M. A. Jabbar
Fernando Ortiz-Rodríguez
Sanju Tiwari
Patrick Siarry (Eds.)

Communications in Computer and Information Science 1818

Applied Machine Learning and Data Analytics

5th International Conference, AMLDA 2022
Reynosa, Tamaulipas, Mexico, December 22–23, 2022
Revised Selected Papers

AMLDA
2022

Springer

Communications in Computer and Information Science 1818

Rationale

The CCIS series is devoted to the publication of proceedings of computer science conferences. Its aim is to efficiently disseminate original research results in informatics in printed and electronic form. While the focus is on publication of peer-reviewed full papers presenting mature work, inclusion of reviewed short papers reporting on work in progress is welcome, too. Besides globally relevant meetings with internationally representative program committees guaranteeing a strict peer-reviewing and paper selection process, conferences run by societies or of high regional or national relevance are also considered for publication.

Topics

The topical scope of CCIS spans the entire spectrum of informatics ranging from foundational topics in the theory of computing to information and communications science and technology and a broad variety of interdisciplinary application fields.

Information for Volume Editors and Authors

Publication in CCIS is free of charge. No royalties are paid, however, we offer registered conference participants temporary free access to the online version of the conference proceedings on SpringerLink (http://link.springer.com) by means of an http referrer from the conference website and/or a number of complimentary printed copies, as specified in the official acceptance email of the event.

CCIS proceedings can be published in time for distribution at conferences or as post-proceedings, and delivered in the form of printed books and/or electronically as USBs and/or e-content licenses for accessing proceedings at SpringerLink. Furthermore, CCIS proceedings are included in the CCIS electronic book series hosted in the SpringerLink digital library at http://link.springer.com/bookseries/7899. Conferences publishing in CCIS are allowed to use Online Conference Service (OCS) for managing the whole proceedings lifecycle (from submission and reviewing to preparing for publication) free of charge.

Publication process

The language of publication is exclusively English. Authors publishing in CCIS have to sign the Springer CCIS copyright transfer form, however, they are free to use their material published in CCIS for substantially changed, more elaborate subsequent publications elsewhere. For the preparation of the camera-ready papers/files, authors have to strictly adhere to the Springer CCIS Authors' Instructions and are strongly encouraged to use the CCIS LaTeX style files or templates.

Abstracting/Indexing

CCIS is abstracted/indexed in DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago, Scopus. CCIS volumes are also submitted for the inclusion in ISI Proceedings.

How to start

To start the evaluation of your proposal for inclusion in the CCIS series, please send an e-mail to ccis@springer.com.

M. A. Jabbar · Fernando Ortiz-Rodríguez ·
Sanju Tiwari · Patrick Siarry
Editors

Applied Machine Learning and Data Analytics

5th International Conference, AMLDA 2022
Reynosa, Tamaulipas, Mexico, December 22–23, 2022
Revised Selected Papers

Springer

Editors
M. A. Jabbar
Vardhaman College of Engineering
Hyderabad, India

Fernando Ortiz-Rodríguez
Autonomous University of Tamaulipas
Ciudad Victoria, Mexico

Sanju Tiwari
Autonomous University of Tamaulipas
Ciudad Victoria, Mexico

Patrick Siarry
University of Paris-Est Creteil
Vitry sur Seine, France

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-031-34221-9 ISBN 978-3-031-34222-6 (eBook)
https://doi.org/10.1007/978-3-031-34222-6

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains the main proceedings of the 2022 edition of Applied Machine Learning and Data Analytics (AMLDA 2022). AMLDA is established as a yearly venue for discussing the latest scientific results and technology innovations related to Machine Learning.

The International Conference on Applied Machine Learning and Data Analytics (AMLDA) is the premier conference in this branch of artificial intelligence. AMLDA renowned for presenting cutting-edge research on all aspects of machine learning as well as important application areas such as healthcare and medical imaging informatics, biometrics, forensics, precision agriculture, risk management, robotics, and satellite imaging. Participants at AMLDA 2022 included academic and industrial researchers and engineers, graduate students, and postdocs.

AMLDA is convened annually to provide a platform for knowledge exchange on the most recent scientific and technological advances in the field of applied machine learning and data analytics.

The main scientific program of the conference comprised 20 papers: 16 full research papers and four short research papers selected out of 89 reviewed submissions, which corresponds to an acceptance rate of 22.4% for the research papers submitted. The program also included two exciting, invited keynotes (Stefan Kramer and Edlira Kalemi), with novel topics.

