

Keratoconus Classification Using Multimodal Imaging Strategy

AATILA Mustapha¹, KARTIT Ali²

¹Cadi Ayyad University, UCA, Faculty of science Semlalia, Laboratory of Computer Science and Smart Systems (LISI), Bd. Prince My Abdellah, B.P. 2390, 40000 Marrakesh, Morocco.

²LTI laboratory, ENSA, Chouaib Doukkali University, El Jadida 1166, Morocco.
E-mail: m.aatila@uca.ac.ma, alikartit@gmail.com

Article history

Received Jun 9, 2020
Revised Nov 20, 2020
Accepted Jan 11, 2021
Published Jan 28, 2021

ABSTRACT

Abstract: Data fusion improves the accuracy and robustness of diagnostic models by combining different types of information. This study presents a multimodal framework for keratoconus classification. It uses numeric and textual features from Pentacam reports, extracted with OCR. These are combined with corneal topographic images processed by a dual-branch deep neural network. The method was tested on 2,924 labeled Pentacam scans. Of these, 1,900 were used for training and 1,024 for testing. Scans were labeled as normal, suspicious, or keratoconus. Results show that combining image and text features improves classification. Deep learning accuracy rose from 96.78% to 98.34%. SVM improved from 93.35% to 95.60%. LDA increased from 92.85% to 94.80%, and KNN from 90.50% to 93.94%. These gains, up to 1.56% for deep learning and 3.44% for KNN, show the value of multimodal data for more accurate keratoconus diagnosis.

Keywords: Keratoconus Classification, Multimodal data fusion, Corneal topography, Machine learning, Deep Learning.

I. INTRODUCTION

Currently, the diagnosis of keratoconus is essentially based on clinical examination and corneal imaging. Corneal imaging techniques, such as Schimpf photography and optical coherence tomography (OCT), can provide detailed information about the cornea and aid in the diagnosis of keratoconus. However, interpretation of corneal images can be subjective and time-consuming. Therefore, there is a need for automated and objective methods for the diagnosis of keratoconus.

With the growing popularity of ML and its accompanying implementations, there are increasing calls to create deployable and relevant decision support tools for the detection and classification of multiple diseases, including keratoconus. Exploring how decisions are made by humans, they rarely rely on a single data point or source of information. This diversity of information sources will give rise to information fusion in the context of multimodal machine learning. For example, a clinical diagnosis made by a physician is typically based on a constellation of labs, signs and symptoms, imaging that, when contextualized by each other, leads to a decision/classification.

Data fusion is a technique of combining data from different sources to create analytical data sets that can be used to make more precise decisions addressing specific

issues by applying in-depth analysis [1].

Data fusion allows various types of data from different sources to be integrated into an analysis, this technique helps to obtain higher quality actionable insights for the analysis process [2]. However, merging two or more data sets often reveals relevant features that could not have been discovered without merging these data sets. These revealed characteristics can offer new insights leading to more informed decisions. This is a very powerful technique for merging multiple datasets to produce an accurate classification compared to classifications using individual datasets.

Three different categories of data fusion can be distinguished, each of these three categories depends in particular on the processing step at which the data fusion occurs. These three categories are:

- **Low-level fusion:** For low-level data fusion, multiple raw data sources provided directly as inputs are combined to produce new raw data. The goal is to produce merged data that is more informative and synthetic than the original individual entries [3];
- **Mid-level fusion:** For mid-level fusion, the fused features are first extracted from other input data before the fusion process. These features can be many things, for example a subset of variables, a

set of latent variables, or other features such as shape or position in an image [3];

- **High-level fusion:** In this type of fusion a supervised model is applied to each individual dataset. Subsequently, the decisions of the different models are combined to obtain higher prediction or classification accuracy [3].

In this paper, we propose a deep learning approach for keratoconus classification using feature fusion. The proposed approach combines several features extracted from corneal topographic images to improve the accuracy of keratoconus classification. The proposed method was evaluated on a dataset composed of 2924 corneal topographic images.

