

Comparison of the PCA and FLD Approaches in Glial Tumors Classification Systems

Miroslav PETROV, Juliana DOCHKOVA-TODOROVA

Faculty of Mathematics and Informatics, St. Cyril and St. Methodius University of Veliko Tarnovo, Veliko Tarnovo, Bulgaria

`m.petrov@ts.uni-vt.bg`, `doskova@ts.uni-vt.bg`

ORCID: 0000-0002-5137-6812, ORCID: 0009-0002-2346-3073

Abstract. Computer-aided diagnosis (CAD) systems based on machine learning methods are an important component of medical practice. Principal Component Analysis (PCA) and Fisher Linear Discriminant (FLD) are among the main linear methods for feature extraction and reduction in recognition tasks. In this regard, a comparative analysis of the efficiency of the systems constructed by PCA and FLD algorithms, respectively, is carried out. As a continuation of our previous research, a classification system model for Magnetic Resonance Images (MRIs) of the brain is built using FLD. By analogy, the built CAD system detects the presence of a glial tumor, with a subsequent two-level and three-level gradation according to the degree of malignancy of the tumor. The input images are again not pre-processed and additional wavelet features are used: the normalized energy of the subimages and their non-normalized Shannon entropy. The comparative analysis of the pair of CAD systems is carried out through the quality measures: F1-score and the Matthews Correlation Coefficient. The validation of the obtained results is based on the diagnosis by three independent radiologists.

Keywords: CAD system, PCA, FLD, MRIs, wavelet transform, minimum distance classifier

1. Introduction

The high resolution of the images obtained by the magnetic resonance technology makes them preferable for medical imaging of the internal organs of the human body. But the vast number of images in medical databases greatly complicates the direct diagnostic activity of radiologists and leads to errors in analysing the specific data. Therefore, the development of computer-aided diagnosis (CAD) systems is an up-to-date task in the field of medical imaging in the relevant diagnostic task.

The main stages of the operation of such systems are:

- pre-processing of the image (noise removal, segmentation, etc.);
- extraction of image characteristics, as well as their eventual reduction;
- training through marked images;
- grouping using a specific classifier.

One of the most important stages in pattern recognition tasks, and therefore in CAD systems as well, is the extraction and reduction of identity features. The transition from the original sample space to a smaller dimensional space is related to the accuracy and

separability of the data. Fisher Linear Discriminant (FLD) and Principal Component Analysis (PCA) are the two main Machine Learning (ML) methods for reducing this dimensionality by linear projection. The goal of PCA is to find the most accurate representation of the data by minimizing the total projection error. The corresponding projection subspace is defined by the eigenvectors of the scattering matrix of the samples. FLD is a classification method indicating the design direction that maximizes the separability of the sample classes and minimizes the internal scatter within the classes, more precisely, it maximizes their ratio (Fukunaga, 1990). The analysis and comparison of their work efficiency with a certain database is a correctly set task. This problem has been mainly addressed in the framework of people identification. Studies whose experimental results show the better performance of FLD recognition algorithms predominate (Belhumeur et al., 1997). However, in (Martinez and Kak, 2001) experiment results are presented showing the superiority of PCA over FLD approaches. This is observed when the classes of the sample contain a small number of representatives or the sample is obtained by uneven selection. In (Eleyan and Demirel, 2006), two face recognition systems based on PCA and FLD algorithms using a neural network for classification are proposed. A comparative analysis of the performance of the proposed systems and the corresponding conventional systems based on the Euclidean classifier is made. In particular, the obtained results confirm the better performance of FLD systems in both classifiers. In (Eleyan, 2008), a comparative analysis is made between the PCA and FLD algorithms in the context of two approaches to solving the face recognition task: the classical approach, considering the whole face, and the method proposed there dividing the face into regions. In particular, the obtained results show that the FLD approach is more effective than the PCA one on large face databases. For example, based on the conducted t-test, within the second approach, this superiority is expressed by more than 16 % . Recently, interest in these problems has been confirmed in (Mostafa and Hossain, 2020), where the performance of PCA, FLD and simple projection approaches for face recognition is investigated. The conducted experiments determine the average efficiency of the three algorithms depending on the dimension of the corresponding projection space. In particular, the better performance of PCA is confirmed when the number of images is small or the sample is uneven with respect to the underlying distribution.

