Assignment 2: Face Classification/Verification

CSE 475: Statistical Methods in AI Monsoon 2019

Nivedita Rufus
2019702002

I. Introduction

The objective of this assignment is to get you familiarize with the problems of "classification" and "verification" with a popular problem space of "face".

II. Given Datasets

There are three datasets given to us:

A. Indian Movie Face Database

This dataset contains 400 face images of 8 Indian movie actors with 50 images each. All the images are selected and cropped from the video frames resulting in a high degree of variability interms of scale, pose, expression, illumination, age, resolution, occlusion, and makeup.

B. Yale Face Database

This dataset contains 165 images of 15 subjects with 11 images each. each subject's face has been captured with a different emotion. this makes it helpful to do an image emotion analysis on the dataset.

C. IIIT Cartoon Face dataset

This dataset contains cartooned faces of 8 different well-known celebrities. These images were procured through Google images searches and were carefully annotated.

III. Reconstruction of images from its Eigen faces

It can be observed that images can be represented by their its projection on their eigen vectors. Each of these projected images are called eigen faces. So this implies that we can represent the images using only a few of the eigen vectors i.e those vectors of correspoding to the most significant eigen values. The number of the eigen vectors to choose can be given found using the eigen spectrum a typical eigen spectrum looks like Fig(1,2,3). It can be observed that only the first few eigen vectors have significant values. The more eigen vectors used for projection lesser is the reconstruction error. The number of eigen vectors chosen for the IMFD dataset is 150, for Yale dataset it is 75, for the IIITcfw dataset it is 300. Their respective reconstruction errors were 0.03, 0.105, 0.06. For example the most difficult personality to represent with fewer eigen vectors, i.e. having the maximum reconstruction error in the IMFD dataset is Shahrukh Khan(can be verified in the code). It can also be observed that the IIITcfw dataset required more eigen vectors to in comparison with the other two datasets i.e it will be difficult to reconstruct the images of this dataset with fewer eigen vectors.

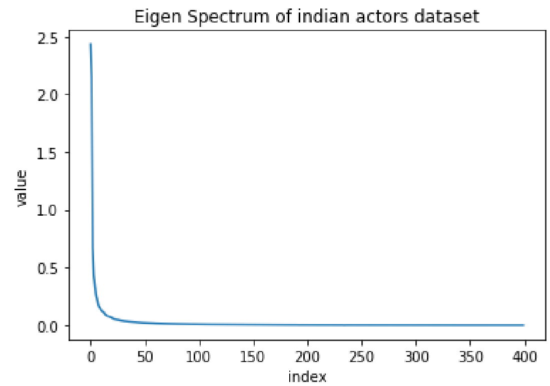

Fig. 1. Eigen spectrum of IMFD dataset

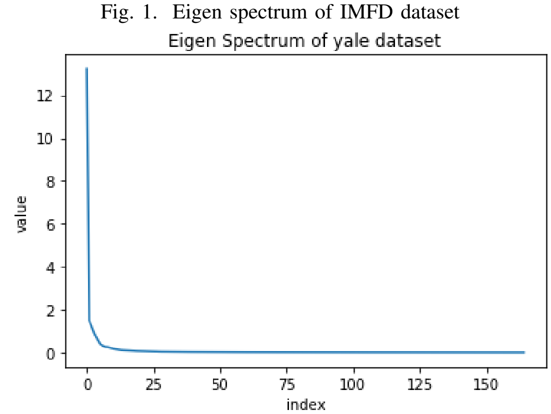

Fig. 2. Eigen spectrum of Yale dataset

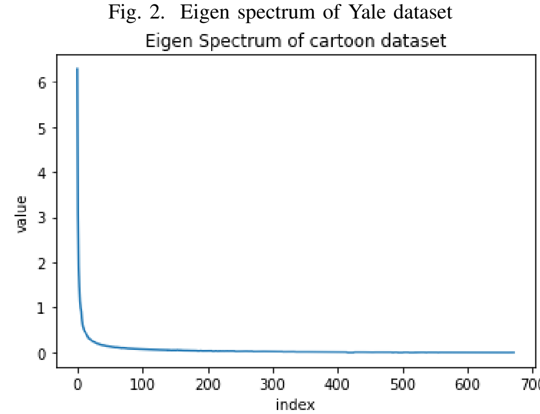

Fig. 3. Eigen spectrum of Carotoon dataset

IV. Comparative study of different classification paradigms on different feature spaces

The classification paradigms analysed in this assignment are :