This work is promising for the development of automated keratoconus detection systems that can aid in the early diagnosis and management of this disease.

The remainder of this paper is organized as follows: Section 2 presents an overview on data fusion in keratoconus classification. Section 3 describes the adopted methodology of keratoconus classification. Section 4 reports and discusses the obtained results. Section 5 provides conclusions of the proposed approach.

II. RELATED WORKS

Various machine learning techniques have been used in the diagnosis and classification of diseases, including keratoconus. Data fusion is one of the most widely used ML techniques in the literature.

The authors of [4] proposed the LKG-Net system for the classification of keratoconus according to the 4 levels Normal, Mild, Moderate and Severe using data fusion. This system is based on a multi-level feature fusion module to merge data from upper and lower levels to obtain more abundant and efficient features. Evaluated on a total of 488 topographic images from 281 people, this model achieved 89.55% for weighted recall, 89.98% for weighted precision and 89.50% for weighted F1-score respectively.

In the study [5], the authors developed a system based on the Xception and InceptionResNetV2 architectures to extract features from three different corneal maps collected from 1371 eyes examined at an eye clinic in Egypt. These features were fused using Xception and InceptionResNetV2 to detect subclinical forms of keratoconus more accurately and robustly. The proposed system achieved an AUC of 0.99 and an accuracy range of 97% to 100% in distinguishing normal eyes from eyes with subclinical keratoconus.

The authors of [6] proposed the KerNet system for detecting keratoconus and subclinical keratoconus based on the raw data of the Pentacam HR system. This system is based on the fusion of five digital matrices, corresponding

to the curvature of the front and rear surfaces, the elevation of the front and rear surfaces and the pachymetry of an eye. The results generated by this system achieved an accuracy of 94.74%, a recall of 93.71%, a precision of 94.10% and an F1-score of around 93.89%.

The keratoconus detection system proposed in [7] is based on the use of a camera of a smart device to capture photographed eye images of the anterior and lateral segments. Using a set of 280 images, the corneal area of the images was segmented in order to extract geometric features. The results obtained by this system showed that the fusion of all features was able to generate an accuracy of 96.05%, a sensitivity of 98.41% and a specificity of 93.65% with the Random Forest classifier.

The authors of [8] proposed a system for identifying the base curve in rigid gas permeable (RGP) lenses based on supervised image processing and classification of the four Pentacam refractive maps in the case of irregular astigmatism. Using a dataset of 247 four labeled Pentacam refractive maps, two novel feature extraction techniques, namely quantization-based radial-sector segmentation (QRSS) and deep convolutional neural networks, were used to extract the different characteristics. Feature fusion was applied and the RGP base curve was identified by the regression layer of a neural network. The results achieved a coefficient of determination of 0.9642 and a root mean square error of 0.0089.

III. METHODOLOGY

The proposed keratoconus classification method is based on the analysis of topographic images of patients. These corneal topographic images are composed of four topographic maps, namely corneal thickness, elevation back, elevation front and sagittal curvature. In addition to these maps, the images include annotations, measurements and text representing certain corneal characteristics.

A. Data collection

The dataset used in this study is composed of a total of 2924 images captured, anonymously, using a Pentacam camera. Each corneal topographic image represents the cornea of a different patient. The retained part of these images consists of the four maps (corneal thickness, sagittal curvature, rear and front elevations) in 1024x729 JPEG format as shown in the figure 1. The captured images were classified and labeled manually, by specialists, considering three corneal classes, which are : normal, suspicious and keratoconus corneal classes.

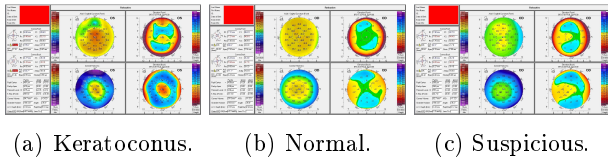


Figure 1: Topographies of the different corneal classes.

