Face Recognition - Eigenfaces

Danny Styvens Cardona Pineda, Juan Camilo Ceballos Arias, Juan Esteban Torres Marulanda, Miguel Angel Mejia Muñoz

ABSTRACT

The aim of the project is to recognize a person's face by making a comparison between its characteristics and those of other people's faces which are already known. To perform this task, Eigenfaces method is implemented together with a previous preparation of the images.

Initially, outlier images are identified within a face and landscape dataset, using metrics such as Manhattan, Euclidean, Chebyshev and Minkowsky distances ($p = \frac{5}{2}$ and $p = \frac{\sqrt{2}}{2}$). Then the results are compared to determine which one of the metrics has better performance in the identification of rare images. In total 2.470 images (2.260 faces and 210 natural landscapes) were used. Secondly, the methodology of Eigenfaces was implemented with a total of 3059 images of faces and then, by distance measurements, distances of new images projected on the subspace were compared with those used to form the Eigenfaces.

Thirdly, a lineal discriminant analysis (LDA) was used on the subspace generated through PCA.

Finally, for the prediction of faces of women, men and landscapes, supervised and unsupervised classification methods were used and compared. For the supervised classification, the Multinomial Logistic Regression and Linear Discriminant Analysis (LDA) were used, and the unsupervised analysis was done using K-means, T-Distributed Stochastic Neighbor Embedding (TDSNE) and Agglomerative Clustering.

The code, datasets, Readme and a document type article are hosted in an open Github repository at https://github.com/mmejiam-eafit/ms_data_science_project_i. It also includes a notebook with all the stuff necessary to replicate the work.

Keywords: Distance, Metric, Outlier, Eigenface, Classification, Supervised, Unsupervised, Clustering, Visualization, Embedding, Dimensionality reduction

1. INTRODUCTION

For the problem of face recognition, it is necessary to identify if a certain face belongs to a set of faces that are already known (recognition), by means of characteristics that can be extracted from the set of images. Different approaches have been proposed such as geometrical characteristics of the faces, and holistic methods. This exercise will focus on the group of holistic methods since they have shown good results, overcoming some of the difficulties related with working on geometrical features such as face symmetry (*Zhao, Chellappa, Phillips, and Rosenfeld, 2003* [12]).

Thanks to technological advances especially in computer science, development and improvement of mathematical models and new algorithms, important advances have been implemented in automatic face recognition starting in the 70s and onwards (Zhao et al., 2003). There are two important topics in automatic face recognition:

1. Detection of a face, for which segmentation methods are used.
2. Features extraction algorithms using linear combinations of characteristics of the images on a dataset, which this project will be focusing on.

Approaches as Eigenfaces and Fisherfaces have showed good results. Eigenfaces approach is based on the Principal Component Analysis (PCA) (*Turk and Pentland, 1991* [11]) and Fisherfaces is based on the Linear Discriminant analysis (*Etemad and Chellappa, 1997* [3]); both methods have advantages and disadvantages depending on the use: classification or image representation (*Belhumeur, Hespanha, and Kriegman, 1997* [2]).

Although the holistic approaches have solved some difficulties, it is still important to address other approaches such as the variation of illumination in the images and facial expressions. These two issues are still a challenge for these methods (Zhao et al., 2003), but methodologies like deep neural network could be a good alternative to deal with them.

This document presents the initial phase of the project, which identifies outlier images within a face and landscape dataset, by using metrics such as Manhattan, Euclidean, Chebyshev and Minkowsky distance ($p = \frac{5}{2}$) (*Macho, 2010* [9]). Additionally Minkowsky similarity ($p = \frac{\sqrt{2}}{2}$) is also included. Afterwards, a comparison is made among the results to determine which one of the metrics has a better performance in the identification of rare images. Then, the Eigenfaces method is developed to make comparisons for face recognition and finally some supervised and unsupervised classification methods are used to predict whether an image belongs to a specific group of images or not.

2. BACKGROUND AND RELATED WORK

2.1 Eigenfaces

Eigenfaces method is developed to make comparisons for face recognition (*Turk and Pentland, 1991* [11]). Basically, what is made is to extract the most important information of a face in order to compare it with a database's. Mathematically, it consists in calculating the images set's covariance matrix eigenvectors so that a space being made of them is created. While working on image recognition the eigenvectors are called Eigenfaces and on its part, the eigenvector space, face space. Every face in the images set can be constructed as a linear combination of the vectors belonging to this, which is also the basis of the whole methodology since the recognition process refers to measuring the distance between an image and this space and depending on how long it is, to classify the face as a known or unknown one.

Note that the process described for eigenfaces is very similar to the analysis of principal components (PCA) [6], in fact, it is a variation of the PCA in which the variables that form the space of work, corresponds to m images considered for the analysis. Consider a set of two-dimensional images collapsed in an array A_{mn}, where the n columns correspond to the number of pixels (high by width) and the m rows correspond to the number of images, $\tilde{A}$ array is defined as:

$$\tilde{A} = A - \Phi \tag{1}$$

where Φ is a factor that centralizes A. This Φ array can be the mean or some robust estimator of centralization (p.e median). Then the covariance matrix is defined as:

$$S = \tilde{A}'\tilde{A} \tag{2}$$

Now it is possible to calculate an efficient decomposition of S by its singular value decomposition, obtaining their respective eigenvalues and eigenvectors.

These eigenvectors (v_i) and eigenvalues (λ_i) allow to identify and build a new subspace in which the greatest possible variability of the space is captured with the least possible number of components. Finally, the corresponding image weights associated with the components of subspace are calculated:

$$w_i = \tilde{\Gamma}_i V \tag{3}$$

where $\tilde{\Gamma}$ is a collapsed and centralized image, and V contains the normalized eigenvectors of subspace. Note that to rebuild the projected image in the subspace, the next operation can be followed:

$$\Gamma_i = w_i V^T + \Phi \tag{4}$$

It has to be taken into account that the total number of possible Eigenfaces is equivalent to the number of images of the whole dataset. Likewise, the faces can be approximated using only the best M Eigenfaces, which are those with the highest eigenvalues associated with the greatest variance in the images. The number M is usually determined holistically, although other proposed methods exist in the literature (*Peres, Jackson and somers, 2005* [10]).