Based on previous research (Petrov, 2023), the task is set to conduct a comparative analysis of the performance of a pair of CAD systems based on PCA and FLD approaches, respectively, for the detection and classification of glial brain tumors.

In (Petrov, 2023), a model is proposed of a two-level CAD system for classifying MRIs of the brain. The classification system uses image descriptors extracted from the PCA projection space. In fulfillment of the given task, a model of an analogous system is built through the FLD approach. The comparative analysis is performed with the same samples of training and test images, keeping the quality measures: F1-score and the Matthews Correlation Coefficient (MCC).

In the rest of the section, some concepts and publications are introduced in order to clarify the content of the sections to come. Timely diagnosis of any tumour entity is extremely important for the outcome of the treatment of the disease. Brain tumours are grouped depending on their origin, place of occurrence, aggressiveness of development, etc. The subject of this work is the task of the classification of glial tumours arising in the auxiliary cells (glia) of the cerebral cortex. Depending on the type of its germ cell, the considered tumours are divided into Astrocytomas, Oligodendrogliomas and

Ependymomas. According to their malignancy and distribution in body tissues, the World Health Organization (WHO) groups them into four classes (Torp et al., 2022):

- Grade I are usually benign and surgically removable tumours;
- Grade II includes astrocytomas, oligodendrogliomas, and oligoastrocytoma (mixed cell type);
- Grade III comprises of anaplastic astrocytomas, anaplastic oligodendrogliomas, and anaplastic oligoastrocytoma;
- Grade IV contains the most aggressive glial tumour called glioblastoma multiforme.

The tumours of the first two classes are defined as low-grade gliomas (LGG) and those from the last two classes are high-grade gliomas (HGG).

There are numerous publications that have proposed various variants of CAD system architectures. In connection with this work, we will give a brief review of two recent overview studies that present sufficiently the engineering methodologies for brain tumour diagnosis.

In (Toufiq et al., 2021), a systematic study of brain tumour classification systems presented in recent years is conducted. The main components of the CAD system are described, as well as the techniques that accompany them. The used classifiers are examined in detail, and a comparative analysis is made between the supervised and unsupervised clustering methods. The review of the brain tumour classification systems in use contains 79 sources. The second publication is (Kaifi, 2023), in which the types of brain tumours and the ways of their detection through imaging methods are presented. An in-depth analysis of the software used in CAD systems is done. A number of brain tumour segmentation and classification methods using the techniques of machine learning and deep learning is reviewed. The corresponding results for the obtained accuracy of these methods are presented, as well as the medical bases used. The literature review contains 127 sources and presents the latest achievements in the field under consideration. In addition, let us note the existence of some CAD systems for the classification of brain tumors, in which the preprocessing step is not performed (Sarhan et al., 2020; Petrov, 2023).

The rest of the document is organized as follows: in the next section, the main steps of Fisher's linear discriminant analysis for extracting the classification features are given; in Section 3, the problem between sample size and data dimensionality for clustering is discussed; the methodology of the proposed system is discussed in the fourth part; and the results of the conducted experiments, their evaluation and the announced comparative analysis are the subject of Section 5. The paper ends with some concluding remarks.