- Support Vector Machine(SVM)
- Multi Layer Perceptron(MLP)

The paradigms mentioned above are tested on the following features for each of the three data sets and the best one was chosen :

- PCA features
- LDA features
- KernelPCA features(rbf kernel)
- KernelLDA features(rbf kernel)
- VGG features
- Resnet features

The performance metrics used to decide the best performing model were accuracy, classification error, reduced dimension space, F1 score etc. The desired conditions are higher accuracy and F1 score, lower dimension space, lower classification error.

A. The best model chosen for IMFD dataset

All the metrics are tabulated in Fig(4). Based on this tabulation the conclusion drawn for the best model is using LDA features with SVM. It can also observed that SVM on KernelLDA features also gives almost identical performance metrics, but SVM along with LDA is preferred as it avoids the computation of the kernelised LDA features making it computationally more efficient. The confusion matrix of the best model for this dataset is shown in Fig(5).

IMFD dataset :

Method	dimensions space	Accuracy	classification error	F1 score
SVM + PCA	150	1	0	1
MLP + PCA	150	0.475	0.525	0.308378
SVM + LDA	7	1	0	1
MLP + LDA	7	0.675	0.325	0.532216
SVM + Kernel PCA	150	0.25	0.75	0.122222
MLP + Kernel PCA	150	0.15	0.85	0.0384615
SVM + Kernel LDA	7	1	0	1
MLP + Kernel LDA	7	0.55	0.45	0.408755
SVM + VGG	4096	1	0	1
MLP + VGG	4096	0.975	0.025	0.977273
SVM + Resent	2048	1	0	1
MLP + Resent	2048	1	0	1

Fig. 4. Analysed models of IMFD dataset

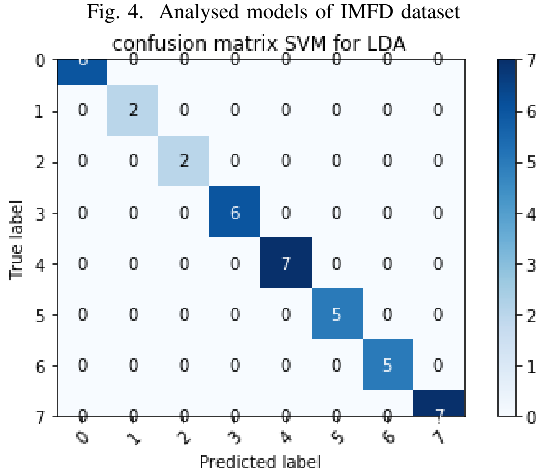

Fig. 5. Confusion matrix of the best model for IMFD dataset

B. The best model chosen for Yale dataset

Based on the tabulation shown in Fig(6) the conclusion drawn for the best model is using LDA features with SVM. It can also observed that SVM on KernelLDA features also gives almost identical performance metrics, but SVM along with LDA is preferred as it avoids the computation of the kernelised LDA features making it computationally more efficient.The confusion matrix of the best model for this dataset is shown in Fig(7).