The General and Program Committee chairs would like to thank the many people involved in making AMLDA 2022 a success. First, our thanks go to the four co-chairs of the main event and more than 50 reviewers for ensuring a rigorous review process that led to an excellent scientific program and an average of three reviews per article.

Further, we thank the kind support of all people from Tamaulipas Autonomous University, particularly the Faculty at the venue, the Reynosa Rodhe campus. We are thankful for the kind support of the staff of Springer. We finally thank our sponsors and our community for their vital support of this edition of AMLDA.

The editors would like to close the preface with warm thanks to our supporting keynotes, the program committee for rigorous commitment in carrying out reviews, and finally, our enthusiastic authors who made this event truly International.

December 2022

M. A. Jabbar
Fernando Ortiz-Rodriguez
Sanju Tiwari
Patrick Siarry

Organization

General Chairs

M. A. Jabbar	Vardhaman College of Engineering, India
Fernando Ortiz Rodriguez	Universidad Autónoma de Tamaulipas, Mexico
Sanju Tiwari	Universidad Autónoma de Tamaulipas, Mexico
Patrick Siarry	University of Paris-Est Créteil, France

Program Chairs

Millie Pant	IIT Roorkee, India
Shishir Kumar Shandilya	VIT Bhopal University, India
José Melchor Medina-Quintero	Universidad Autónoma de Tamaulipas, Mexico
Andries Engelbrecht	University of Stellenbosch, South Africa

Publicity Chairs

P. Sivakumara	University of Malaya, Malaysia
Vania V. Estrela	Universidade Federal Fluminense, Brazil
Patience Usoro Usip	University of Uyo, Nigeria

Workshop Chairs

Yusniel Hidalgo Delgado	Universidad de las Ciencias Informáticas, Cuba
Shikha Mehta	JIIT Noida, India
C. A. Dhawale	P R Pote Patil College of Engineering and Management, India

Special Session Chairs

Tanima Datta	IIT BHU, India
David Hicks	Texas A & M University Kingsville, USA
Antonio De Nicola	ENEA, Italy
Patrice Boursier	International Medical University, Malaysia

Tutorial Chairs

Hari Prabhat Gupta	IIT BHU, India
Ayush Goyal	Texas A & M University Kingsville, USA
Smita Shandilya	Sagar Institute of Research and Technology, India

Program Committee

Fatima Zahra Amara	Abbès Laghrour University of Khenchela, Algeria
Dridi Amna	Birmingham City University, UK
Ivan Armuelles Voinov	University of Panama, Panama
Debajyoty Banik	Kalinga Institute of Industrial Technology, India
Bilal Ben Mahria	Sidi Mohamed Ben Abdellah University, Morocco
Semih Bitim	İzmir Bakırçay Üniversitesi, Turkey
Orchid Chetia Phukan	University of Tartu, Estonia
Gerard Deepak	Manipal Institute of Technology Bengaluru, India
Burcu Devrim İçtenbaş	Ankara Bilim University, Turkey
Onur Dogan	İzmir Bakırçay University, Turkey
Houda El Bouhissi	University of Bejaja, Algeria
Orhan Er	İzmir Bakırçay University, Turkey
Christian Estay-Niculcar	FUNIBER, Spain
Abid Ali	Halmstad University, Sweden
Devottam Gaurav	Indian Institute of Technology Delhi, India
Orlando Toledano	Universidad de las Ciencias Informáticas, Cuba
Sven Groppe	Universität zu Lübeck, Germany
Shankru Guggari	BMS College of Engineering, India
Mounir Hemam	Khenchela University, India
Yuseniel Hidalgo	Universidad de las Ciencias Informáticas, Cuba
Hanieh Khorashadizadeh	Universität zu Lübeck, Germany
Ajit Kumar	Soongsil University, South Korea
Kusum Lata	NIT Hamirpur, India
Yoan Antonio Lopez	Universidad de las Ciencias Informáticas, Cuba
Víctor López	Universidad Tecnológica de Panama, Panama
Thangavel M.	VIT Bhopal University, India
Melchor Medina	Universidad Autónoma de Tamaulipas, Mexico
José Eladio Medina Pagola	Universidad de las Ciencias Informáticas, Cuba
M. A. Jabbar	Vardhaman College of Engineering, India
Shikha Mehta	JIIT Noida, India
Vladimir Milián Núñez	Universidad de las Ciencias Informáticas, Cuba
Mohamed Madani Hafidi	Abbès Laghrour University of Khenchela, Algeria
Mounir Hemam	Abbès Laghrour University of Khenchela, Algeria