B. Proposed approach

The idea of the approach proposed in this chapter is based on recovering the textual values contained in these images and saving them in separate text files before cleaning these images, then combining these recovered values and the cleaned images during the classification phase in order to improve the predictive performance of the different classifiers used. Thus, the proposed architecture consists of two branches, the first branch for extracting textual features included in images, and the second branch for extracting features from cleaned images. The overview of the different steps of the methodology adopted in this study is illustrated in figure 2.

1. Data preprocessing

The data preprocessing step for diseases classification, including keratoconus, is very essential to ensure better accuracy of the model. Thus, corneal topographic images were preprocessed to correct distortions and to eliminate image noise from the entire data set. As shown in figure 1, topographic images contain, in addition to the topography of the corneas, digital measurements and textual annotations on the images.

Before proceeding with the classification of keratoconus, the next step is then to retrieve these textual measurements from the images and save them in .CSV files separately before excluding them from the images.

2. Textual features extraction

For text extraction, the adopted technique is based on optical character recognition (OCR) tool using python [9]. This tool allows reading plain text and tables from image and PDF files using an OCR engine and provides intelligent post-processing options to ensure results in the desired formats.

The textual features extraction branch used in this study is composed of the following layers:

- **An Embedding layer** : This embedding layer can be seen as transforming data from a higher dimensional space to a lower dimensional space, or it can be seen as a mapping of data from a space from lower dimension to higher dimensional space;
- **An LSTM layer** : with 100 units, this layer essentially implements an LSTM type recurrent model,

which is one of the variants of recurrent neural networks;

- **A dense layer** : (text_features) with 100 units;
- **A ReLU activation function.**

The extracted text values for each corneal class were stored in a separate .CSV file. Thus, three resulting csv files were created, namely suspect.csv, normal.csv and keratoconus.csv. To ensure better classification performance and improve the robustness of the model, images of poor quality were removed.

3. Text Removal

Figure 4 (a) represents an original corneal topography, including annotations and text values. Once the text values have been extracted from the images and saved separately, these images will be cleaned. For this purpose, these images were subjected in the next step to a process of eliminating these annotations by following a particular treatment detailed in the following section.

In order to exclude text that is potentially harmful to the learning process, these corneal topography images must be subjected to inpainting techniques and algorithms to be able to repair and restore them. Thus, the OpenCV tool was used for this purpose [10], the images are downloaded and then converted to grayscale, subsequently a rectangular kernel of (6x6) is built and a blackhat operation that allows finding dark regions on a light background is applied to these images to detect the textual values existing on the images.

The images resulting from this step are subsequently subjected to binary thresholding with a threshold of 10. Thus, all pixel values above the specified threshold 10 are set to a maximum value of 255 and all pixel values below 10 are set to 0.

The mask image indicates where the damage to exclude is located in the image. This image must have the same spatial dimensions (width and height) as the input image. Non-zero pixels correspond to areas that need to be painted (i.e. fixed), while zero pixels are considered normal and do not need to be painted. Figure 4 (b) below shows a mask image of the original image 4 (a).

Subsequently, the image inpainting technique based on Telea's fast walking method is applied [11], using the original images and the previously thresholded images as the inpainting mask. The return image is a restored image without text, as shown in the figure 4 (c), representing an example of a topographic image after extraction and elimination of text values. This process is applied to all images in the dataset.

Thus, the inputs of the future keratoconus classification model will be of two types: cleaned corneal topographic image data and textual data previously extracted from these same images before eliminating them.

