# Fast and Efficient Random Forest Technique for Face Verification

Amina K, Lekshmy P L

**Abstract**— Face verification systems aim to find out whether two face images belong to the same subject. The existing system uses a generative model such as Joint Bayesian method, it formulate verification task as a binary Bayesian decision problem. This method treats the samples as random variables respecting certain data-generating models, and treats the subjects of the samples as latent variables. However in Joint Bayesian method, accuracy is low due to the inaccurate feature selection. And the classifier does not support for highly non linear data. In this paper uses a random forests technique to deal with the above challenges in the face recognition. The proposed algorithm first extracts features from the face images using SIFT, LBP, Gabor and HoG. And uses a Non-negative Matrix Factorization (NMF) method for appropriate feature selection. And then uses the random forest algorithm to classify the images based on the selected features. Non-negative matrix factorization is a linear, nonnegative approximate data representation. This can be used to reduce the feature size. Random forest can also support for highly nonlinear data. And it has a good convergence property and achieves a higher verification rate than both the Joint Bayesian method and other state-of-the-art classifiers on the labeled faces in the wild face database.

**Index Terms**— AJB (Advanced Joint Bayesian), Gabor, HoG, LBP (Linear Binary Pattern), NMF (Non-negative Matrix factorization), Random forest, SIFT (Scale Invariant Feature Transform).

---

## 1 Introduction

FACE recognition is a rapidly growing filed today for many applications of biometric authentication, security, ect. The security of information is becoming very significant and difficult. Security cameras are presently common in airports, Offices, University, ATM, Bank, and in any locations with a security system. Face recognition is a biometric system used to identify or verify a person from a digital image.

The critical step in face recognition is face verification. Face verification system aims to find out whether two face images belong to the same subject. A typical face verification system consists of several stages: image preprocessing, feature extraction, feature selection and classification.

Recently a generative model such as Advanced Joint Bayesian (AJB) classifier [1] is used for face verification task. Three recent pieces of work [2],[3],[4] use this method to design the classifiers of their systems. DeepFace [5], an innovative algorithm, which uses a deep learning structure and reduces the verification error and also states that the Joint Bayesian method is the “currently most successful system”. However, the AJB method uses EM-like algorithm to estimate the model parameters. In this algorithm, accuracy is low due to the inaccurate feature selection. And less efficient for highly nonlinear data.This work attempts to use random forests to deal with the above challenges in the face recognition. The proposed algorithm first extracts features from the face images using SIFT [6], LBP [7], Gabor [8] and HoG [9]. And uses a Non-negative Matrix Factorization (NMF) method [10], [11] for appropriate feature selection. This can be used to reduce the feature size. And then uses the random forest algorithm to classify the images based on the selected features. Random forests are proven to be fast [12],[13],[14],[15] which is suitable with our motivation. Further, a random forest is a powerful statistical framework [16] with very high generalization accuracy.

## 2 Literature Survey

In 2004, Raphael mare, pierre geurts, Justus piater, and Louis wehenkel [12] evaluate a generic machine learning algorithm based on decision tree ensembles. And then introduce an extension of this algorithm that augments its generality even further. Both forms of the algorithm operate directly on the pixel values and do not extract any task specific features. To demonstrate the performance of the algorithms, [12] have chosen four typical problems in image classification. And in all four cases, the extended algorithm produces results competitive with the state-of-the-art.

In 2005, Raphael mare, Pierre geurts, Justus piater, and Louis wehenkel [13] compare five tree-based machine learning methods within the generic image classification framework. It is based on random extraction of sub windows and their classification by decision trees. It shows that this general and conceptually simple framework yields good results for object recognition task combined with ensembles of decision trees.

In 2006, Pierre Geurts, Damien Ernst, Louis wehenkel [14] propose a new tree induction algorithm that selects splits, both attribute and cut-points, totally or partially at random. The high variance of decision and regression tree splits suggests investigating whether higher randomization level could improve accuracy with respect to existing ensemble methods.