Ron Ojino	Cooperative University of Kenya, Kenya
Fernando Ortiz-Rodriguez	Universidad Autónoma de Tamaulipas, Mexico
Piyush Pareek	Nitte Meenakshi Institute of Technology, India
Krishna Ponnekanti	Anil Neerukonda Institute of Technology & Sciences, India
Hector Raúl	Universidad de las Ciencias Informáticas, Cuba
Shishir Shandilya	VIT Bhopal University, India
Serge Sonfack	INP-Toulouse, France
Tagaram Soni Madhulatha	Telangana Social Welfare Residential Degree College for Women, India
Sanju Tiwari	Universidad Autónoma de Tamaulipas, Mexico
Patience Usoro Usip	University of Uyo, Nigeria
Edlira Vakaj	Birmingham City University, UK
David Valle-Cruz	Universidad Autónoma del Estado de México, Mexico
Miguel Vargas-Lombardo	Universidad Tecnológica de Panama, Panama
Sonali Vyas	University of Petroleum and Energy Studies, Dehradun, India
Bezza Youcef	Abbès Laghrour University of Khenchela, Algeria

Contents

Keratoconus Classification Using Feature Selection and Machine Learning Approach

E. Smily Yeya Jothi[1], J. Anitha[2], and Jude Hemanth[2](✉)

[1] Department of Biomedical Instrumentation Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India
[2] Department of Electronics and Communication, Karunya Institute of Technology and Sciences, Coimbatore, India
judehemanth@karunya.edu

Abstract. The development and design of new tools that can help identify the disease at an early stage may aid in preventing or delaying the progression of the disease. The preservation of young populations' vision is crucial because Keratoconus (KC) typically affects people at puberty, with children being the most affected. The goal of this research is to identify the most pertinent variables in relation to the various keratoconus classifiers employed in the Harvard Dataverse keratoconus dataset. Out of 3162 observations, a total of 446 parameters are examined by 3 feature selection techniques, Sequential Forward, and Backward Selection (SFS and SBS), Genetic Algorithm (GA) and Particle Swarm Optimisation (PSO). The prediction is done through machine learning algorithms. The experimental results implied that Sequential forward selection and PSO provided the highest classification accuracy by the application of Negative Correlation based Deep Ensemble Network (NC-DEN), with an accuracy of 96.14%.

Keywords: Keratoconus · SFS · GA · PSO · SVM · Ensemble · Neural network

1 Introduction

Noninflammatory corneal disease known as keratoconus frequently affects both eyes. Usually, the cornea is shaped like a ball and is round. However, occasionally the cornea's structure is insufficiently robust to maintain this spherical shape. The eye's normally round surface might develop a cone-like outgrowth over time. Keratoconus is the medical term for this condition. In keratoconus, the cornea is bent into a conical shape and the stroma is subsequently thinned as portrayed in Fig. 1. The development of an uneven astigmatism, which is frequently difficult to regulate and typically results in vision deterioration. Progressive keratoconus may result in a slow loss of eyesight that eventually affects the patient's quality of life [1]. Figure 2 shows the Corneal Topography Image. Recent advancements in refractive surgery have accelerated the development of imaging techniques for the cornea. The pertinent principles of corneal optics are addressed in order to comprehend the significance of novel imaging techniques. Placido disc patterns or mires, which are reflected off the anterior cornea's tear film and converted

M. A. Jabbar et al. (Eds.): AMLDA 2022, CCIS 1818, pp. 39–51, 2023.
https://doi.org/10.1007/978-3-031-34222-6_4

to colour scales, are used to assess topography. Since the image is derived from the tear film, tear film imperfections can have a considerable impact on the accuracy and quality of a Placido disc topography. Second, poor patient fixation may degrade the topographic image's quality. Finally, posterior elevation results are less accurate, particularly following refractive surgery [2–6].