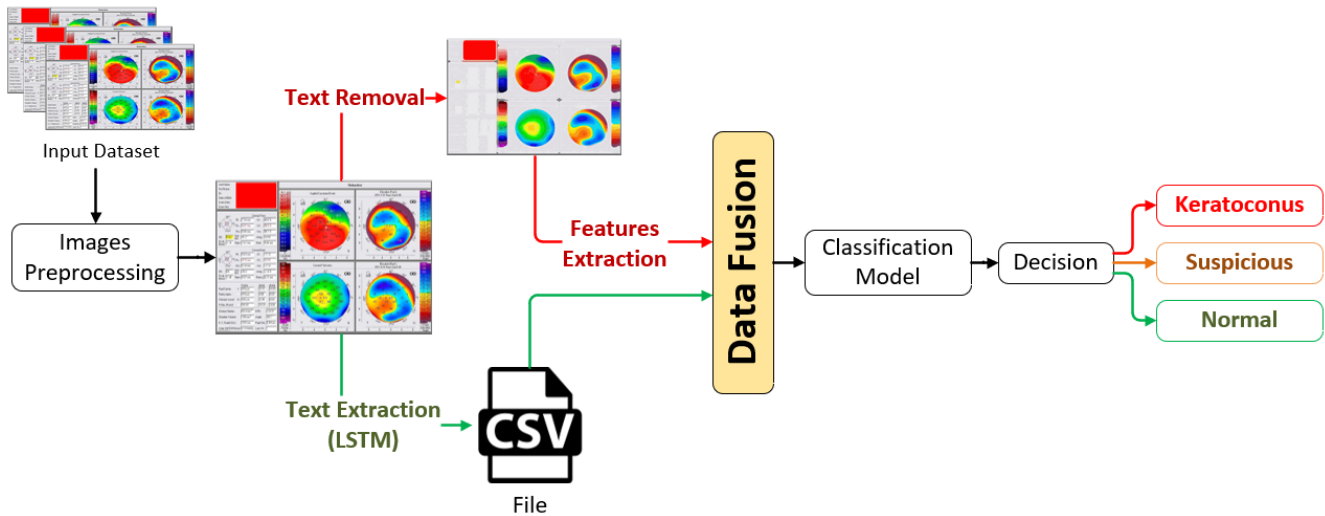


Figure 2: Adopted data fusion approach.

4. Features extraction from cleaned images

The image feature extraction branch starts with a convolutional layer (conv1_1) that has 32 filters with a kernel size of 3x3, followed by a ReLU activation. The output is then passed to a max-pooling layer (pool1_1) with a pool size of 2x2. This process is repeated with another convolutional layer (conv1_2) with 64 filters and a ReLU activation, followed by another max-pooling layer (pool1_2) with a pool size of 2x2. The output of this branch is then flattened.

The output of this branch is then merged with the result of the text extraction branch from images. Indeed, these two branches are then merged by concatenating the image features and the text features (merged_features). The output is then passed through two fully connected layers (fc1 and fc2) with 128 and 64 units respectively and a ReLU activation. The final output is a dense layer with 3 units and a softmax activation.

5. Data fusion (Images et text)

Feature fusion refers to the process of combining information from multiple sources, images and text, to improve the predictive performance of the learning model used for keratoconus classification. This could involve combining information from multiple modalities, such as combining information about the shape of the cornea and the size of certain features, or combining information from multiple layers of the deep learning model. By combining information from multiple sources, the model is able to make more accurate predictions and improve its classification accuracy. The article would have described the specific feature fusion techniques used and how they were implemented in the model.

In this chapter, the classification of keratoconus con-

sists of combining data from two different formats, images and text data to obtain more accurate results.

6. Classification model

Once the textual data and the corneal topographic image data were extracted and merged, by concatenating these data, the resulting dataset was used as input for the models responsible for classifying corneas into three different classes (normal, suspect and keratoconus). Four different models were adopted for the classification of keratoconus, these models were tested on the data before and after data fusion. These models are :

- Support Vector Machine (SVM);
- Linear Discriminant Analysis (LDA);
- K-Nearest Neighbors (KNN).
- A Deep learning (DL) model: This classification model consists of three different dense layers. The first fully connected layer (fc1) with a total of 128 units and a ReLU activation function. The second dense layer (fc2) is composed with 64 units and a ReLU activation function. The final classification of corneas is ensured by the third fully connected layer fc3 with 3 units and a Softmax activation function. To avoid the overfitting problem when training the adopted DL model, a Dropout of 0.2 is applied in order to ignore 20% of the nodes of the layers at random during training.