YALE dataset :

Method	dimensions space	Accuracy	classification error	F1 score
SVM + PCA	75	1	0	1
MLP + PCA	75	0.764706	0.235294	0.676923
SVM + LDA	14	1	0	1
MLP + LDA	14	0.411765	0.588235	0.288462
SVM + Kernel PCA	75	0.764706	0.235294	0.704396
MLP + Kernel PCA	75	0.117647	0.882353	0.0161943
SVM + Kernel LDA	14	1	0	1
MLP + Kernel LDA	14	0.352941	0.647059	0.223443
SVM + VGG	4096	1	0	1
MLP + VGG	4096	0.764706	0.235294	0.665934
SVM + Resnet	2048	1	0	1
MLP + Resnet	2048	1	0	1

Fig. 6. Analysed models of Yale dataset

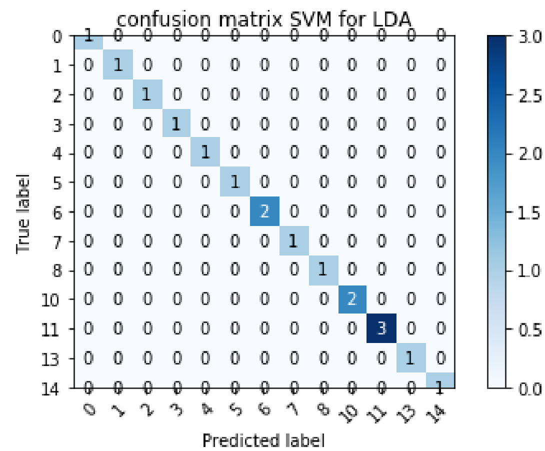

Fig. 7. Confusion matrix of the best model for Yale dataset

C. The best model chosen for IIIT Catoon face dataset

Based on the tabulation shown in Fig(8) the conclusion drawn for the best model is using PCA features with SVM. This model has the maximum accuracy and lowest classification error making it the ideal choice for classification.The confusion matrix of the best model for this dataset is shown in Fig(9).

V. T-SNE Based Visualisation

T-distributed Stochastic Neighbor Embedding (t-SNE) is a machine learning algorithm for visualization. It is a nonlinear dimensionality reduction technique well-suited for embedding high-dimensional data for visualization in a low-dimensional space of two or three dimensions. Specifically, it models each high-dimensional object by a two- or three-dimensional point in such a way that similar objects are modeled by nearby points and dissimilar objects are modeled by distant points with high probability. The 2-D t-SNE visualization can seen for all the

CARTOON dataset :

Method	dimensions space	Accuracy	classification error	F1 score
SVM + PCA	300	1	0	1
MLP + PCA	300	0.838235	0.161765	0.814874
SVM + LDA	7	0.985294	0.0147059	0.988337
MLP + LDA	7	0.323529	0.676471	0.179118
SVM + Kernel PCA	300	0.294118	0.705882	0.149009
MLP + Kernel PCA	300	0.308824	0.691176	0.106999
SVM + Kernel LDA	7	0.985294	0.0147059	0.988337
MLP + Kernel LDA	7	0.75	0.25	0.594621
SVM + VGG	4096	0.970588	0.0294118	0.974868
MLP + VGG	4096	0.75	0.25	0.626592
SVM + Resnet	2048	1	0	1
MLP + Resnet	2048	1	0	1

Fig. 8. Analysed models of IIIT Catoon face dataset

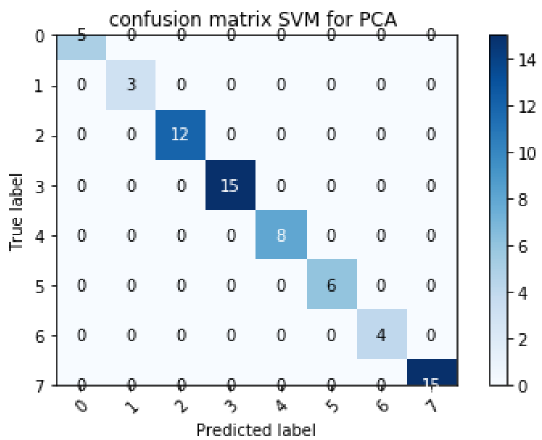

Fig. 9. Confusion matrix of the best model for Yale dataset

datasets in Fig(10,11,12). It can be observed that for the Yale data set shows the best distinction between classes. The IMFD dataset also shows decent distinction between classes. But for the IIIT cartoon database this representation does not help in distinguishing the classes.