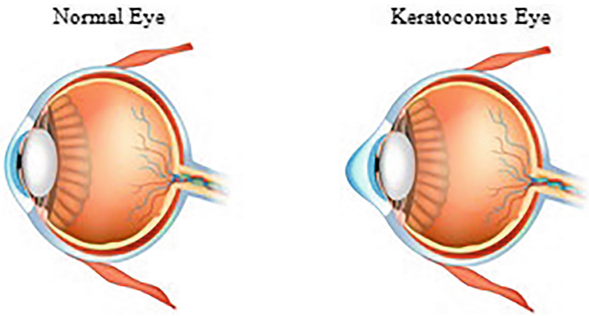

Fig. 1. Normal vs keratoconus eye

The keratoconus has been identified and categorised using several topographical indices. However, due to the sheer volume of indices, manual identification is challenging. Choosing more pertinent indices is also necessary for computerised diagnosis in order to cut down on computation time and error. Because it is so simple to gather data in a hospital clinic, the majority of prevalence studies have been done there. These results provide an approximation of prevalence, but they are probably to understate the true incidence of the disease because hospital patients are frequently symptomatic, making early stages of the condition difficult to detect [4].

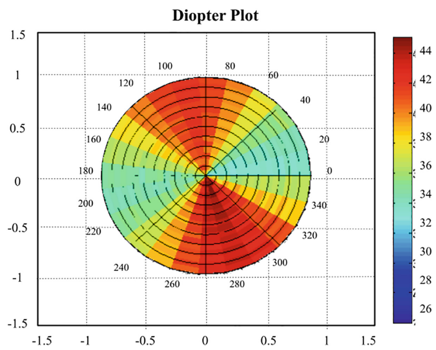

Fig. 2. Corneal Topography Image

The frequency and prognosis of keratoconus vary widely around the world, according to epidemiological research, with Middle Eastern and Asian ethnic groups and those in their 20s and 30s often experiencing the highest rates. The use of novel imaging techniques for the human cornea has improved our knowledge of the condition [5]. The severe stages of keratoconus can be distinctly seen and recognizable as Munson's sign, Vogt's striae, Fleischer's Ring, etc. [6]. Observation of corneal topography and continuous monitoring of its thickness are therefore essential for the early diagnosis of keratoconus. A novel diagnostic method and device should be developed to improve life quality and save lives in children with keratoconus since it affects children so young. The effectiveness of categorization will be impacted by the capabilities that various models bring to the table when classifying data. To achieve the best accuracy, several feature selection techniques are also applied. To a certain extent, dimensionality reduction approaches have been used to cut down on the number of features or indices. The development of deep learning and machine learning architectures has aided in keratoconus diagnosis [13, 14].

The following is a summary of this study's significant contributions. First, the most important parameters are extracted from the examination of other parameters, particularly for the analysis of categorization data. The second step is a comparison of various machine learning models, including the Support Vector Machine (SVM), Ensemble Network (EN), and Neural Network based Classifier (NN). The rest of the paper is outlined as follows, Sect. 2 reviews the existing and current methods involved in the detection of KC. Section 3 provides detailed description of the dataset, feature selection methods and classifiers. Section 4 discusses the outcomes of the study and Sect. 5 concludes the research.

2 Related Works

Over the past few years, machine and deep learning analysis has drawn a lot of attention in a variety of medical fields, including ophthalmology. Using machine learning to analyse clinical and topographical data from 124 KC patients. With reference to the current Amsler-Krumeich (A-K) classification method, supervised multilayer perceptron and unsupervised variational autoencoder models were both employed to categorise KC patients. High accuracy was achieved by both techniques, albeit the unsupervised technique performs better. The outcome demonstrated that the unsupervised method using a choice of 29 variables might be an effective tool to give clinicians an automatic classification tool [1]. Ensemble Deep Transfer Learning (EDTL) technique along with Pentacam Indices (PI) was used in the classification of KC along with four pre-trained models. An improved accuracy of 98.3% was achieved with AlexNet. The imbalance in the PI was corrected using Synthetic Minority Over-Sampling Technique (SMOTE). In the pre-processing stage, the topographic map and the colour code were extracted [4].

PSO was considered for the optimisation of the segmented image. PSO, DPSO (Discrete Particle Swarm Optimization) and FPSO (Fraction order Particle Swarm Optimization). Pre-trained Convolutional Neural Network (CNN) – VGG16 was used for the training and testing. Three distinct classes of images were considered, Normal, Subclinical and Keratoconus. When optimization was performed, as well as when compared to the results from the literature with three different classifications of keratoconus, the results revealed increased performance in terms of accuracy, sensitivity, and specificity [6]. Keratoconus Harvard dataset was used in the study for the classification of KC. Eleven different feature selection techniques were used along with seven different classifiers. The authors experimented that Sequential forward selection algorithm along with