For the training and validation of the model used, the data are sent as and when. Indeed, a data normalization has been applied and uniform batches of a size of 32 elements are used for loading the data. This allows to

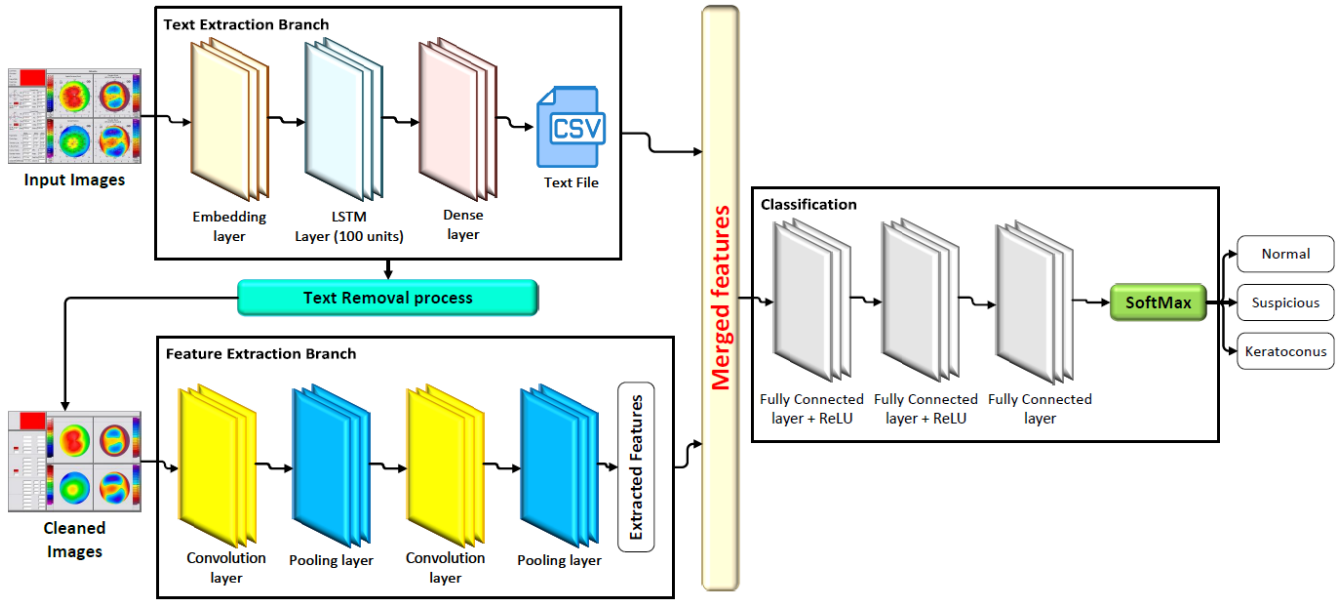


Figure 3: Text and features extraction process.

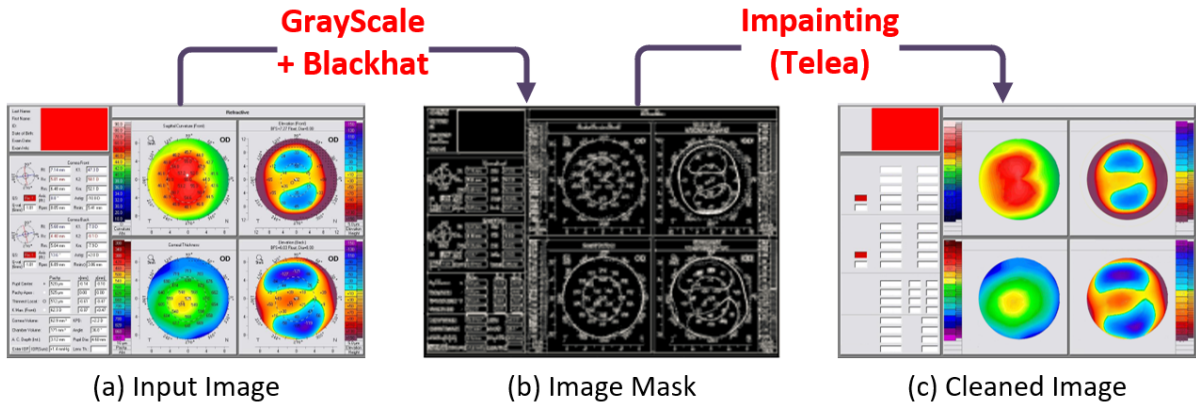


Figure 4: Text removal process.