2. Fisher's Linear Discriminant

We will now briefly discuss FLD as a supervised feature extraction method in the projection space. The considered MRIs can be represented as a one-dimensional vector of its pixels by sequentially connecting the rows (columns) of the $N \times N$ matrix of these pixels, as $x_i = [p_1, p_2, \dots, p_d]^T$, where $d = N^2$. Then the training sample containing n images from the K class can be written in the form $X = \{x_i, l_i\}_{i=1}^n$, where $x_i \in R^d$, and the labels $l_i \in \{1, 2, \dots, K\}$. If $X_k = \{x_i | 1 \leq i \leq n_k, l_i = k\}$, then $X = \bigcup_{k=1}^K X_k$ and

$n = \sum_{k=1}^K n_k$, where $n_k = |X_k|$. To formulate Fisher's criterion (Fukunaga, 1990), it is necessary to define the following two matrices: the within-class scatter matrix –

$$S_w = \sum_{k=1}^K \sum_{x \in X_k} (x - \mu_k)(x - \mu_k)^T; \quad (1)$$

and the between-class scatter matrix –

$$S_b = \sum_{k=1}^K n_k (\mu_k - \mu)(\mu_k - \mu)^T, \quad (2)$$

where μ_k and μ are respectively the mean values of the k -th grade data and the entire training sample.

Fisher's criterion gives the optimal direction of projecting the original features in a low-dimensional subspace in which the between-class scatter is as large as possible and the within-class scatter is the smallest possible. If the subspace for the FLD is determined by the set of vectors

$$W = [w_1, w_2, \dots, w_p] \in R_1^{d \times p}, \quad p \leq \min(d, K-1), \quad (3)$$

then the solution is obtained by maximizing the function

$$J(W) = \frac{\det(W^T S_b W)}{\det(W^T S_w W)}. \quad (4)$$

The matrix (3) is constructed from the generalized eigenvectors w by (S_b, S_w) , corresponding to the generalized eigenvalues $\lambda = \frac{w^T S_b w}{w^T S_w w}$, i.e.

$$S_b w = \lambda S_w w. \quad (5)$$

In the case where matrix S_w is invertible, equation (5) can be written as a standard equation for finding eigenvectors and eigenvalues of matrix $S_w^{-1} S_b$ –

$$S_w^{-1} S_b w = \lambda w. \quad (6)$$

Under binary classification ($K = 2$), the optimal design direction can be obtained directly from equation (6') –

$$w = S_w^{-1} (\mu_1 - \mu_2). \quad (6')$$

The representation of the original data X_k in the space R^p generated by the vectors $\{w_1, \dots, w_p\}$, is given by the formula (7) –

$$Y_k = W^T X_k, \quad k = 1, \dots, K. \quad (7)$$

Next, each test image x_i needs to be projected in an analogous way into space R^p . The distribution of x_i is based on the Minimum Distance Classifier (MDC) by assigning it the label l_k , where

$$k^* = \arg \min_k \left\| w^T (\mu_k - x_t) \right\|_{\mathbb{R}^p}. \quad (8)$$

3. The Small Sample Size Problem

The problem of the small sample size (SSS) is a major challenge when using FLD. If the dimensionality of the original data exceeds their number ($d > n$), then the within-class scatter matrix is singular – $\det(S_w) = 0$. Numerous methods have been developed to overcome the SSS problem, which appears in tasks from various fields, such as face identification, text recognition, bioinformatics, seismology, etc. A detailed overview of these methods can be found, for example, in (Sharma and Paliwal, 2015). Two such methods, which are used in this work, are presented.

3.1. The Moore-Penrose pseudoinverse matrix

A brief description of the Moore-Penrose (MP) pseudo-inverse for the matrix $S_w \in \mathbb{R}_+^{d \times d}$ is given in (Wu, 2017). For this purpose, the spectral decomposition (diagonalization) $S_w = U \Lambda U^T$ is used, where Λ is a diagonal matrix containing the eigenvalues of S_w , and the columns of the orthogonal matrix U contain their respective eigenvectors. Let $\Lambda = \text{diag}([\lambda_1, \lambda_2, \dots, \lambda_d]^T)$, then its MP pseudo-inverse is set as $\Lambda^+ = \text{diag}([\lambda_1^+, \lambda_2^+, \dots, \lambda_d^+]^T)$, where

$$\lambda_i^+ = \begin{cases} 0 & \text{if } \lambda_i = 0 \\ \lambda_i^{-1} & \text{otherwise} \end{cases}, \quad (9)$$

and the MP pseudo-inverse of S_w is $S_w^+ = U \Lambda^+ U^T$. It should be noted that, if S_w is not singular, then $S_w^+ = S_w^{-1}$.

