Gender Recognition Challenge

Manasi Dubey	Akshat Mehta	Mahima Modi	Mehul Prajapati
201501051	201501053	201501060	201501079

***Abstract*— Gender recognition is a method of distinguishing genders of human beings based on different aspects of human traits. In our case, we are classifying genders based on facial features. In gender recognition of images given to us, clustering and classification is the most important task. In both gender clustering and classification, one the most vital processes carried out is the selection of features. For that purpose, we have first found the facial features using SURF. Then by using K-means, images are clustered based on similar features. In the end using SVM (Support Vector Machine), learning process is done using which the genders are labelled as male or female.**

***Keywords*— Gender recognition, clustering and classification, SURF(Speeded Up Robust Feature), K-means, Support Vector Machine.**

I. Introduction To Face Recognition

Gender recognition has recently become an area of research in the domain of computer vision that provides many interesting application areas. Among the potential application areas for gender recognition are human computer interaction, person categorization, improving biometric recognition. Even in the commercial world, there are many applications where gender recognition is useful. An instance is when a retailer wants to evaluate the percentage of a particular gender that stopped by to view the displayed products. All these potential applications of gender recognition have led to the necessity of finding practical solutions to the problem of gender recognition for electronic systems. For humans, recognizing gender is an easy task, under any kind of constrain. But, for electronic systems, performing accurate gender recognition is a daunting task, especially when facial features cannot be well extracted as a result of several challenges such as low-quality image and improperly aligned images.

Till date, the work has been emphasized on gender recognition through visual observation, but now, it has to be emphasized to computer, to perform this task. It is observable that our behavior and social interaction are greatly influenced by genders of people whom we intend to interact with. Hence a successful gender recognition system could have great impact in improving human computer interaction systems in such a way as to make them be more user-friendly and acting more human-like. Over the past decades, there have been significant advances in facial image processing, especially, in a face detection area where a number of fast and robust
algorithms have been proposed for practical applications. As a result, a number of research areas attempting to extend the works have been emerging, face recognition, facial expression recognition and gender recognition.
Based on the types of features used, facial feature extraction approaches can be roughly divided into two different categories: geometric feature-based methods and appearance-
based methods. Geometric features refer to distance between various facial features such as eyes, nose, chin and lips. Facial features can be extracted from facial image using Viola Jones algorithm that returns the coordinates of various features.

Machine Learning is a field of computer science that evolved from field of pattern recognition and computational learning theory in Artificial Intelligence Machine Learning explores the study of algorithms that can learn and make predictions from available raw data. This article has exploited the capabilities of machine learning.

II. Literature Review

A. Revisiting Linear Discriminant Techniques

It discusses why linear techniques are not achieving competitive results and shows how to obtain state-of-the-art performances. Their work confirms previous results reporting very close classification accuracies for Support Vector Machines (SVMs) and boosting algorithms on single-database experiments. They have proven that Linear Discriminant Analysis on a linearly selected set of features also achieves similar accuracies. They perform cross-database experiments and prove that single database experiments were optimistically biased.

B. Gender Recognition System from Facial Images using SURF based BoW Method

This paper is principally aimed to use SURF based BoW and SVM method pair for frontal, left and right pose of face image-based gender recognition system unlike previous work. Firstly, face regions are detected from well-known FERET face database. Secondly, the features are extracted using SURF based BoW and SIFT based BoW algorithms from various face images and finally extracted features are classified with the SVM method.

C. Deep Convolutional Neural Networks and Support Vector Machines for Gender Recognition

They explored the applicability of deep convolutional neural networks on face gender recognition. We showed that despite the challenging nature of the problem, state-of-the-art classification rates can be achieved using relatively short training times. On both datasets, the best results were obtained when using the fine-tuned networks. Oversampling by averaging class scores of the final classifiers was shown to improve classification rates in all cases.

D. Gender Recognition using PCA and LDA with Improve Preprocessing and Classification Technique

This paper explains the gender recognition system through a human facial image by using the basic method of Principal Component Analysis (PCA) combined with Linear Discriminant Analysis (LDA). PCA+LDA method performance can be improved by improvising the preprocessing techniques such as resizing the image, equalizing the histogram, and removing the variation of the image background by adding oval masking face. Furthermore, in classification process, using 9 nearest neighbors gives the better recognition accuracy rather than using only 1 nearest neighbor. The highest accuracy results obtained with the proposed method is superior to get 89.70% when compared to the PCA + LDA method without adding masking face, which only reached 84.16%.

