial image. Following figure 4 shows the sum of the cumulative performance with different Eigenvalues in each facial image.

Table 1 shows the performance results of the proposed approach in terms of time with different facial images.

<table>
<thead>
<tr>
<th>Time for different datasets</th>
<th>Proposed Approach</th>
<th>Naïve Bayesian Classifier</th>
<th>Random Forest Classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>FERET</td>
<td>1.0</td>
<td>1.3</td>
<td>1.7</td>
</tr>
<tr>
<td>LFW</td>
<td>1.1</td>
<td>2.2</td>
<td>1.6</td>
</tr>
<tr>
<td>ORL data set</td>
<td>1.2</td>
<td>3</td>
<td>1.6</td>
</tr>
<tr>
<td>FG Net Data Set</td>
<td>0.8</td>
<td>1.5</td>
<td>1.2</td>
</tr>
<tr>
<td>Private FG Net dataset</td>
<td>1.5</td>
<td>2</td>
<td>1.8</td>
</tr>
</tbody>
</table>

Figure 5. Time comparison results with different data sets.

Figure 6 and table 2 follows accuracy results with different data sets with different facial images.

Table 2. Accuracy values for different facial data sets

<table>
<thead>
<tr>
<th>Classification Accuracy</th>
<th>Naïve Bayesian Classifier</th>
<th>Random Forest Classifier</th>
<th>Proposed Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>FERET</td>
<td>78.30</td>
<td>83.6</td>
<td>87.4</td>
</tr>
<tr>
<td>LFW</td>
<td>78.49</td>
<td>87.2</td>
<td>90.6</td>
</tr>
<tr>
<td>ORL data set</td>
<td>81.5</td>
<td>78.6</td>
<td>88.5</td>
</tr>
<tr>
<td>FG Net Data Set</td>
<td>75.39</td>
<td>83.9</td>
<td>89.3</td>
</tr>
<tr>
<td>Private FG Net dataset</td>
<td>78.6</td>
<td>83.6</td>
<td>94.5</td>
</tr>
</tbody>
</table>

Figure 6. Performance of accuracy with facial data sets. Based on total discussion table 1and table 2, output results show better and efficient results using MSIFT and SVM[23][26]. Accuracy of the proposed approach is higher than 90%, i.e. 98.7 %.

Table 1: Performance Metrics Comparison

<table>
<thead>
<tr>
<th>PARAMETERS</th>
<th>Naïve Bayesian Classifier</th>
<th>Random Forest Classifier</th>
<th>Proposed Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>1.56</td>
<td>1.78</td>
<td>0.86</td>
</tr>
<tr>
<td>SSE</td>
<td>0.94</td>
<td>0.56</td>
<td>0.015</td>
</tr>
<tr>
<td>RMSE</td>
<td>1.24</td>
<td>1.334</td>
<td>0.927</td>
</tr>
<tr>
<td>PRECISSION</td>
<td>0.803</td>
<td>0.95</td>
<td>0.967</td>
</tr>
<tr>
<td>RECALL</td>
<td>0.0965</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>F MEASURE</td>
<td>0.93</td>
<td>0.945</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Error rate minimization was happened when compared with other classifiers using MSIFT feature turning process. Our approach is performing 14% better when compared with other classifiers too.

V. CONCLUSION

We present Human detection and Face based gender Recognition System which consists of both MSIFT and Distance-based SVM for the reduction of dimensionality of facial attributes for classification of gender for different facial data sets. Classification related approach is also taken to prove assumptions that given input images align with background clutter image data. Our proposed approach extracts features after performing dimensionality based reduction using SVM with classification using Gaussian filter model with different sequences. Results of proposed approach varying facial data, image intensity with time, classification accuracy and Eigen values classification of the image also presented. Based on the above results, our proposed approach gives better and efficient accuracy with time performed on different data sets. Further improvement of the proposed approach is to support different object label classification with a clustering approach on different facial image data sets.
REFERENCES


2. Chandrakamal sinha, “Gender Classification from Facial Images using PCA and SVM”, Department of Biotechnology and Medical Engineering National Institute of Technology Rourkela Rourkela – 769 008, India, proceedings in Ethics 2013.


AUTHORS PROFILE

G. D. K. Kishore, is a Research Scholar in Krishna University and working as an Assistant Professor in SRKIT, Vijayawada. His research interests in Data Mining, Machine Learning and Image Processing. He had life time membership in CSI.

Mukkamalla. Babu Reddy is currently working as an Assistant Professor & HOD (i/c) in the Department of Computer Science and also acts as Officer In-Charge (Exams) in Krishna University, Machilipatnam, Krishna Dt. Andhra Pradesh. He received Ph.D in Computer Science & Engineering from Acharya Nagarjuna University, Andhra Pradesh, India, which was awarded in March 2010. He has a strong publication record to his credit, and he is also the reviewer for good flagship Conferences and Reputed Journals. His research is broadly in the field of Data Mining, Machine Learning and Neural Networks.