3.2. Robust FLD Model

In (Deng et al., 2006), the Robust Fisher Linear Discriminant Analysis (RFLDA) method is proposed for the case of a singular (or close to singular) matrix S_w . Again, the spectral decomposition of the within-class scattering matrix is considered, and its eigenvalues are sorted – $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$, as well as their corresponding eigenvectors. In this case, the last eigenvalues of the matrix can compromise the results of the discriminant analysis, so in RFLDA they are replaced by a certain unified value. Based on statistical analysis, the number d^* of the main eigenvalues is determined, and the remaining $(d - d^*)$ are either very small or equal to zero. After conducting experiments, the authors propose the determination of d^* by minimizing the function

$$E(d^*) = \left(\sum_{i=1}^{d^*} \lambda_i \right) \left(\sum_{j=1}^d \lambda_j \right)^{-1}, \quad (10)$$

so that its values remain within the interval $(0.9, 0.99)$. The remaining $(d - d^*)$ eigenvalues are then replaced by $\lambda^* = \frac{1}{d - d^*} \sum_{j=d^*+1}^d \lambda_j$. Thus, the matrix S_w is evaluated by

$$S_w^* = U \Lambda^* U^T, \quad (11)$$

where $\Lambda^* = \text{diag}([\lambda_1, \dots, \lambda_{d^*}, \lambda^*, \dots, \lambda^*]^T)$, and the design directions are set by the generalized eigenvectors of (S_b, S_w^*) .

4. Methodology

In this part, the methodology of the proposed system is explained. Its design is presented in Fig.1. At first, the brain MRIs selected in the training sample are divided into normal and abnormal, according to the absence or presence of a glial tumour. In the next step, the abnormal images are grouped into two or three classes, depending on the extent of the glial tumour present.

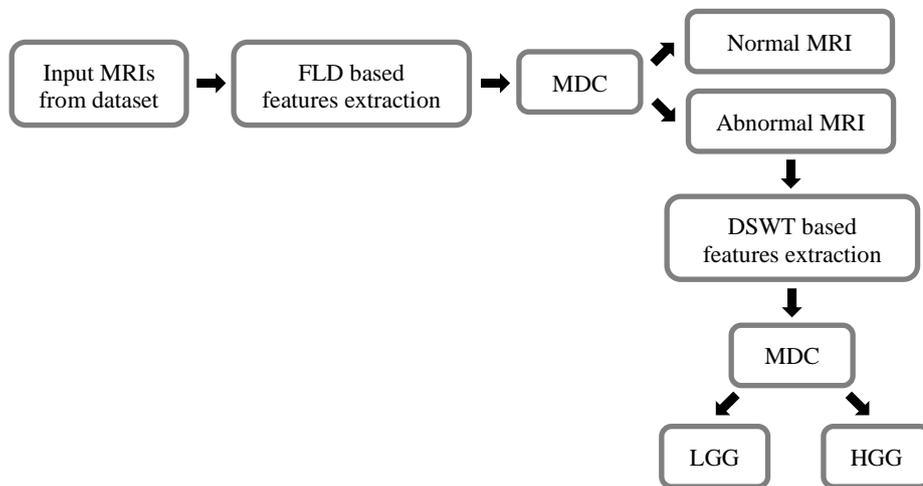


Figure 1. Block diagram of the proposed system.