The process of facial recognition can be summarized in these 4 steps [11]:

1. Calculate the eigenfaces of some images, thus defining the face space.
2. When a new image is introduced, weights of incoming image and the M eigenfaces are calculated by projecting the incoming image in each Eigenface.
3. Determine if the image is a face. This is defining whether the image is relatively close to the face space or not.
4. If the new image is a face, its weight pattern is classified to a previously known or unknown person.

2.2 Linear Discriminant Analysis (LDA)

Linear discriminant analysis (LDA) can be considered as a special case of regression, where the dependent variable is categorical and the independent variables are continuous. It consists in finding the linear combination between the independent variables that best discriminates the groups. To work on LDA, the following conditions need to apply [5]:

1. The response variable is categorical and the predictions are continuous
2. There are at least two groups and for each group two or more cases are needed.
3. p < n-2, where p is the number of variables and n the number of subjects
4. No discriminant variable is a linear combination of other discriminant variables.
5. The maximum number of discriminant functions is equal to the minimum between the number of variables and the number of groups minus 1.
6. The covariance matrices within each group should be approximately equal.
7. Continuous variables must follow a normal multivariate distribution.

3. DATASETS

For the development of this project, a group of images hosted in a free online database will be used. This database has four different directories holding the images in different levels of difficulty as follows: faces94, faces95, faces96 and grimace. The last two are more complex due to the images variation on background and scale and the type of facial expressions in them.

The whole set has 7900 images belonging to 395 individuals. Different genders and races are shown, people wearing glasses and beards are also taken into account and, regarding the age range, most of data corresponds to first-year undergraduate students between 18 and 20 years old, even though some older people are present in the data as well. For this first phase, Dataset Faces94 will be used.

Faces94 is a collection of images consisting of a wide range of people's pictures taken speaking in front of camera. Because of the speech, this set is an introduction to the variation in facial expression. Faces94 has 153 individuals images using portrait format. It contains pictures of male, female and male staff in separate directories. The pictures background is plain green. It does not have any individual's variation on head scale and image lighting, but it does have a few on head turn, tilt and slant, and considerable on facial expression. Additionally, there is no individual hairstyle variation as the images were taken in a single session.

Figure 1. Faces94 face images

The methodology of Eigenfaces was implemented with a total of 3059 images of Faces94 (faces94 male,faces94 female,faces94 malestaff).

Additionally, natural landscape images were included. These images were obtained from ImageNet database and each one of their links are online `http://image-net.org/api/text/imagenet.synset.geturls?wnid=n13104059`. Cv2 package was used to read and resize the images and then a Numpy array was created with a gray scale of the images.

Figure 2. Natural landscape images

4. ARCHITECTURE, MODELS AND DATA PREPARATION

Only images in JPG format were used in the project. They were originally colored (3-D) and converted into grayscale (2-D) due to practical reasons of computation, but it is important to mention that the model could be generalized to the case of 3-D.

Python programming language is used for the development of the project which considers some stages as follows:

a) Data preparation, in which the colored images are converted into grayscale and resized in such a way that they can be compared. This process allows to have a data matrix where each element corresponds to a pixel of the image; cv2 package in python was used for that purpose.
b) Functions are defined to read and manipulate images inside the dataset which contains subfolders (nested structure).
c) Functions are written to do distances calculation, outliers detection, accuracy of the metrics determination and finally graphing the results.
d) A Jupyter Notebook was created in which functions are run and results are depicted. These other packages were used: Pandas, Numpy, Matplotlib and Collections.

The datasets, the functions and the Notebook are located in a Github repository, so that each member of the team makes their contributions after developing and testing locally.

5. RESULTS

5.1 Metrics comparison and outliers detection

In this project, comparison measures of efficiency between metrics were used. Table 1 depicts the global accuracy efficiency between metrics. In all cases, the measure was greater than 87% for both criteria detection ($Q3 + 1.5 * IQR$ and $90th$ percentile).

Table 1. Accuracy of metrics based on the mean and median image.

	Outlier selection criteria*			
	Percentile 90		**Q3 +1.5*IQR**	
Metrics	**Distance to the mean**	**Distance to the median**	**Distance to the mean**	**Distance to the median**
Manhattan ($p = 1$)	88.06	89.51	91.58	92.20
Euclidean ($p = 2$)	89.68	90.49	92.63	93.40
Minkowsky ($p = 3$)	89.27	90.00	93.04	93.00
Chebyshev ($p = inf$)	89.19	88.54	90.57	91.00
Minkowsky ($p = \frac{5}{2}$)	89.51	90.16	93.52	93.20
Minkowsky ($p = \frac{\sqrt{2}}{2}$)	87.49	89.35	91.13	92.10

* accuracy (%) = (true positive + true negative)/N

Table 2 depicts false negatives and true negatives. For true negatives, the percentage with respect to the total of outliers (210) was calculated. Overall outlier detection was less than 53%.

Table 2. Number of true and false negatives obtained, based on the mean and median of images.

	True negatives (% true outlier accuracy)		**False negatives**	
	Percentile 90		**Percentile 90**	
Metrics	**Distance to the mean**	**Distance to the median**	**Distance to the mean**	**Distance to the median**
Manhattan ($p = 1$)	81 (38.57)	99 (47.14)	166	148
Euclidean ($p = 2$)	101 (48.09)	111 (52.86)	146	136
Minkowsky ($p = 3$)	96 (45.71)	105 (50.00)	151	142
Chebyshev ($p = inf$)	95 (45.24)	90 (42.86)	152	163
Minkowsky ($p = \frac{5}{2}$)	99 (47.14)	107(50.95)	148	140
Minkowsky ($p = \frac{\sqrt{2}}{2}$)	74 (35.24)	97 (46.19)	173	150

No large differences were observed when comparing the distances to the mean and the median of the images (tables 1 and 2).