RF yielded best outcomes. Considering the execution time, the SFS algorithm showed promising results in the performance. This work aimed at the identification of KC through corneal topographic maps [7]. The anterior and lateral portion of the eye was captured by the smartphone camera and was used to identify KC. A total of 280 images were considered in the study. The automated contour model and semiautomatic spline model were used to extract the geometric features from the segmented portion of the image. The feature selection was done using an infinite latent feature selection method and three different classification techniques were utilised in the study. The method proved to be efficient in the identification of KC based on fusion features [8]. SVM was used for the classification of the early stage of KC using topographic, pachymetry and aberrometry features. Two different classes of classification were one: former is KC with normal and latter is early KC with subclinical KC [9]. 25 different machine learning algorithms were used for the prediction of KC. The algorithms were tested based on real-time data which includes, corneal elevation, topography, and parameters of pachymetry. The dataset was split and the performance was evaluated on 2 class and 3 class techniques. In both the predictions, Support Vector Machine (SVM) achieved maximum accuracy of 94% [10].

A hybrid deep learning model was proposed for detecting KC using corneal maps. An average of 5024 images were collected from 280 subjects (540 eyes). Seven deep learning models were developed based on the extracted features. The available dataset was split into three classes, the developmental dataset with 542 eyes, independent test dataset with 150 eyes, and Merged dataset consisting of 692 eye images. 2 class prediction and 3 class prediction were done for the identification of KC [11]. Table 1 describes the different features extracted and the methods used for the extraction by various researchers, along with their performance metrics.

Table 1. Comparative Study

Reference	Number of Images	Features Extracted	Feature Selection Method	Classification	Inference
Al-timemy et al., 2021	5024	• Anterior and Posterior Eccentricity • Anterior and Posterior elevation • Anterior and Posterior sagittal curvature • Corneal thickness	EfficientNet-b0	EfficientNet-b0	AUC-0.99 F1 – 0.99 Accuracy – 98.8%

(*continued*)

Table 1. (*continued*)

Reference	Number of Images	Features Extracted	Feature Selection Method	Classification	Inference
Lavric A et al., 2020	3151	• Higher Order irregular astigmatism • Maximum keratometric power • Best fit sphere • Standard deviation of pachymetry • Higher-order aberrations • Aberrations parameters in coma orders 5 • Aberrations parameters in Sphere orders 5	-	25 - ML	SVM Accuracy – 94%
Shanthi. S et al., 2021	205	• Corneal Topography of the anterior face -6 features • Corneal Topography of the posterior face -21 features	Recursive Feature Elimination	SVM	Accuracy-91.8%, sensitivity-94.2%, specificity-97.5%
Daud.M et al., 2020	280 images (140 KC and 140 normal)	• Horizontal Visible Iris Diameter • Vertical Visible Iris Diameter • Eccentricity • Asphericity • Major Line • Minor Line • Orientation	Infinite Latent Feature Selection	Random Forest with n = 50 SVM-Linear, Quad, Cubic, RBF KNN	Accuracy – 96.5%, Sensitivity-98.4%, Specificity-93.6%

(*continued*)

Table 1. (*continued*)

Reference	Number of Images	Features Extracted	Feature Selection Method	Classification	Inference
Aatila M et al., 2021	3162 observations	446 features from.csv file	Wrapper Embedded Filter and Hybrid	LR LDA KNN CART NB SVM RF	RF – 98%
Subramanian. P, Ramesh.G 2022	1500 images	-	PSO, DPSO, FPSO	CNN – VGG16	Accuracy-95.9%
Al-timemy et al., 2022	444 (226 right eye, 218 left eye cases)	-	SMOTE	EDTL, SqueezeNet, AlexNet, Shuffle Net, Mobile Net-v2, PI	-
Hallet N et al., 2020	124	29 Features collected from Questionnaire, PI, Clinical, Medical records	-	MLP VAE	Accurcy-73% and 80%

3 Methodology

The work model of the proposed study is shown in Fig. 3. The dataset description, the feature selection models and classifiers are discussed in this section. Based on the previous research studies, the feature selection algorithms are chosen. Three algorithms are chosen to identify and classify KC. The work is implemented using MATLAB R2021a.

3.1 Dataset

The dataset used in this study is taken from [12]. This dataset, which is organized as a csv file, contains 446 features across 3162 rows. According to Table 2 as shown below, there are four categories for eyes [7].

3.2 Feature Selection

The representation and appropriateness of the data that a machine learning system uses are two aspects that have an impact on how well that system performs. In general, not all learning information is always pertinent to the system. However, the learning system places a high value on the selection of pertinent features by removing less useful, pointless, or even irrelevant factors.