4.1. Tumour Detection Stage

Let the training sample of MRIs that enter the input of the CAD system be $X = \bigcup_{k=1}^2 X_k$, where $X_k = \{x_i | 1 \leq i \leq n_k, l_i = k\}$ and $n_k = |X_k|$. The required image features are extracted using FLD. Projecting the original data X into space R is performed according to equation (7). For the classification stage, it is necessary to determine the centroids $w^T \mu_k$ of each of the two classes $X_k, k = 1, 2$. Next, the image received for classification x_i is projected in an analogous way in R and its belonging to one of the two classes is determined by MDC and the corresponding weighted metric.

4.2. Tumour Classification Stage

Let X_a be the array of MRIs with malignant entities obtained at the first stage of the operation of the CAD system. Then the training sample for the second classifications will have the type $X_a = \bigcup_{k=1}^K X_{a_k}$, where $K = 2 \vee K = 3$ and $X_{a_k} = \{x_{a_i} | 1 \leq i \leq n_k, l_i = k\}$, $n_k = |X_{a_k}|, k = 1, \dots, K$. In addition to FLD, the Discrete Stationary Wavelet Transform (DSWT) is also used to extract the required image features (Mallat, 1998). The detailed wavelet coefficients are indicators of the local peculiarities of the signals, therefore only the three sub-bands LH, HL and HH are considered in this work (see Fig.2).

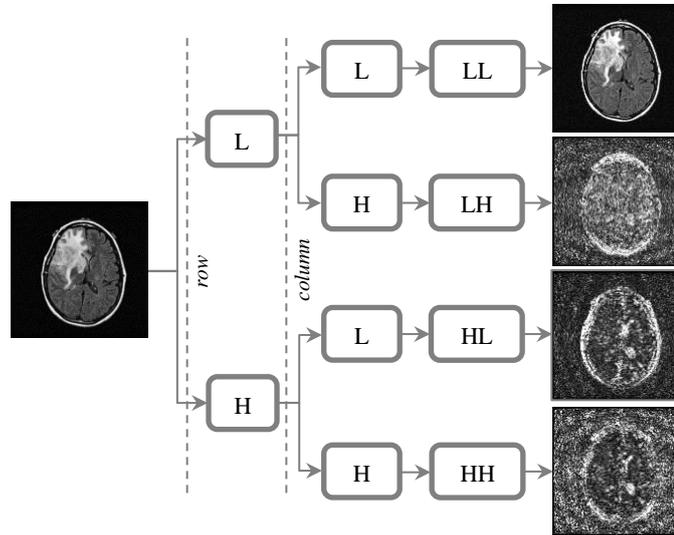


Figure 2. MRI decomposition by DSWT.

The extracted energy and entropy features are added to those obtained by FLD to construct the image descriptor vector. For the second part of this vector, the normalized energy of the sub-image $P_{r_s} - E_{r_s} = d^{-2} \sum_{q,m} D_{r_s}^2(q,m)$, is used, where the corresponding wavelet coefficients are indicated by D_{r_s} . The resulting energy characteristics are $\{E_{r_s}\}$, $r = 1, 2, 3$ and $s = 1, \dots, L$, where L is the maximum decomposition level. The third part of the descriptor contains the entropy features $\{H_{r_s}\}$, $r = 1, 2, 3$, $s = 1, \dots, L$, where $H_{r_s} = -\sum_{q,m} D_{r_s}^2(q,m) \cdot \log D_{r_s}^2(q,m)$ is the unnormalized Shannon entropy of the subimage P_{r_s} .

Similarly, the feature vector of a random test image x_t , is obtained, and its belonging to the corresponding class is again determined by MDC. The metric used in formula (8) is the Normalized Euclidean Distance (NED).