VI. Face verification using KNN algorithm

The KNN algorithm was applied to three features spaces of each of the datasets namely:

- PCA features
- LDA features
- KernelLDA(rbf kernel) features

A. KNN for the IMFD dataset

As mentioned above the KNN algorithm was done for three of the above mentioned feature spaces. These results are tabulated in Fig(13). Based on the metrics used for analysis it can be observed that KNN with LDA works the best for verification giving highest accuracy on the test set. It can also observed that KNN on KernelLDA features also gives almost identical performance metrics, but KNN along with LDA is preferred as it avoids the computation of the kernelised LDA features making it computationally more efficient.The confusion matrix of the best model for this dataset is shown in Fig(14).

B. KNN for the Yale dataset

These results of the KNN algorithm on the three feature spaces of the Yale dataset are tabulated in Fig(15). Based on

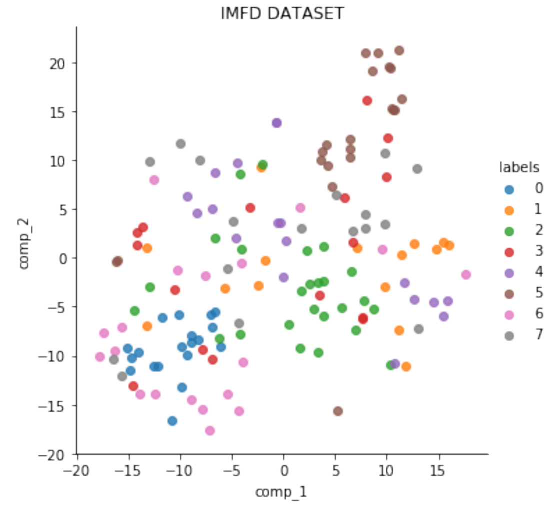

Fig. 10. 2D t-SNE representation of IMFD dataset

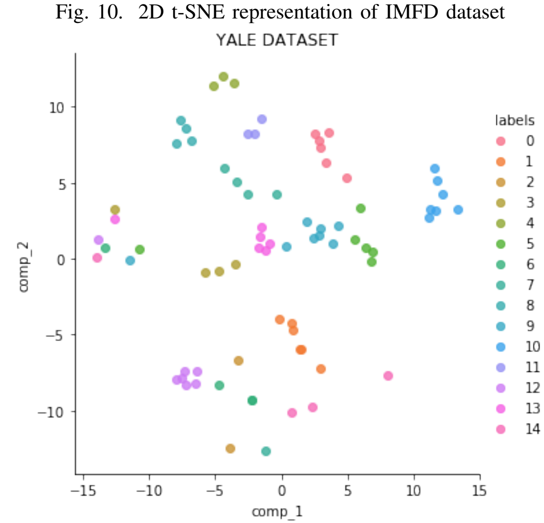

Fig. 11. 2D t-SNE representation of Yale dataset

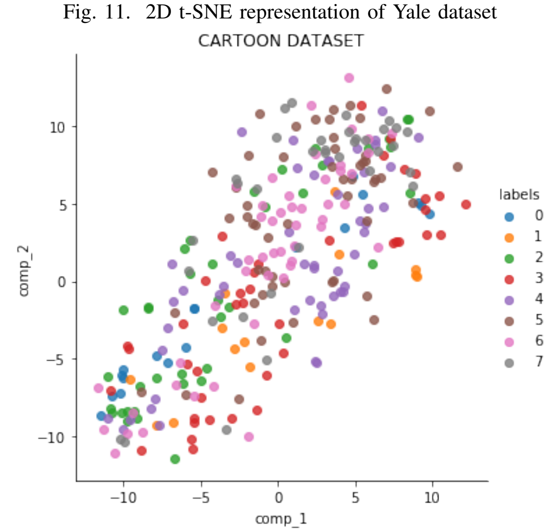

Fig. 12. 2D t-SNE representation of Carotoon dataset

IMFD dataset:

Method	dimensions space	Accuracy	Verification error	Precision
KNN + PCA	150	0.675	0.325	0.671789
KNN + LDA	7	1	0	1
KNN + Kernel LDA	7	1	0	1

Fig. 13. Analysed models using KKN for IMFD dataset

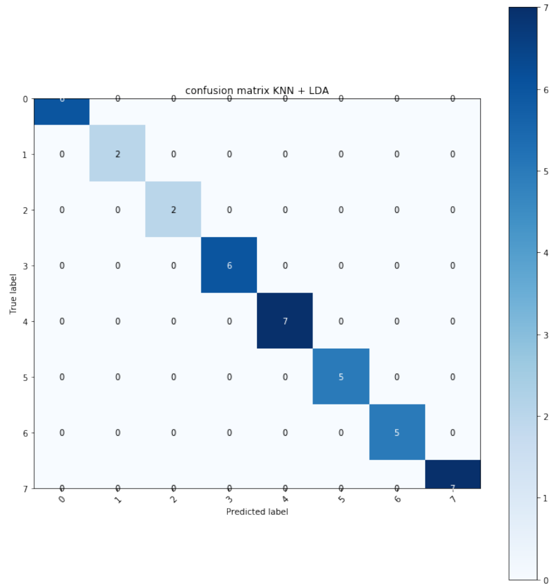

Fig. 14. Confusion matrix of the best model of KNN for IMFD dataset

YALE dataset:

Method	dimensions space	Accuracy	Verification error	Precision
KNN + PCA	75	0.764706	0.235294	0.668498
KNN + LDA	14	0.823529	0.176471	0.689796
KNN + Kernel LDA	14	0.823529	0.176471	0.689796

Fig. 15. Analysed models using KKN for Yale dataset

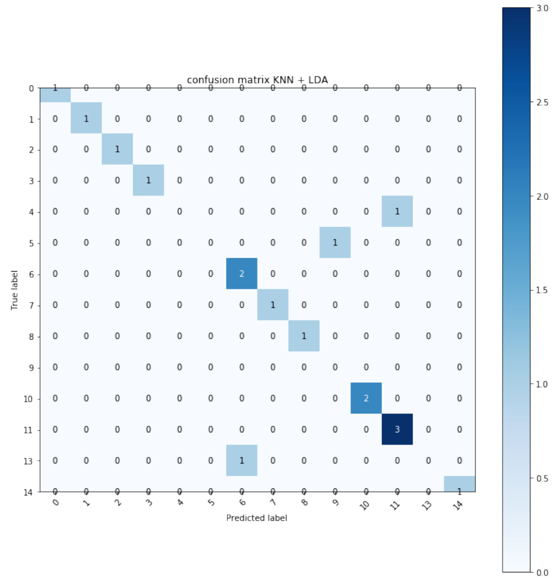

Fig. 16. Confusion matrix of the best model of KNN for Yale dataset

the metrics used for analysis it can be observed that KNN with LDA works the best for verification giving highest accuracy on the test set. It can also observed that KNN on KernelLDA features also gives almost identical performance metrics, but KNN along with LDA is preferred as it avoids the computation of the kernelised LDA features making it computationally more efficient.The confusion matrix of the best model for this dataset is shown in Fig(16).

C. *KNN for the IIIT Cartoon dataset*

These results of the KNN algorithm on the three feature spaces of the IIIT cartoon dataset are tabulated in Fig(17). Based on the metrics used for analysis it can be observed that KNN with LDA works the best for verification giving highest accuracy on the test set. It can also observed that KNN on KernelLDA features also gives almost identical performance metrics, but KNN along with LDA is preferred as it avoids the computation of the kernelised LDA features making it computationally more efficient.The confusion matrix of the best model for this dataset is shown in Fig(18).