For the calculation of the real median it was necessary to compute and sum the distances of all images to each other, in order to select the minimum result.

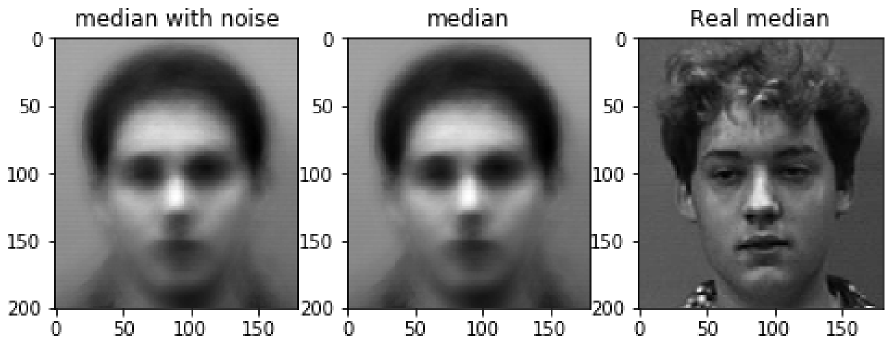

Figure 3. Mean face

Table 3. Confusion matrix of metrics/similarity, according to the mean image

Metric/similarity	**True positive**	**False negative**	**False positive**	**True negative**
Manhattan $(p = 1)$	2094	166	129	81
Euclidean $(p = 2)$	2114	146	109	101
Minkowsky $(p = 3)$	2109	151	114	96
Chebyshev $(p = inf)$	2108	152	115	95
Minkowsky $(p = \frac{5}{2})$	2112	148	111	99
Minkowsky $(p = \frac{\sqrt{2}}{2})$	2087	173	136	74

Table 4. Confusion matrix of metrics/similarity, according to the median image

Metric/similarity	**True positive**	**False negative**	**False positive**	**True negative**
Manhattan $(p = 1)$	2112	148	111	99
Euclidean $(p = 2)$	2124	136	99	111
Minkowsky $(p = 3)$	2118	142	105	105
Chebyshev $(p = inf)$	2097	163	120	90
Minkowsky $(p = \frac{5}{2})$	2120	140	103	107
Minkowsky $(p = \frac{\sqrt{2}}{2})$	2110	150	113	97

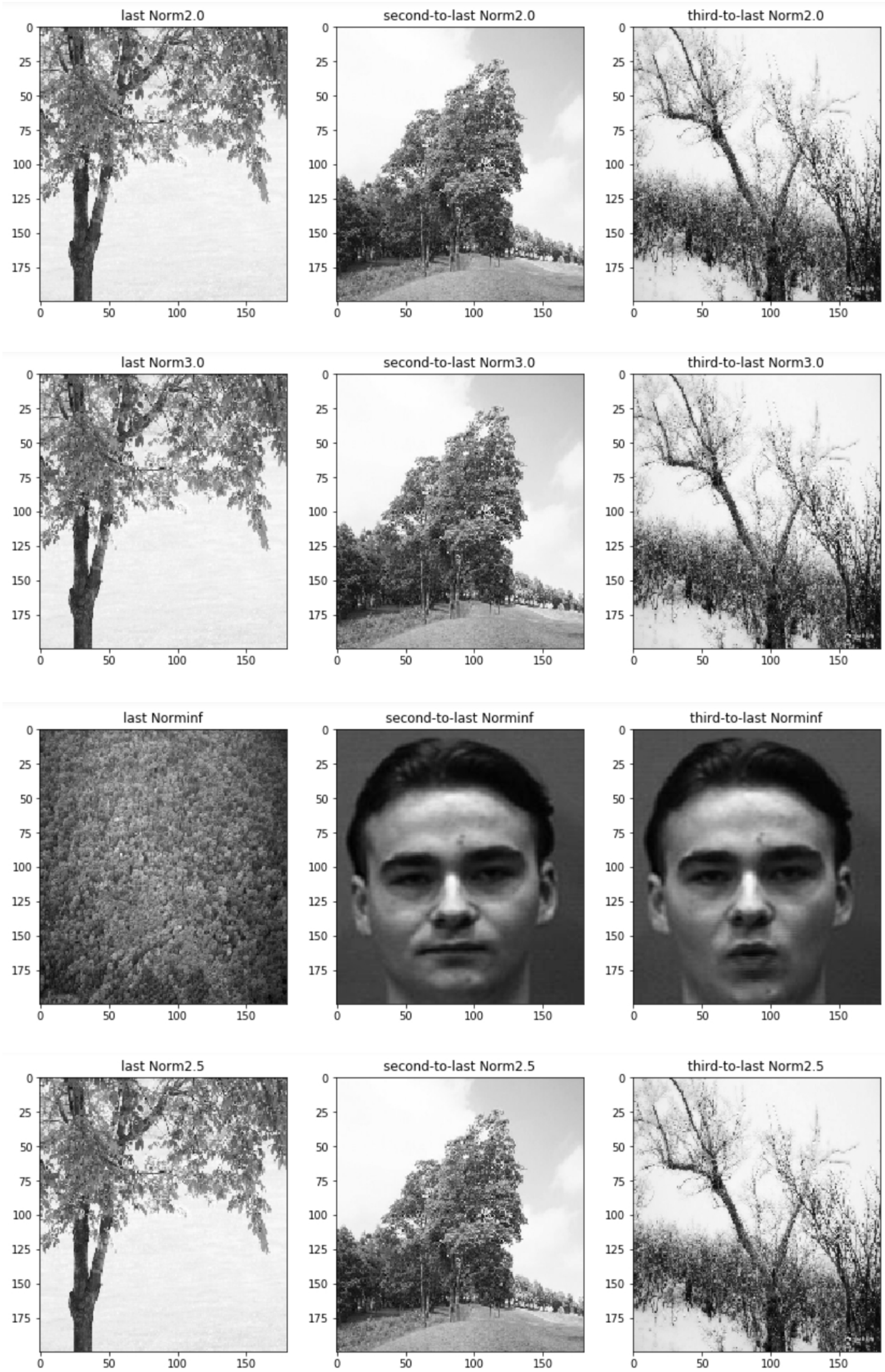

Figure 4. Outliers - against mean

5.2 Eigenfaces

The first step was to centralize the data in order to calculate the covariance matrix. This was made by means of calculating the mean image for later subtracting it to the whole dataset and performing a dot product of this last matrix and its transposed to get the covariance one. The product was done to have the covariance matrix dimensions not in terms of the dataset's images height and width but the number of images wise, as a strategy to reduce the size of the result.