5. Comparative Analysis and Discussions

In the proposed system, supervised learning is carried out, allowing to take into account the possible errors of the classifier. The performance of the CAD system is evaluated by the measures F1-score and MCC, which are predetermined by the goals set in this work. The first one is the harmonic mean of precision and recall, proportional to the quality of the classifier. The second measure reflects the relationship between the observed and predicted data using the entire confusion matrix and is unaffected by the dimensionality of the classes. Besides, it should be noted that, when conducting the comparative analysis of the performance of the pair of CAD systems, the values of the quality measures were obtained using the expert opinion of three radiologists.

5.1. Accuracy in Tumour Detection

The presented results needed to perform the requested comparative analysis were obtained with the same collection of 340 brain MRIs (Petrov, 2023), 200 of which represent the training sample. T1-weighted (T1W), T2-weighted (T2W) and T2-sensitive (T2F) images were used, each of 256×256 pixels in size in DICOM format. The training samples were labelled by three independent experts, the sample itself being balanced. The data were obtained from the following publicly available medical databases (Pedano et al., 2016; Scarpace et al., 2019; Erickson et al., 2017) and from the Imaging Department of Dr Stefan Cherkozov Hospital of Veliko Tarnovo. Fig. 3 shows MRIs containing glial tumours of the following types: astrocytoma, oligodendroglioma and glioblastoma.

When using the F1-score and MCC measures to evaluate binary classifications and their corresponding confusion matrices, their sensitivity to the balance of the dataset should be considered. MCC uses all elements of the confusion matrix, making it robust to unbalanced samples.

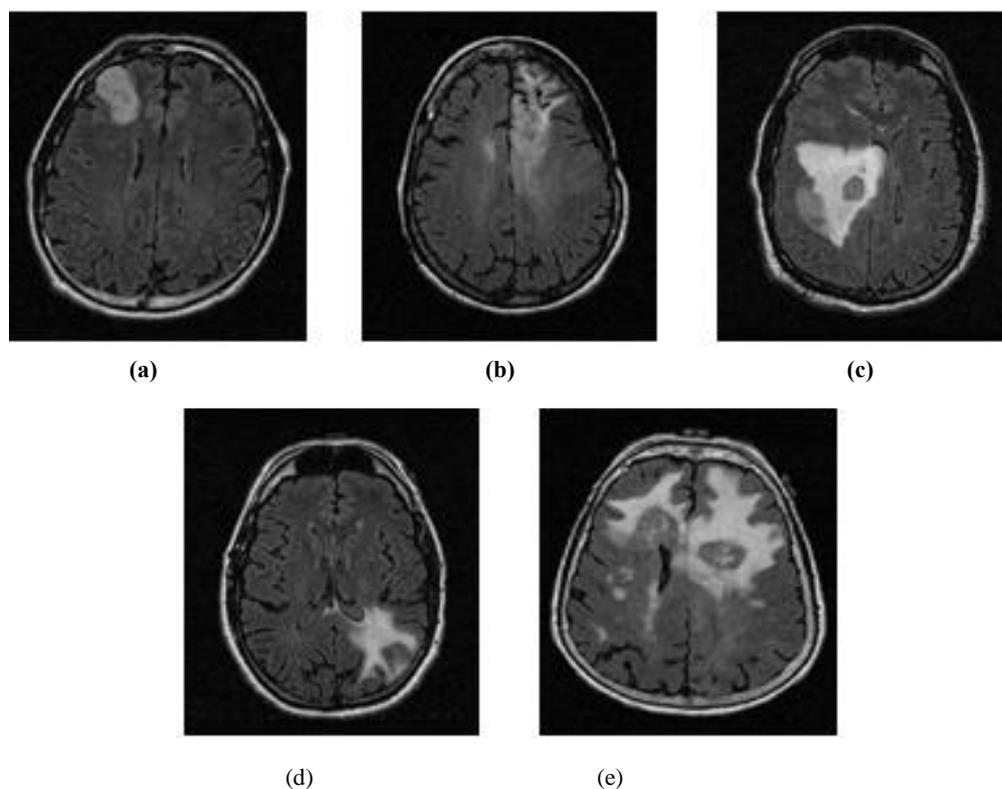


Figure 3. Glial tumors: a) low-grade astrocytoma; b) low-grade oligodendroglioma; c) high-grade astrocytoma; d) high-grade oligodendroglioma; e) glioblastoma.