CARTOON dataset:

Method	dimensions space	Accuracy	Verification error	Precision
KNN + PCA CARTOON dataset	300	0.558824	0.441176	0.537229
KNN + LDA	7	0.911765	0.0882353	0.908547
KNN + Kernel LDA	7	0.911765	0.0882353	0.908547

Fig. 17. Analysed models using KKN for IMFD dataset

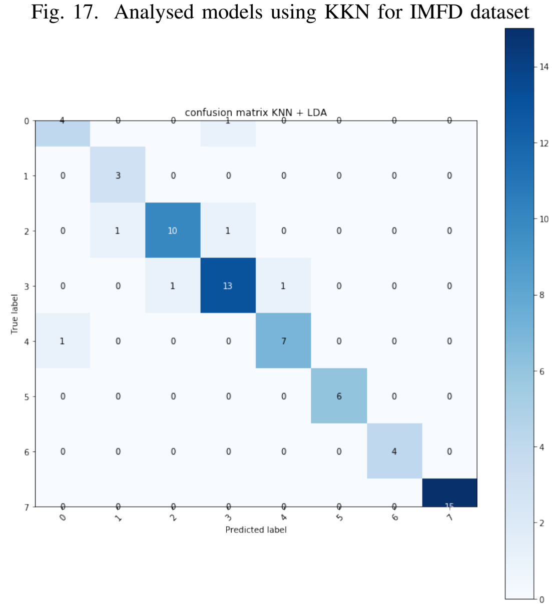

Fig. 18. Confusion matrix of the best model of KNN for IIIT Cartoon dataset

VII. Extension/Application: Cartoon vs Real

The task taken up is to identify if an image is a cartoon or real. For this purpose the IIIT Cartoon Face dataset and the IMFD dataset were concatenated and shuffled. The labels for all images in the IMFD dataset are set as 0 and those of the cartoon dataset were set as 1. To choose no. of significant eigen vectors the eigen spectrum of this dataset was plotted and 400 eigen vectors were chosen to represent the dataset in PCA feature space. The eigen spectrum of the data set is shown in Fig(19) and the PCA representation of the dataset is shown