E. Bag of Words Based Surveillance System Using Support Vector Machines

They have proposed an automated surveillance system for detecting fire weapons in cluttered scene. First SIFT features are extracted from the collection of images. Second, K-means clustering is adopted for clustering the SIFT features. Third, a word vocabulary based histogram is implemented by counting occurrences of the extracted clusters in each image. The histogram is the input to Support Vector Machine that will be trained on the collection of images. Finally, the trained SVM is the system classifier that will decide if new image contains a weapon or not. The main contributions of the paper is to adopt the visual words classification scheme in detecting fire weapons. In addition, we used RANSAC to reduce the matching outliers. The system showed high accuracy in detecting fire weapons in images and video surveillance systems

III. Our Approach

- We detected faces using Viola Jones.
- Extract the SURF local feature vectors from the set of the training image. SURF features are used in this approach because of their strength in classification and immunity to illumination, rotation and scaling.

  SURF works in the following way –

  I. Find features/ key-points that are likely to be found in different images of same object again.

  II. Find the right "orientation" of that point so that if the image is rotated according to that orientation, both images are aligned in regard to that single key-point.

  III. Computation of a "descriptor" that has information of how the neighbourhood of the key-point looks like (after orientation) in the right scale.

  Now the Euclidean distance computation is done only on the descriptors, not on the key point locations!

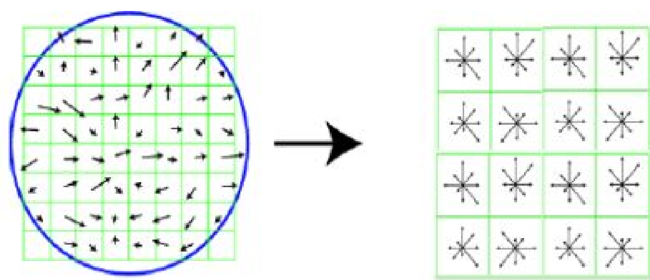


Figure 1. Key Point Descriptor, (left) the Gradient Map (right) with a 4x4 Location Grid and 8 Orientations (128 Dimensions)

- Put all extracted local feature vectors into a single set (You do not need to know for which image the local feature is obtained).

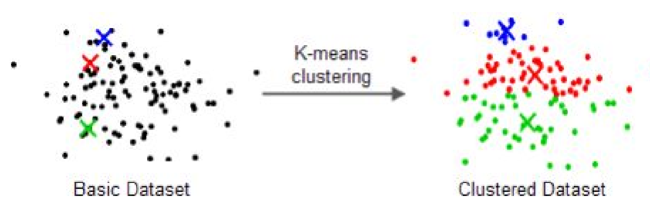


Figure 2. K-means Clustering. (left) the Dataset without Clustering, Cross Marks show the Means (centers) that are Selected in the First Iteration, (right) Data after Clustering and Assigning each Value to the Nearest Center

- Apply a clustering algorithm (the k-means algorithm is used for its efficiency) over the set of local feature vectors in order to obtain a set of clusters; each cluster represents a single visual word.

- Obtain the spatial histogram (spatial histogram is the global feature vector that counts how many times each visual word occurred in each image).
- Pass the obtained histogram features for each image to the SVM for training. Classify new images by using the trained SVM

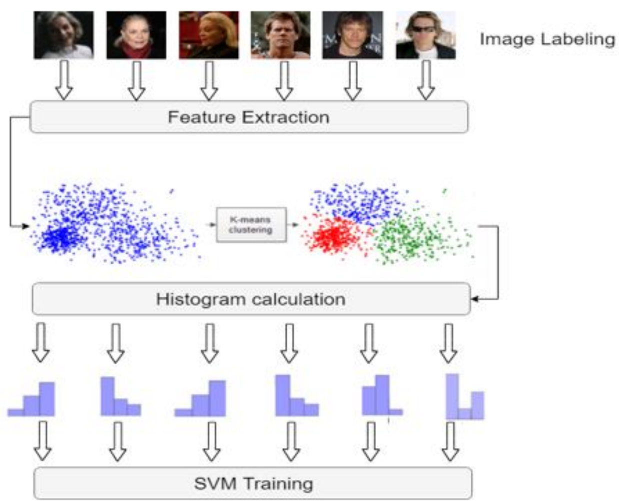


Figure 3. The Proposed System Architecture, The Training Phase of BoWSS, an Example of 3 Visual Words