The results of the comparative analysis in the tumour detection stage are presented in Table 1. Its data show that the FLD methods outperform the PCA method by an average of 3.23% and 11% , respectively, for F1-score and MCC. Furthermore, the RFLDA method shows better classification than the FLD–MP method in both evaluation metrics.

Table 1. Values of the performance measures of the classifier for tumour detection

Methods	Performance metrics	
	MCC	F1-score
PCA (Petrov, 2023)	0.82	0.93
Proposed FLD - MP	0.89	0.95
Proposed RFLDA	0.93	0.97

5.2. Accuracy in Tumor Classification

In this part, the performance of the CAD system in categorizing the glial tumour grade is investigated. The training sample is obtained from the original one, keeping the malignant tumour MRIs. The test set contains 100 slices of low-grade and high-grade gliomas.

The results of the comparative analysis for the second classification are presented in Tables 2 and 3. The data show that for these classifiers too, the FLD methods perform better, on average by 3.8% (F1-score) and by 3.2% (MCC). Again, the better performance of the second method is confirmed. The data from the additionally performed three-level classification show the superiority of the RFLDA method over the FLD-MP method, with

Table 2. Indicators in the two-stage classification of the tumor

Methods	Performance metrics	
	MCC	F1-score
PCA (Petrov, 2023)	0.78	0.92
Proposed FLD - MP - MP	0.8	0.95
Proposed RFLDA	0.81	0.96

Table 3. Indicators in the three-stage classification of the tumor

Methods	Accuracy in tumour grading [%]		
	II grade	III grade	IV grade
Proposed FLD - MP	88	79	91
Proposed RFLDA	91	85	94

the corresponding percentage expression being 3.4% (for II grade), 7.6% (for III grade) and 3.3% (for IV grade).

5.3. Discussions

The objective set in Section 1 and the descriptors used justify the comparative analysis between the PCA and FLD methods in the tasks of glial tumours detection and clustering. These are two projection methods of ML to reduce the dimensionality of the original data space. PCA maximizes the accuracy of the samples in the projection space while preserving the variance of the original data. FLD is a supervised classification method that tries as much as possible to preserve the necessary information to separate the classes.

The basis for the comparative analysis is that in the present work both methods are used to determine the centroids of the classes. From the data presented in the above tables, it can be seen that the classifiers based on FLD methods are more efficient than those using PCA. The values in each of these tables are obtained by averaging the results of twenty tests with data randomly generated from the respective test samples. From a computational

point of view, the mathematical implementation of the FLD algorithm requires a significant amount of RAM, even for images with a resolution of 256×256 pixels. An additional difficulty is the singularity of the scattering matrix S_w due to the SSS problem.

6. Concluding Remarks

Because FLD is a direct class separation method and PCA is a method representing data as a whole, the former is usually assumed to be superior in recognition tasks. This hypothesis is also confirmed by the comparative analysis carried out in the previous section. But as it was stated in Section 1 there are cases when PCA outperforms FLD in some tasks. For example, if there is a small-size training sample (unrepresentative) or if it is unevenly distributed across classes (unbalanced). The proposed analysis can be extended by examining the performance of the methods: as a function of the sample size; at different class distributions or at additional wavelet features obtained with other multiscale transformations.