In 2007, Frank moosmann, Bill trigs and Frederic jurie [15] contributes two main ideas. One is that ensembles of trees eliminate many of the disadvantages of single tree based coders without losing the speed advantage of trees. The second is that classification trees contain a lot of valuable information about locality in descriptor space that is not apparent in the final class labels. One can exploit this by training them for classification then ignoring the class labels and using them as clustering trees. Comparing these two ideas introducing an Extremely Randomized Clustering Forests (ERC-forests) [15]. Which are the ensembles of randomly created trees. And they provide much faster training and testing and more accurate results than conventional k-means in several state-of-the-art image classification tasks.

In 2010, Dingding Wang Tao Li [10] proposed an unsupervised approach that combines keyword selection and document clustering (topic discovery) together. The proposed approach extends non-negative matrix factorization (NMF) by incorporating a weight matrix to indicate the importance of the keywords. The proposed approach is further extended to a weighted version in which each document is also assigned a weight to assess its importance in the cluster.

## 3 Methods And Techniques

### 3.1 Pre-processing

The image is first processed in order to extract the features, which describe its content. The pre-processing involves filtering, normalization, segmentation, and object identification. [17] The output of this stage is a set of significant regions and objects. This can enhance the visual appearance of image.

### 3.2 Feature Extraction

Feature extraction is related to dimensionality reduction. Here we use multiple feature extraction techniques including SIFT [6], LBP [7], Gabor [8], and HoG [9]. And divide the image into non-overlapping blocks and then extract uniform patterns.

### 3.3 Feature Selection

Non-whitening PCA [18] is used to reduce the feature dimension. It is a statistical procedure that uses an orthogonal transformation to convert a set of observations into a set of values of linearly uncorrelated variables.

### 3.4 Advanced Joint Bayesian Method

Advanced joint Bayesian (AJB) is a generative classifier method. This method treat the samples as the random variables respecting certain data generating models, and treat the subjects of samples as latent variable [1]

However, the joint Bayesian method uses EM-like algorithm to estimate the model parameters, which deviates from standard EM algorithm. [1] This algorithmic deviation will cause the parameter estimator of AJB to have some undesired properties. And it is less efficient for highly nonlinear data.

## 4 Proposed Method

To deal with the above challenges in the face recognition, a fast and efficient face recognition system has been needed. To develop an efficient application, Random forest method is used in classification step. Extending the literature above to develop random forest for face recognition presents an interesting research goal. In the second part, random forest are used to classify images represented using non negative matrix factorization (NMF) [10],[11].

### 4.1 Random Forest

Random forest are defined as a combination of tree predictors such that each trees depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest [19]. Random forest is powerful statistical framework [19] with high generalization accuracy.

Random trees are grown recursively top down. And it divides the training data on completely random attributes. This randomization mainly addresses high dimensional data and generalization [20],[21],[19].

The random forest algorithm can be summarized as follows: [22]

1. Draw $n_{tree}$ bootstrap samples from the original data.
2. For each of the bootstrap sample, grow an unpuned classification tree, with the following modification: at each node, rather than choosing the best split among all predictors, randomly sample $m_{trv}$ of the predictors and choose the best split from among those variables.
3. Predict new data by aggregating the predictions of the $n_{tree}$ trees ( i.e, majority votes for classification).

The underlying nature of random forests is that of building classifiers independently. This makes their construction inherently parallel. Its result as, referred to as a parallel learning algorithm [23]. Random forest creates a large number of independently trained classifiers by the random sampling of features.

### 4.2 Non-Negative Matrix Factorization (NMF)

Non-negative matrix factorization (NMF) [11] is relatively a new way of reducing the dimensionality of data into a linear combination of bases. NMF has non-negative constraints, so it can be used to represent data with non-negative features quite well. Non-negative matrix factorization (NMF) also non-negative matrix approximation is a group of algorithms in multivariate analysis and linear algebra where a matrix V is factorized into (usually) two matrices W and H, with the property that all three matrices have no negative elements. This non-negativity makes the resulting matrices easier to inspect [10]

Let's assume that our data consists of T measurements of N nonnegative scalar variables. Denoting the (N-dimensional) measurement vectors $v^t$ (t =1, . . .,T), a linear approximation of the data is given By

$$V^t \approx \sum W_i h_i^t = Wh^t$$

Where W is an N×M matrix containing the basis vectors $w_i$ as its columns. Note that each measurement vector is written in terms of the same basis vectors. The M basis vectors $w_i$ can be thought of as the 'building blocks' of the data, and the (M-dimensional) coefficient vector ht describes how strongly each building block is present in the measurement vector $v^t$.