process image data with normalized values between 0 and 1, and no longer on their entire RGB color scales, which extends from 0 to 255, which allows a better understanding of the classification model used for the discrimination of keratoconus.

For the different machine learning models SVM, LDA and KNN, the 10-fold cross-validation technique was used to avoid the overfitting problem.

7. Evaluation metrics of adopted models

To evaluate the classification performances of the models adopted in this proposed approach in an objective and complete manner, the evaluation metrics used are:

- Accuracy;
- Precision;

- Recall;
- F1-score;
- Confusion matrix.

These metrics are used to examine the ability of the models used to more precisely and accurately predict the different classes of patients' corneal topographic images.

IV. SIMMULATION RESULTS

A. Dataset description

The initial dataset used in this study consists of corneal topographic images in JPEG format, with a total of 2924 images representing the three corneal classes normal, suspect and keratoconus. Table 1 details the dataset used in this study.

Table 1: Description of training and testing data.

Classes	Training	Test	Total
keratoconus	221	120	341
Normal	1102	593	1695
Suspicious	577	311	888
Total of Images	1900	1024	2924
Percentage	65%	35%	100%

After the quality control step, the dataset was divided into two subsets. A training subset consisting of 65% of the images from the initial dataset, i.e. 1900 images. The second subset is the validation subset, consisting of 35% of the images from the initial dataset, i.e. 1024 images.

B. Description of the computer used

The simulation results are obtained using a computer equipped with an Intel(R) Core(TM) i5-6300U CPU @ 2.40GHz 2.40GHz, a RAM of 8.00 GB, the Windows 10 Professional operating system, tensorflow and the keras library in Python 3.7.4 on Jupyter notebooks.

C. Results obtained

To train and validate the adopted CNN model, on the two training and validation sets respectively, the optimizer used is adam with a learning rate of 0.001 and using batches of uniform images of 32 elements over 20 epochs. Figure 5 represents the evolution curve of the classification accuracy before fusion (A) and after fusion of the data (B) for the adopted DL model.

Figure 6 shows the classification performances by the different ML models (i.e. SVM, LDA and KNN) used for the classification of keratoconus before fusion (A), then after fusion (B) of the data. It should be recalled that a 10-fold cross-validation was used for the different models in order to avoid over-fitting problems.

Figures 5 and 6 clearly indicate that the predictive performance of the adopted DL and ML models have been significantly improved by using the data fusion technique. Indeed, concatenating the textual data extracted from the topographic images with the original images has resulted in good performance compared to the results obtained by using only the images during classification.

V. DISCUSSION

The aim of using the data fusion technique in this study is to improve the accuracy of keratoconus classification by combining the data representing the topographic images and the textual annotations and metrics present on these images. Thus, the metrics and textual annotations present on the different images were retrieved and saved

in separate files, then combined with the cleaned images for a more accurate classification of corneas.

Figure 5, representing the classification accuracy curve of the adopted DL model, clearly shows that the classification performance of the associated model, during the test phase, has been significantly improved after using data fusion. Indeed, the overall classification accuracy has increased from 96.78% for classification without fusion, to 98.34% for classification using data fusion, with an increase of around 1.56%. Table 5.2 describes in detail the classification performance of the DL model used before and after adopting the data fusion technique.

Figure 6 part(a) shows that the classification of keratoconus by the SVM, LDA and KNN models, using only topographic images generated good classification performances, with an accuracy of the order of 93.35%, 92.85% and 90.50% for the SVM, the LDA and the KNN respectively.