References

- Belhumeur, P., Hespanha, J., Kriegman, D. Using discriminant eigenfeatures for image retrieval. *PAMI*, **19**(7), 711–720, 1997.
- Deng, W., Hu, J., Guo, J. (2006). Robust Fisher Linear Discriminant Model for Dimensionality Reduction, *18th International Conference on Pattern Recognition (ICPR'06)* **4**, 699-702, doi: 10.1109/ICPR.2006.211.
- Eleyan, A., Demirel, H. (2006). PCA and LDA Based Face Recognition Using Feedforward Neural Network Classifier. In: Gunsel, B., Jain, A.K., Tekalp, A.M., Sankur, B. (eds) *Multimedia Content Representation, Classification and Security. MRCS 2006. Lecture Notes in Computer Science*, **4105**. Springer, Berlin, Heidelberg, https://doi.org/10.1007/11848035_28.
- Eleyan, M. (2008). PCA and LDA based face recognition using region-division with majority voting. <http://library.neu.edu.tr/Neutez/4713191966.pdf>.
- Erickson, B., Akkus, Z., Sedlar, J., Korfiatis, P. (2017). Data from LGG-1p19qDeletion(version 2) [data set]. The cancer imaging archive, <https://doi.org/10.7937/K9/TCIA.2017.DWEHTZ9V>.
- Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition*, Academic Press, San Diego, California, USA.
- Kaifi R. (2023). A Review of Recent Advances in Brain Tumor Diagnosis Based on AI-Based Classification. *Diagnostics (Basel)* **13** (18):3007, 1-32, doi: 10.3390/diagnostics13183007.
- Mallat, S. (1998). *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, California, USA.
- Martinez, A. M., Kak, A. C. (2001). PCA versus LDA, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23** (2), 228-233, doi: 10.1109/34.908974.
- Mostafa, F. B., Hossain, S. (2020). Revisiting the performance of PCA versus FDA versus Simple Projection for Image Recognition. *Global Journal of Engineering and Technology Advances* **8** (1), 084-095.
- Pedano, N., Flanders, A. E., Scarpace, L., Mikkelsen, T., Eschbacher, J. M., Hermes et al., (2016). The cancer Genome Atlas Low Grade Glioma Collection (TCGA-LGG) (version 3) [data set]. The cancer imaging archive, <https://doi.org/10.7937/K9/TCIA.2016.L4LTD3TK>.
- Petrov, M., (2023). A model of a two-stage classification system for glial tumors in magnetic resonance imaging, *2023 International Conference Automatics and Informatics (ICAI)*, 418 - 422, doi: 10.1109/ICAI58806.2023.10339009. ISBN 979-8-3503-1291-1.

- Sarhan, A. M. (2020). Brain Tumor Classification in Magnetic Resonance Images Using Deep Learning and Wavelet Transform, *J. Biomedical Science and Engineering*, **13** (6), 102-112.
- Scarpace, L., Flanders, A. E., Jain, R., Mikkelsen, T., Andrews, D. W. (2019). Data from REMBRANDT[dataset]. The cancer imaging archive, <https://doi.org/10.7937/K9/TCIA.2015.588OZUZB>.
- Sharma, A., Paliwal, K. (2015). Linear discriminant analysis for the small sample size problem: an overview. *Int. J. Mach. Learn. & Cyber* **6**, 443–454, <https://doi.org/10.1007/s13042-013-0226-9>.
- Torp, S. H., Solheim, O., Skjulsvik, A. J. (2022). The WHO 2021 Classification of Central Nervous System tumours: a practical update on what neurosurgeons need to know—a minireview, *Acta Neurochir* **164**, 2453–2464, <https://doi.org/10.1007/s00701-022-05301-y>.
- Toufiq, D. M., Ali Makki Sagheer A. M., Veisi H. (2021). A Review on Brain Tumor Classification in MRI Images, *Turkish Journal of Computer and Mathematics Education* **12** (14), 1958- 1969.
- Wu, H. (2017). Fisher’s linear discriminant. <https://cs.nju.edu.cn/wujx/paper/FLD.pdf>.

Received August 13, 2024, revised January 20, 2025, accepted January 23, 2025