Arranging the measurement vectors $v^t$ into the columns of an N×T matrix V, we can now write,

$$V \approx WH$$

Where each column of H contains the coefficient vector $h_t$ corresponding to the measurement vector $v^t$. Written in this form, it becomes apparent that a linear data representation is simply a factorization of the data matrix. Principal component analysis [18], and non-negative matrix factorization (NMF) [11] can all be seen as matrix factorization, with different choices of objective function and/or constraints.

$$E(W,H) = \|V - WH\|^2 = \sum (V_{ij} - (WH)_{ij})^2.$$

NMF is similar to PCA in that it assigns weights to a set of bases to blend a representative observation, but the weights are constrained to be positive.

## 5 Result And Analysis

In this section, we compare our Random Forest with both the joint Bayesian method and other classifiers, using LFW face database for illustration.

### 5.1 Experimental Settings

The LFW database [24] is the database designed for face verification. It presents large variations in pose, expression and lighting. From the database, we consider 30 images to be tested. And the performance of the classifier is compared with the help of confusion matrix. We use multiple features to conduct our experiments, including SIFT [6], LBP [7], Gabor [8], and HoG [9]. From the extracted features, we select 100 features by NMF [11]. The accuracy of the classifier is obtained from adding true positive and true negative value from the confusion matrix. And these values are divided by the total number of testing inputs.

### 5.2 Comparison with the Joint Bayesian

The result is shown in TABLE 2. It can be seen clearly that the Random forest method improves verification accurary, compared with the joint Bayesian method in TABLE 1.

TABLE 1
Confusion Matrix Of AJB

| | | Predicted class | | |
|---|---|---|---|---|
| | | Class 0 | Class 1 | |
| Actual class | Class 0 | 16 (53.3%) | 2 (6.7%) | TP rate 88.9% |
| | Class 1 | 2 (6.7%) | 10 (33.3%) | FP rate 16.7% |
| | | Precision 88.9% | Error rate 13.3% | Accuracy 86.7% |

From the TABLE 1, AJB method results the accuracy with 86.7%. And the error rate is13.3%. AJB is a generative method, and it generates each value at its training time. So it is more time consuming.

TABLE 2
Confusion Matrix Of Random Forest

| | | Predicted class | | |
|---|---|---|---|---|
| | | Class 0 | Class 1 | |
| Actual class | Class 0 | 17 (56.7%) | 1 (3.3%) | TP rate 94.4% |
| | Class 1 | 0 (0.0%) | 12 (40.0%) | FP rate 0.0% |
| | | Precision 100% | Error rate 7.7% | Accuracy 96.7% |

The TABLE 2 shows the performance of random forest method. Here the accuracy is 96.7%, and the error rate is 7.7% only. So it gives more accurate result than AJB.

### 5.3 Comparison with the State of the Art

Finally, we combine our four Random forest classifier, which are SIFT [6], Gabor [7], LBP [8] and HoG [9] features. And the classifier takes the linear combination of these four scores for the verification. And the verification rate reaches to 0.967.

TABLE 3
COMPARISON OF ALGORITHMS ON THE LFW DATABASE

| METHODS | VERIFICATION RATE |
|---|---|
| PLDA | 0.901 |
| FISHER VECTOR FACE | 0.930 |
| AJB | 0.925 |
| RANDOM FOREST | 0.967 |

Table 3 shows the verification rate of state of the art classifiers. From the table, we clearly seen that random forest classifier has higher verification rate compared with the state of the art.

## 6 CONCLUSION

In this research has been done to the performance of a face recognition system by making the use of classifier with Random forests. The whole system mainly consists of four parts, namely preprocessing, feature extraction, feature selection, and classification. In this paper, we also propose a new method in feature selection step, NMF (Non-negative Matrix Factorization). This can be used to reduce the feature size. In the classification, the face image is compared with the images from the database. Random forest is one of the best classification technique available today and has been show to perform very well compared to other classifiers.