Figure 6 part(b), representing the classification performances of keratoconus by the same ML algorithms by applying the data fusion technique, shows a remarkable increase in the predictive capacities of the different models used. Indeed, the adoption of the data fusion technique allowed to reach a classification accuracy of the order of 95.60%, 94.80% and 93.94% for the SVM, the LDA and the KNN respectively.

Table 5.3 presents more details regarding the classification performances of the SVM, the LDA and the KNN models used without and with adoption of the data fusion technique.

VI. CONCLUSION

Fusing numeric and textual data from Pentacam reports with corneal topography images significantly improves keratoconus classification. A two-branch deep network was used to combine these modalities. On a dataset of 2,924 scans, accuracy increased across all models. Deep learning improved from 96.78% to 98.34%. SVM rose from 93.35% to 95.60%. LDA went from 92.85% to 94.80%. KNN increased from 90.50% to 93.94%. Combining different data types produces more accurate and reliable diagnostic models. This approach captures complementary disease markers and supports earlier, more confident clinical decisions.

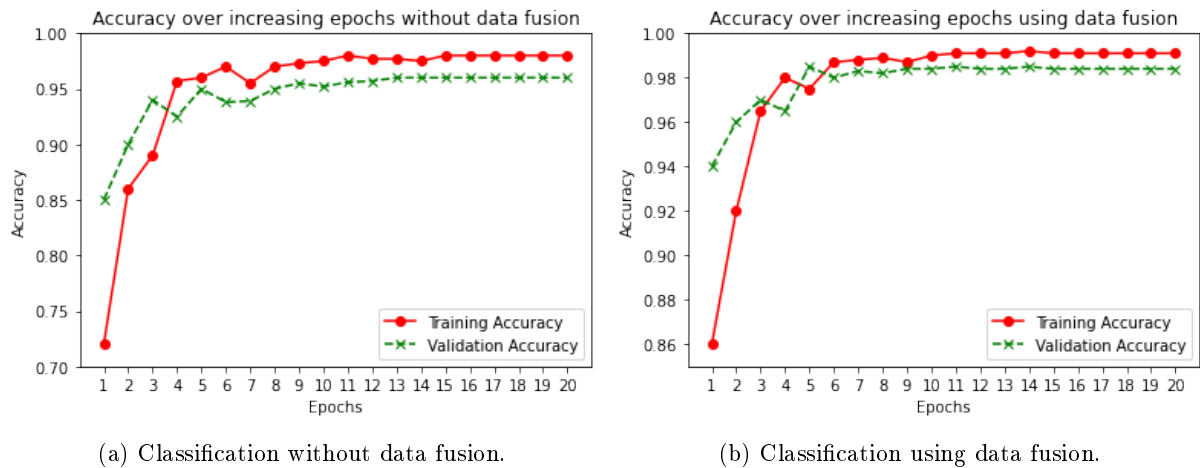


Figure 5: Keratoconus classification using DL model.