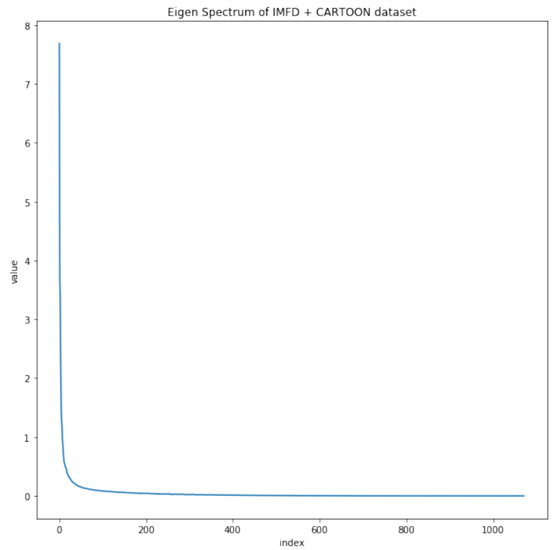

Fig. 19. Eigen Spectrum of the IMFD + IIIT Cartoon Dataset

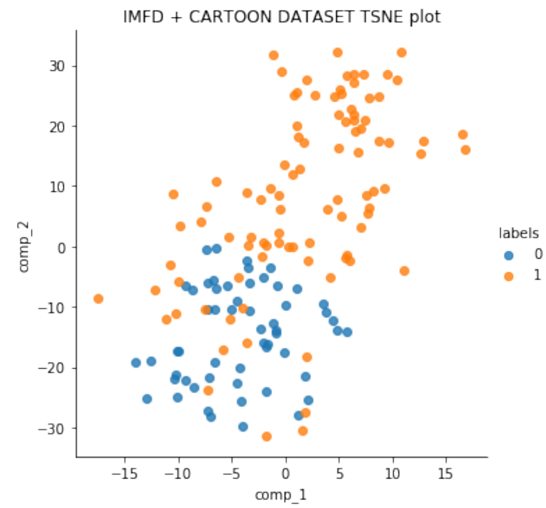

Fig. 21. 2D t-SNE representation of the IMFD + IIIT Cartoon Dataset

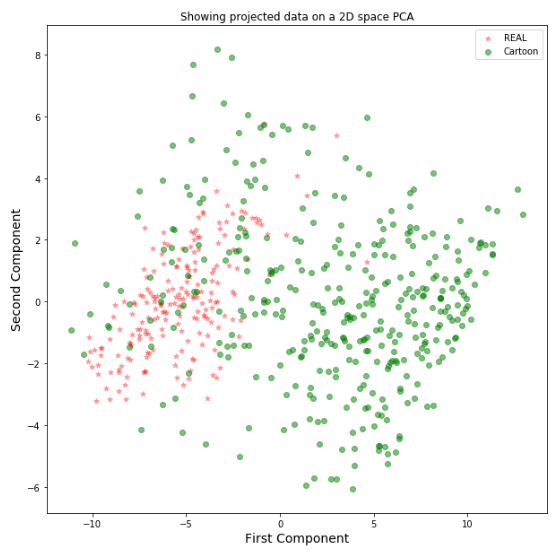

Fig. 20. 2D PCA features of the IMFD + IIIT Cartoon Dataset

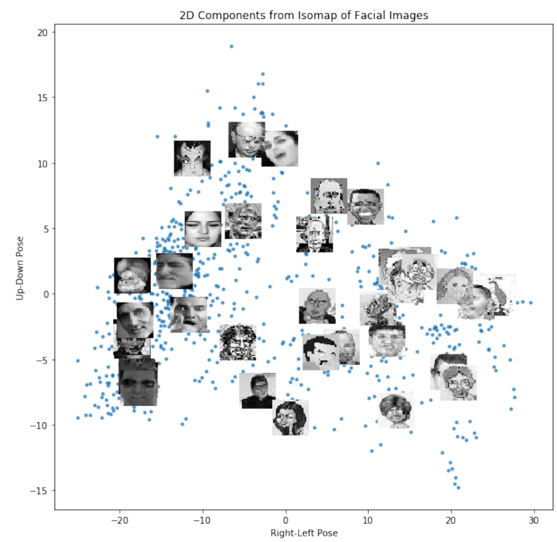

Fig. 22. 2D isomap of facial features of the IMFD + IIIT Cartoon Dataset

in Fig(20). Note that Fig(20) shows the representation of the dataset along only the first two principal components, but for classification purpose all 400 features were used. The t-SNE plots for the dataset was also analysed and this is shown in Fig(21). The 2D isomap plot of the facial images of the dataset also is shown in Fig(22). The dataset is divided into train, validation and test sets. The classifier used for this purpose was SVM. Since there only two classes it makes sense to use a linear kernel for the SVM. This accuracy obtained on the test set was 0.94. The confusion matrix for the classification is shown in Fig(23).

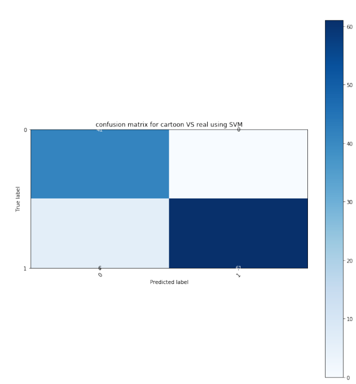

Fig. 23. Confusion matrix of SVM on the IMFD + IIIT Cartoon Dataset