## ACKNOWLEDGMENT

We are greatly indebted to our principal Dr. JAYAMOHAN J ,Dr. V GOPAKUMAR., Professor, Head of the Department of Computer Science and Engineering, Mr. MANOJ KUMAR G., Associate Professor, Department of Computer Science and Engineering, LBS Institute of Technology for Women who has been instrumental in keeping my confidence level high and for being supportive in the successful completion of this paper.We would also extend our gratefulness to all the staff members in the Department,also thank all my friends and wellwishers who greatly helped me in my endeavor. Above all,we thank the Almighty God for the support, guidance and blessings bestowed on us, which made it a success.

• *Amina k, Mtech computer science and engineering, LBS Institute of Technology for women poojappura,Trivandrum,India, PH-9744135876. E-mail: aminak064@gmail.com*
• *Lekshmy P L,Assistant Professor, LBS Institute of Technology for women poojappura,Trivandrum,India, Country, PH-9497000337. E-mail: lekshmyvinod@gmail.com*

## REFERENCES

[1] Yicong Liang, Xiaoquing Ding, Fellow, IEEE, and Jing-Hao Xue, "Advanced Joint Bayesian Method for Face Verification", IEEE Transactions on Information Forensics and Secirity, Vol.10, No.2,February 2015.
[2] X. Cao, D. Wipf, F. Wen, G. Duan, and J. Sun, "A Practical Transfer Learning Algorithm for Face Verification", IEEE Int. Conf.Comput.Vis. (ICCV), Dec 2013, pp.3208-3215.
[3] D.Chen, X.Cao, F.Wen, and J.Sun," Blessing of Dimensionality: High Dimensional Feature and its Efficient Compression for Face Verification", IEEE conf.Comput.Vis>Pattern Recognit(CVPR), Jun 2013, pp.3025-3032.
[4] K.Simonyan, O.M.Parkhi,A.Vedaldi, and A.Zisserman, "Fisher Vector Faces in the Wild", in proc.Brit.Mach.Vis.Conf.(BMVC),2014,pp.7-18.
[5] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 1701–1708.
[6] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 7th IEEE Int. Conf. Comput. Vis. (ICCV)*, Sep. 1999, pp. 1150–1157.
[7] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 467–476, Apr. 2002.
[8] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.
[9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2005, pp. 886–893.
[10] Dingding Wang Tao Li, Chris Ding," Weighted Feature Subset Non-Negative Matrix Factorization and its Applications to Document Understanding", 2010 IEEE International Conference on Data Mining.
[11] Yi-Hung, Huang, Chun-Nan, Hsu," An efficient high dimensional supervised feature selection method based on non-negative matrix factorization".
[12] Raphal Maree, Pierre Geurts and Louis Wehenkel,"A Generic Approach for Image Classification based on decision tree ensemble", 6th asian conference on computer vision, 2004.
[13] Raphal Maree, Pierre Geurts and Louis Wehenkel," Decision trees and random subwindows for object recognition", ICML workshop on Machine learning Techniques for processing media,2005.
[14] Pierre Geurts, "Extremely randomized trees", In machine learning, page 2006.
[15] Frank Moosamann, Bill Triggs, and Freeric Jurie," Fast discriminative visual codebooks using randomized clustering forests", NIPS 2007.
[16] Leo Breimen and E. Schapire, Random forests. In Machine Learning, pages 5-32, 2001.
[17] Vincent Luboz,"Image Segmentation and Preprocessing", Department of surgery and cancer, Imperial college London.
[18] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 1991, pp. 586–591.

ISSN 2229-5518


[19] Leo Breiman and E.Schapire, "Random Forests", In Machine Learning, pages 5-32, 2001.

[20] Leo Breiman, "Bagging Predictors", In Machine Learning, pages 123-140, 1996.

[21] R. Olsen, L.Breiman, J.Friedman and C.Stone, "Classification and Regression Trees", Wadsworth and Brooks, Monterey, CA, 1984.

[22] A. Liaw and M.Wiener, "Classification and Regression by Random Forest", R News, 2(3):18-20, 2002.

[23] T K Ho, "The Random Subspace Method for Constructing Decision Forests", IEEE Transactions on Pattern analysis and machine intelligence,(20):832-844, 1998.

[24] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled-faces in the wild: A database for studying face recognition in unconstrained environments," Univ. Massachusetts Amherst, Amherst, MA, USA, Tech. Rep. 07-49, 2007.