5.2.1 Subspaces

After having the covariance matrix, the next step was to compute its singular value decomposition to get the eigenvectors and eigenvalues. Then, using the results, the face space started to be created by selecting the number of components of the subspace. That selection was run using two options:

• Investigator's criteria of variability captured, where the number of principal components were obtained from a prior set representation percentage. This option raises from how disperse the values are and the percentage of variability captured, so cumulative sums of variance were performed until getting the desired representation, in this study, 98%

• Investigator's criteria of threshold contribution value, where the approach to choose the number of components was to analyze the whole variation as a percentage.

Finally, option 2 was used to set the total of components to create the face space. They were the first 305 components, with 97.59% of variability captured and from which the contribution was less than 0.01%.

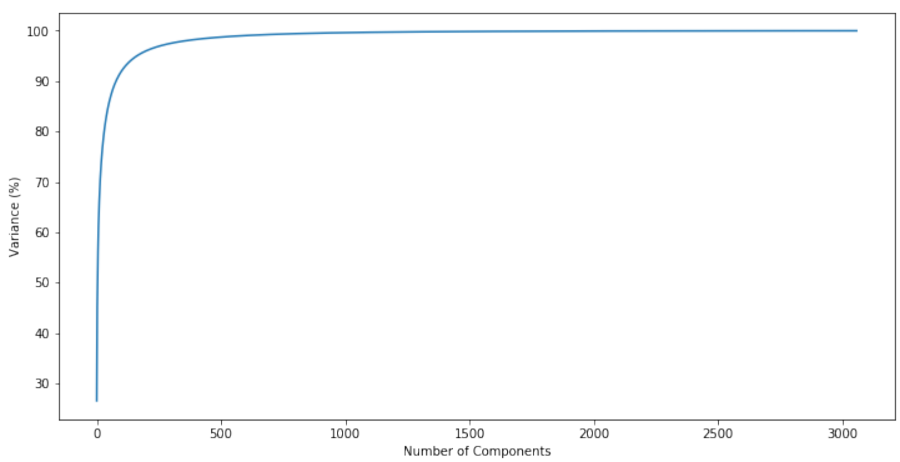

Figure 5. Cumulative summation of the explained variance

5.2.2 Face space

After setting the number of principal components that were going to be used to create the face space, the calculations were done: the dot product of the data and the eigenvectors selected was performed and the result was normalized.

Figure 6. First sixteen components of face space

5.2.3 Projection of an image on the face space

As every image of the dataset can be constructed as a linear combination of the face space's vectors, some image projections were made to observe how they were reconstructed by the model. To begin with, the weight (w) of each Eigenface in the generated subspace due to the original image, was computed. Then, that image was reconstructed by a dot product of w and the face space, and later adding the mean image.

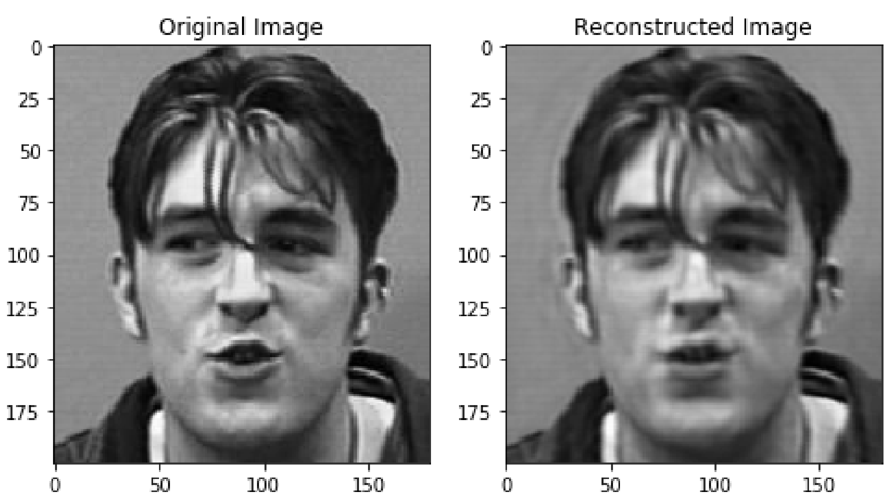

Figure 7. Projection of an image on face space

5.2.4 Distances and outliers

The distances between the original images and those reconstructed from the subspace, were calculated and atypical data were identified (greater distances between the original image and those reconstructed), the ones with the smallest distances (the best reconstructed ones) were also identified.