IV. Implementation And Results

Results obtained real time image –

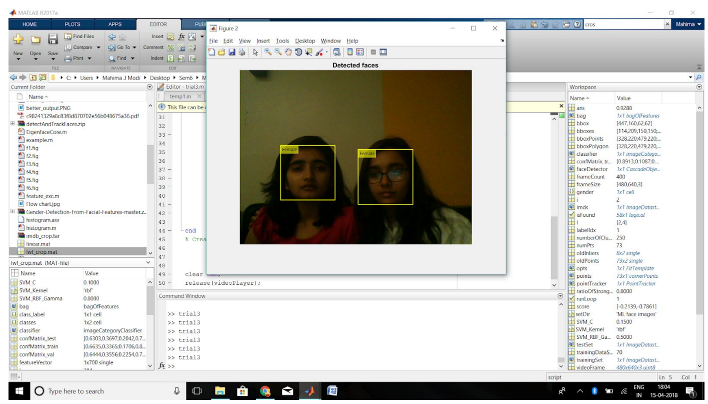
Following is for default values -

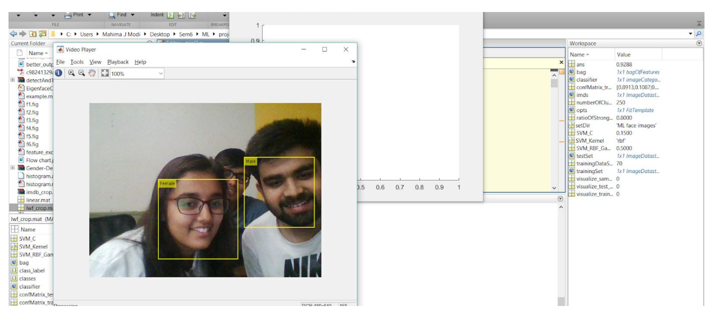
1. Kernel - RBF
2. C = 0.15, gamma = 0.6, K= 500(No. of clusters),
   Training set = 60%, Test set = 20%, Validation set = 20%
   Interest points = 80%

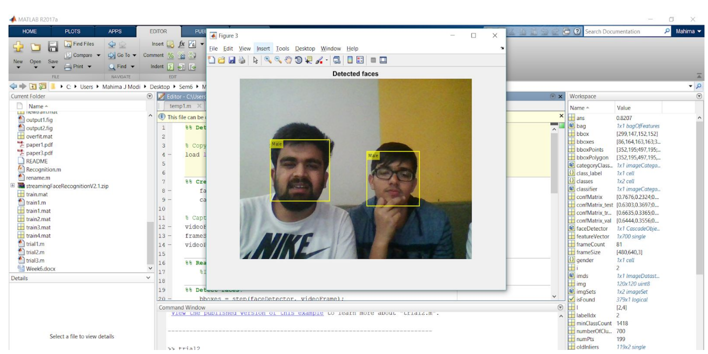
DATASET	TOTAL IMAGES	MALE IMAGES	FEMALE IMAGES
IMBD	4788	2394	2394
CLASS	548	396	152
LFW	2926	1463	1463

DATASET	VALIDATION ACCURACY			TEST ACCURACY		
	Average accuracy (%)	Accuracy in male detection(%)	Accuracy in female detection(%)	Average accuracy(%)	Accuracy in male detection(%)	Accuracy in female detection(%)
IMDB	68	77	57	67	74	60
Class	71	73	80	68	63	73
Labelled Faces in the Wild-(funneled)	76	70	69	72	69	66
LFW_cropped	82	86	80	80	83	79

Result for real time video surveillance-







We have used linear SVM with k = 250, interest point = 80 and on average the accuracy is 93%.

V. Conclusion

In order to detect the gender from face images efficiently, we have used a surveillance technique called Bag of Words Surveillance System, and we have managed to apply properly visual words to help detect gender. As a result, the algorithm presented accuracy of about 70%. We are using SURF instead of SIFT. because SURF extracts interest point faster and accurate than SIFT. IMDB dataset contains mostly facial images of European people. So, while evaluating for Asian people, the accuracy decreases. While working with the dataset of college students the accuracy increases because the data was cleaned.

VI. Future Plan

In the future work, we are planning to use the same algorithm for real-time surveillance. However, this leads us to explore more efficient, accurate and faster classification especially in case of clustering and histogram calculations.

VII. Acknowledgement

We are very thankful to Prof. Mehul Raval for giving us the opportunity to explore the field of machine learning to such a depth. Mr. Aditya Parikh was also very helpful whenever we faced any difficulty. By doing this project, we did nothing but increased our depth of knowledge. It was really an experience.

VIII. References

1. Ferizal, Reza, Sunu Wibirama, and Noor Akhmad Setiawan. "Gender recognition using PCA and LDA with improve preprocessing and classification technique." In *Engineering Seminar (InAES), 2017 7th International Annual*, pp. 1-6. IEEE, 2017.
2. Hatipoglu, B. and Kose, C., 2017, October. A gender recognition system from facial images using SURF based BoW method. In *Computer Science and Engineering (UBMK), 2017 International Conference on* (pp. 989-993). IEEE.
3. Halima, Nadhir Ben, and Osama Hosam. "Bag of words based surveillance system using support vector machines." *International Journal of Security and Its Applications* 10.4 (2016): 331-346.
4. van de Wolfshaar, Jos, Mahir F. Karaaba, and Marco A. Wiering. "Deep convolutional neural networks and support vector machines for gender recognition." In *Computational Intelligence, 2015 IEEE Symposium Series on*, pp. 188-195. IEEE, 2015.
5. Bekios-Calfa, J., Buenaposada, J. M., & Baumela, L. (2011). Revisiting linear discriminant techniques in gender recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(4), 858-864.