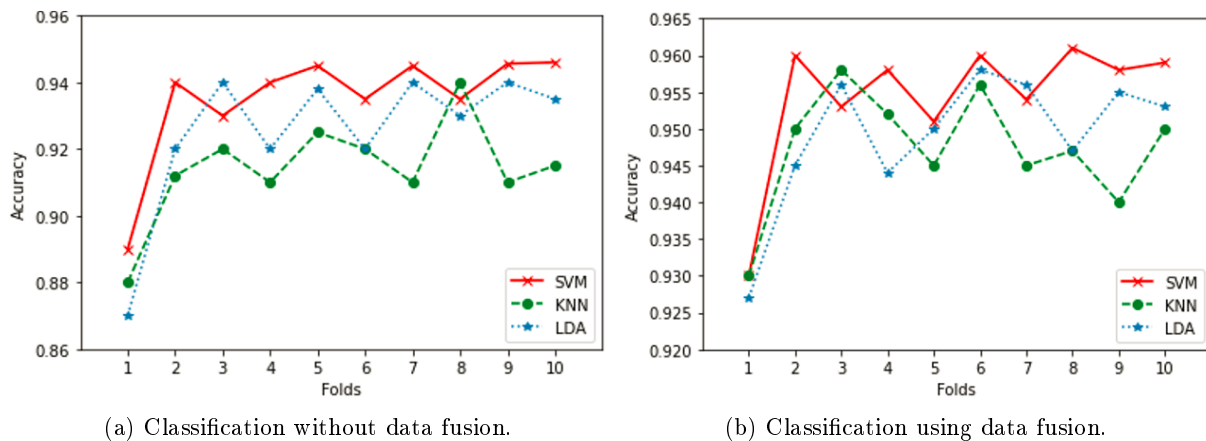


Figure 6: Keratoconus classification using different ML models.

REFERENCES

- [1] W. Chango, J. A. Lara, R. Cerezo, and C. Romero, "A review on data fusion in multimodal learning analytics and educational data mining," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 12, p. e1458, 2022.
- [2] W. Nsengiyumva, S. Zhong, M. Luo, Q. Zhang, and J. Lin, "Critical insights into the state-of-the-art nde data fusion techniques for the inspection of structural systems," *Structural Control and Health Monitoring*, vol. 29, p. e2857, 2022.
- [3] A. Smolinska, J. Engel, E. Szymanska, L. Buydens, and L. Blanchet, "Chapter 3 - general framing of low-, mid-, and high-level data fusion with examples in the life sciences," vol. 31, pp. 51–79, 2019.
- [4] S. Gao, Y. Chen, F. Shi, Peng *et al.*, "Lkg-net: lightweight keratoconus grading network based on corneal topography," *Biomedical Optics Express*, vol. 14, pp. 799–814, 2023.
- [5] A. H. Al-Timemy, L. Alzubaidi, Mosa *et al.*, "A deep feature fusion of improved suspected keratoconus detection with deep learning," *Diagnostics*, vol. 13, p. 1689, 2023.
- [6] R. Feng, Z. Xu, X. Zheng, H. Hu, Jin *et al.*, "Ker-net: a novel deep learning approach for keratoconus and sub-clinical keratoconus detection based on raw data of the pentacam hr system," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, pp. 3898–3910, 2021.
- [7] M. M. Daud, W. M. D. W. Zaki, A. Hussain, and H. A. Mutalib, "Keratoconus detection using the fusion features of anterior and lateral segment photographed images," *IEEE Access*, vol. 8, pp. 142 282–142 294, 2020.

Table 2: Classification performanc of adopted DL model.

Technique	Class	Precision	Recall	F1-score	Accuracy
Without Fusion	Keratoconus	92.00%	87.00%	89.00%	97.56%
	Normal	98.00%	99.00%	99.00%	98.34%
	Suspicious	96.00%	96.00%	96.00%	97.66%
	Global Classification Accuracy				96.78%
With Fusion	Keratoconus	97.00%	94.00%	95.00%	98.93%
	Normal	99.00%	100.0%	99.00%	99.32%
	Suspicious	98.00%	97.00%	97.00%	98.44%
	Global Classification Accuracy				98.34%

- [8] S. Hashemi, H. Veisi, E. Jafarzadehpur, R. Rahmani, and Z. Heshmati, “An image processing approach for rigid gas-permeable lens base-curve identification,” *Signal, Image and Video Processing*, vol. 14, pp. 971–979, 2020.
- [9] C. Thorat, A. Bhat, P. Sawant, I. Bartakke, and S. Shirsath, “A detailed review on text extraction using optical character recognition,” *ICT Analysis and Applications*, pp. 719–728, 2022.
- [10] A. B. Abadi and S. Tahcfulloh, “Digital image processing for height measurement application based on python opencv and regression analysis,” *JOIV: International Journal on Informatics Visualization*, vol. 6, no. 4, pp. 763–770, 2022.
- [11] R. K. Cho, K. Sood, and C. S. C. Channapragada, “Image repair and restoration using deep learning,” in *2022 4th International Conference on Artificial Intelligence and Speech Technology (AIST)*. IEEE, 2022, pp. 1–8.