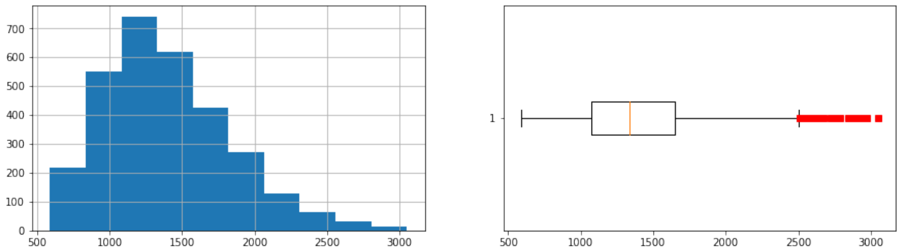

Figure 8. Distances (norm 2) distribution

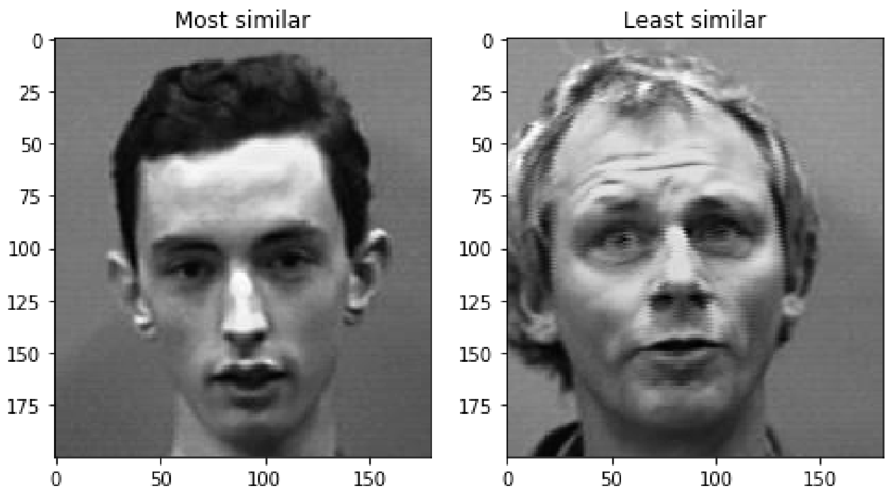

Figure 9. Most and least similar reconstructed faces to original image

Figure 10. Least nine similar reconstructed faces to original image

5.2.5 Face recognition

To prove whether or not it is possible to detect if an image is a face or not using the generated subspace, a test is performed by projecting images of landscapes and its distance is calculated. Large distances are observed (outside the range of distances), it is concluded that the image is not a face.

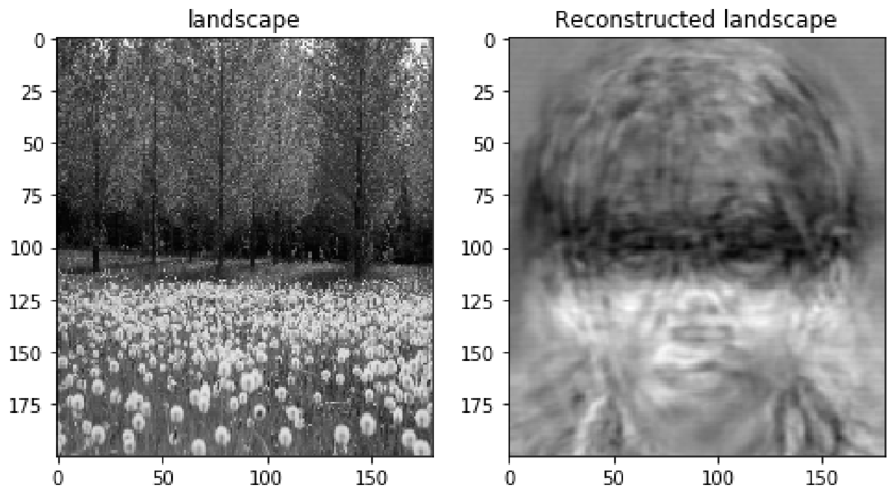

Figure 11. Image does not belong to the set of faces (distance>max(faces to face-space)

5.3 supervised classification

Three classes were created in the dataset, the first one is landscapes (label: 0), the second includes both "Male" and "Malestaff" (label: 1) and third represents women (label: 2).

The adjustment of the model was calculated considering the w_i of the images on the subspace generated by Eigenfaces. At this time, 62 components were selected, which capture 85% of the variability of the entire space. The component number were selected according to first criterion shown in section 5.2.1.

Note that in this case the required component number is greater compared to the subspace generated from the dataset that does not include the landscapes, this is an indication that with the inclusion of landscapes, more noise was introduced into the original space.

The predictive variables correspond to the respective weights w_i of each image on the generated subspace, this operation generates a matrix $A_{3268,62}$ in which each row corresponds to an individual (image), and each column corresponds to the weights obtained by component.

The percentage of data chosen for training and test was 70% and 30% respectively.

5.3.1 Multinomial logistic regression

The multinomial logistic regression (or multiclass) can be considered as a generalization of the binomial logistic regression, where the probability of occurrence follows a multinomial probability distribution[4].

An accuracy of 98.06% was observed, in table 5 the precision is shown for each of the categories

Table 5. Classification Measures of multinomial logistic regression

Measures	Precision	Fisher score	Recall
Males	0.99	0.99	0.99
Females	0.97	0.98	0.98
Landscapes	0.97	0.93	0.90

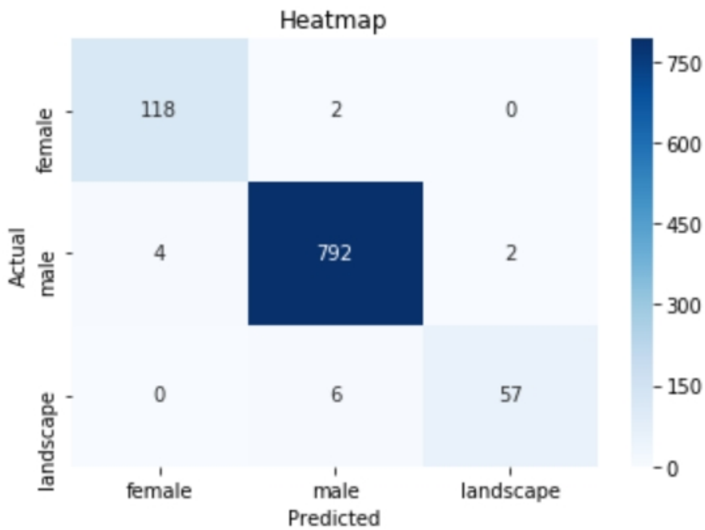

Figure 12. heatmap Logistic regression

5.3.2 Linear Discriminant Analysis

Some authors suggest the usefulness of using PCA and LDA together for classification[7]. Below are the results of applying LDA to the space generated with PCA.

Through LDA a subspace with two components was obtained and it is observed that the methods are capable of discriminating the three created classes (females, males, landscapes).

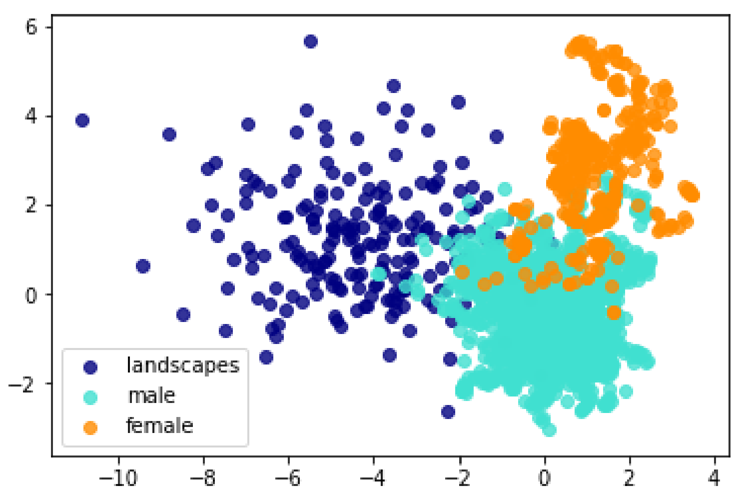

Figure 13. LDA of EigenFaces distribution

Table 6. Classification Measures of Fisher discriminant analysis

Measures	Precision	Fisher score	Recall
Males	0.96	0.97	0.97
Females	0.83	0.81	0.80
Landscapes	0.93	0.87	0.83

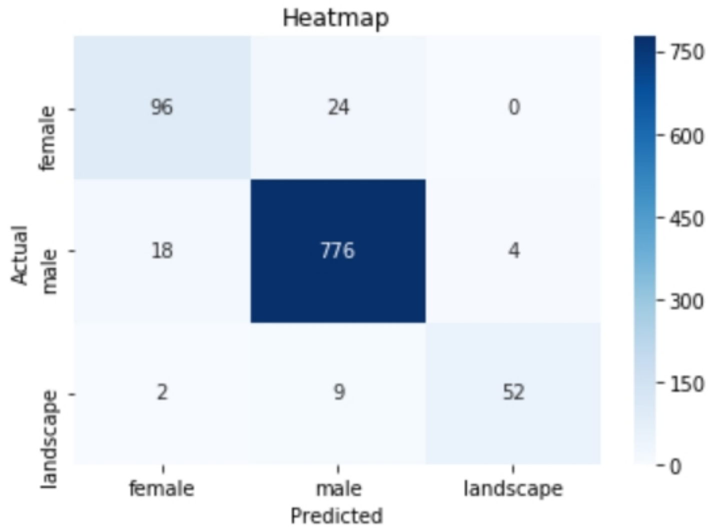

Figure 14. heatmap LDA

5.4 unsupervised classification

5.4.1 Kmeans

Kmeans is an iterative method similar to the Hierarchical cluster, a similarity criteria are used to define groups, but this time the similarity is defined according to how close a point of a "centroid" is, which forces to define a-priori cluster numbers. The principle of this method is to assume that there are K centers in the space such that around said centroids the data are concentrated (minimum distances from each point to a centroid) [1].

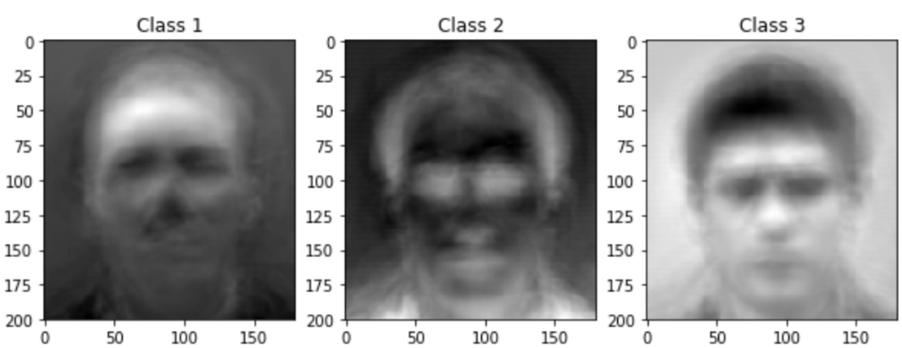

Figure 15. K means Classes

In class 1 the k-means model classifies images giving importance over forehead in faces. The class 2 focuses on the background of the images. And class 3 focuses in the upper part of the images, specifically in the hair.

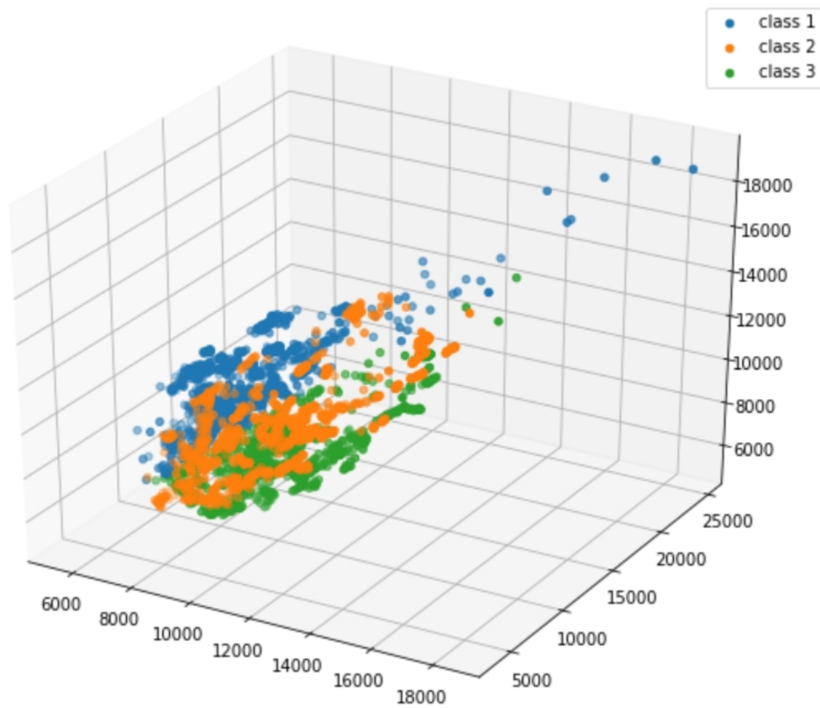

Figure 16. k means visualization

The K-means model couldn't classify initial labels (male, female and landscape). The data has a strong concentration in male's images (overlapping) in contrast to female and landscape's images, so the model is sensitive to initial conditions of images and the proportion to categories.

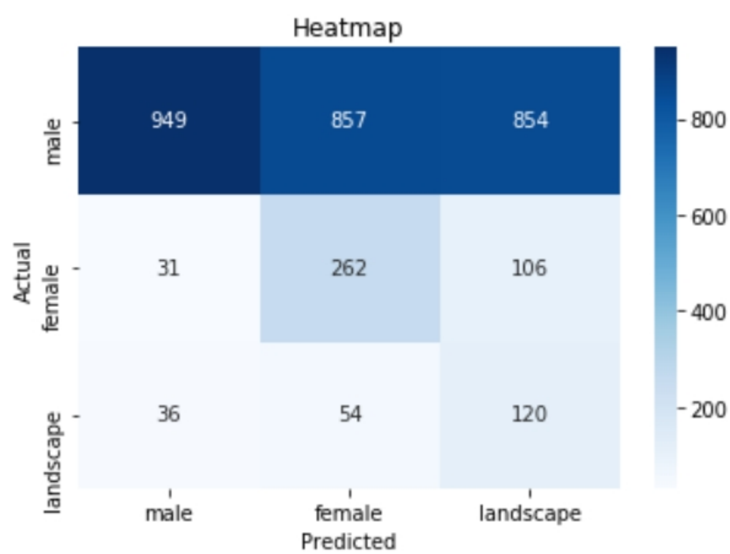

Figure 17. heatmap k-means

Table 7. Classification Measures of Kmeans

Measures	**Precision**	**Fisher score**	**Recall**
Males	0.93	0.52	0.36
Females	0.22	0.33	0.66
Landscapes	0.11	0.19	0.57

5.4.2 t-SNE

t-Stochastic Neighbor Embedding (t-SNE for short) is a non-linear dimensionality reduction/visualization method for high dimensional data on a low dimensional embedding (generally a 2 or 3-dimensional space). This subspace embedding is constructed by in such a way as to minimize the loss of information from the probability distribution of the original data when using a new probability distribution from modelling the data on a lower dimensionality. This means that all objects in dimension $\mathbb{R}^n$ with a distribution $p(x)$ that are similar based on a notion of distance have a similar value $q(x)$ on the new dimensionality $\mathbb{R}^k$ (where $k < n$), attempting to maintain this reasoning of distance from the original set. This is done by minimizing the Kullback-Leibler divergence between the two distributions. This algorithm uses the Euclidean distance as a measure for proximity of data points.

Applying this method to the result of the previous dimensionality reduction using PCA, we try to find a 2-dimensional embedding of the vectors that represent each of the images on the eigenface subspace. This embedding then will give us a representation of the original data (i.e. the projection of each image on the eigenface subspace) that follows a similar distribution on our new 2-dimension space. This approach, although it provided valuable insights as to what the shape of the projected data is (i.e. casting a shadow of the 62-manifold on our 2 dimension space), it showed no pragmatic value in regards as to what classification methods are concerned. In fact, given the embedding that was shown, and because the amount of data-points for the male class surpass the amount of the other classes by an order of magnitude, we see that the cloud generated by these data-points dominates the embedding, rendering moot any classification efforts using this visualization. [8].

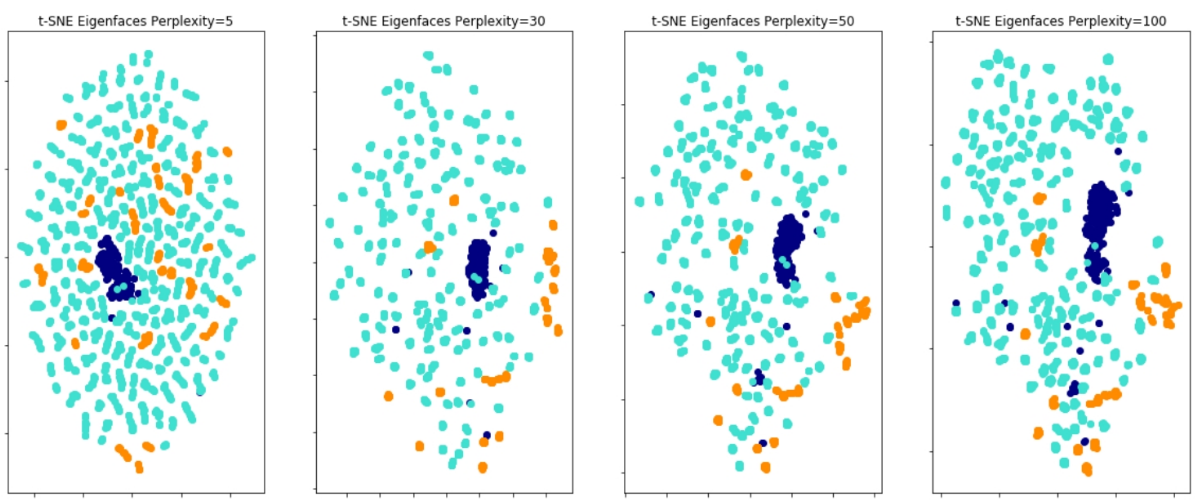

Figure 18. t-SNE visualization

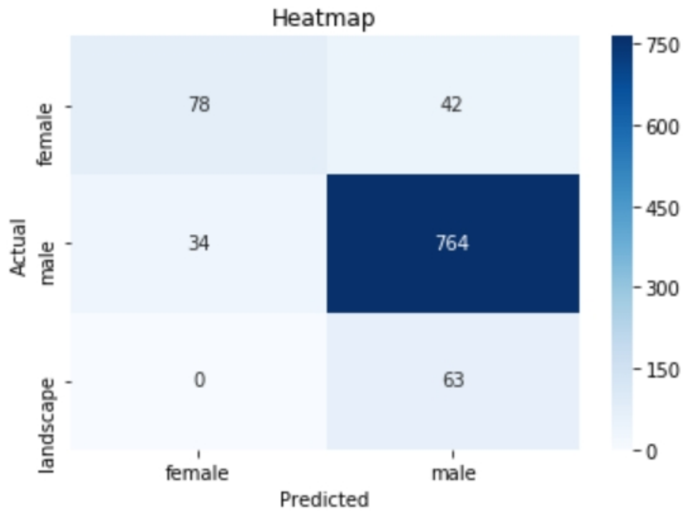

Figure 19. heatmap t-SNE

5.4.3 Agglomerative clustering

We used an agglomerative clustering approach as another unsupervised method for clustering the data projected unto our eigenface subspace, and a dendrogram was used in order to visualize the clustering of our data-points based on a weighted distance approach. Although this method warrants further exploration of how the data is distributed in our 62-manifold of data-points, we see that no insight was gained from performing this study other than its academic and exploratory purposes for our specific data-set.

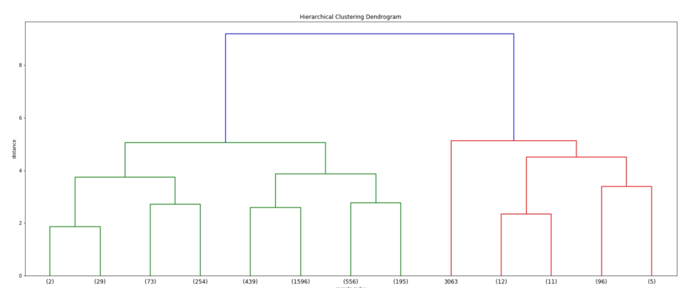

Figure 20. Agglomerative clustering

6. CONCLUSIONS

The outlier detection presented low accuracy rate, none of the metrics detect more than 53% of natural landscapes. Due to face recognition was not done in this first phase of the project, it is possible that some image pixels around faces could have been interpreted as landscapes. Also, image lighting could be a noise factor that disturbs the outlier identification.

Distances to mean and median were similar, although about 10% of images were outliers, this suggests that it is necessary to do some type of preprocessing in images to improve estimations.

To improve estimations, in next phase, other distances will be calculated, such as Mahalanobis, for which it is necessary to reduce dimensionality of data through principal component analysis (PCA) to estimate the inverse of the covariance matrix or precision matrix.

In general, all metrics behaved in a similar way on detecting true positives and negatives. When detecting false positives and negatives, outlier detection was less consistent compared to true positive and negatives detection.

The worst norm was the Chebyshev one (p = inf) with nule true outlier detection.

The Eigenfaces method allowed to reduce the original dimensionality of the problem without much loss of information. With only 305 (out of 36.000) components, 97.59% of variability was captured.

Eigenfaces allowed to identify in a successfully way whether an image is a face or not, by projecting that image to the face space.

By including the images of Malestaff to the images of Males, a little more noise was introduced to the images of the men, the results show that this noise did not affect the preitions of the model, on the contrary, the model learned to better recognize the faces of men

Using methods to reduce dimensionality allowed us to define an embedding to the manifold where the eigenface subspace was defined. Various methods were shown according to different criteria on how the projection of the data to a lower dimensionality space ought to be performed. We could see that projecting the data in such a way as to maximize the variance per cluster, while minimizing the variance in-between classes, allowed us a good generalization and an excellent visualization tool in order to manually detect the proportion of the data. We saw that using an embedding such that it contains a similar probability distribution from the manifold of eigenfaces will not produce a space with practical applications, given that the dominance of the male data-points compared to the other classes will absorb and therefore skew any bi-dimensional embeddings that try to project the original manifold unto a new space.

PCA allowed us to perform particularly well on linear classification models, which allow us to reduce the classification of images to a linear problem by projecting them to a vector space generated by the weights of the eigenfaces.

References

[1] Jain AK. Data clustering: 50 years beyond K-means. *Pattern Recognit Lett*, 31(8):651–66, 2010.

[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection). *IEEE Xplore*, 19(7):711–20, 1997.

[3] K. Etemad and R. Chellappa. Discriminant analysis for recognition of human face images). *J. Opt. Soc. Am. A*, 14(8):1724–33, 1997.

[4] Sturdivant R Hosmer D., Lemeshow S. Logistic regression models for multinomial and ordinal outcomes. In *Applied Logistic Regression*, pages 269–310, Hoboken, New Jersey, 2013.

[5] Sturdivant R Hosmer D., Lemeshow S. Discrimination and classification. In *Applied Multivariate Statistical Analysis RichardJohnson DeanWichern*, pages 575–670, United States of America, 2014.

[6] Sima L Karl W. Principal components analysis. In *Applied Multivariate Statistical Analysis*, pages 319–57, Berlin, Germany, 2014.

[7] Xu Z. Xiuming M Lan G, Wenya L. Improved PCA + LDA Applies to Gastric Cancer Image Classification Process. *Physics Procedia*, 24:1689–95, 2012.

[8] L.v.d. Maaten and G. Hinton. *Visualizing data using t-SNE*. Journal of Machine Learning Research, Vol 9(Nov), pp. 2579—2605., 2008.

[9] M. Macho. *Topología de Espacios Métricos*. 2010. Retrieved from http://www.ehu.eus/ mtwmastm/TEM0910.pdf.

[10] Somers K Peres P., Jackson D. How many principal components? stopping rules for determining the number of non-trivial axes revisited.). *Computational Statistics Data Analysis*, 49(4):974–97, 2005.

[11] M. A. Turk and A. P. Pentland. *Face recognition using eigenfaces*. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings, pp.586-591, 1991. https://doi.org/10.1109/CVPR.1991.139758.

[12] W. Zhao, R. Chellappa, P. J. Phillips, and A Rosenfeld. Face Recognition: A Literature Survey.). *ACM Comput. Surv.*, 35(4):399–458, 2003.