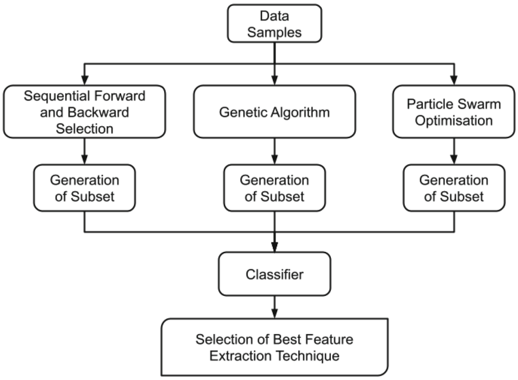

Fig. 3. Work Model

Table 2. Description of the Dataset

Size of the original dataset		Class	Number of rows
Number of features	Number of rows		
446	3162	1	264
		2	2595
		3	221
		4	82

3.2.1 Sequential Forward and Backward Selection

An iterative algorithm called sequential forward selection (SFS) begins with an empty subset of variables. The FFS algorithm assesses each variable separately for each iteration, keeping just the one that has the greatest impact on the model's performance. When a new variable no longer improves the system's performance, the selection process comes to an end. An iterative approach called sequential backward selection (SBS) uses all the dataset's attributes at the beginning of the process. Each round of BFE eliminates the least important variable until no performance gain is seen [7].

3.2.2 Genetic Algorithm

Iterative algorithms based on the genetic evolution process are known as genetic algorithms (GA). GA creates chromosomes from a starting population by suggesting potential answers to the topic being researched. To arrive at the optimal solution, this initial population of solutions evolves utilizing three operators (selection, crossing, and mutation operators) [7].

3.2.3 Particle Swarm Optimisation

Particle Swarm Optimisation (PSO) is used for the linearization of the database. With fewer parameters to alter, PSO is simple to use and faster. The particle swarm optimization can reach a solution more quickly and uses only fundamental mathematical operations rather than difficult-to-implement derivatives. A swarm or cluster of particles used in particle swarm optimization each represents a potential solution. Here, each particle's intensity N is compared to each particle's intensity L from the input image [6].

Condition 1: $L > N$, variable α is incremented by 1, and the corresponding intensity is added to β.

Condition 2: $L < N$, variable f is incremented by 1, and the corresponding intensity is added to g.

The fitness function of the optimised network is calculated as follows in Eq. 1,

$$Fitness\,Fn = f * \alpha * ((g/f) - (\beta/\alpha))^2 \tag{1}$$

3.3 Classification

3.3.1 Negative Correlation Based Deep Ensemble Network (NC-DEN)

Multiple separate models are combined using ensemble learning to improve generalization performance. Deep learning architectures are currently outperforming shallow or standard models in terms of performance. Deep ensemble learning models combine the benefits of both deep learning models and ensemble learning, improving the generalization performance of the resulting model. An essential method for educating the learning algorithms is Negative Correlation Learning (NCL). The NCL's basic idea is to promote variation among the individual models in the ensemble in order to help them understand the many facets of the training set. By minimizing the error functions of the individual networks, NCL minimizes the empirical risk function of the ensemble model. Both classification and regression tasks include evaluating NCL. On classification tasks, simple averaging, and winner-takes-all measures, as well as simple average combination approaches for regression issues, were utilized in the evaluation [15]. Figure 4 shows the flowchart of NC-DEN network.

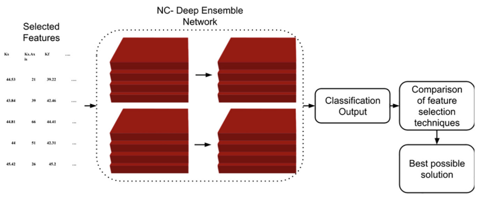

Fig. 4. Flowchart of NC-DEN

3.3.2 Neural Network-Based Classifier

Complex models called neural networks aim to replicate how the human brain creates classification rules. A neural network is made up of several different layers of neurons, each of which receives input from the layers below it and sends output to the layers above it [14]. Every output serves as an input for a subsequent function, and the neurons within the network communicate with the neurons in the layer below. Every function, including the starting neuron, uses an internalized function that incorporates the addition of a bias term that is specific to each neuron to produce a numeric output. The output is then multiplied by the appropriate weight to create the numeric input for the function in the following layer. This continues until the network produces its result.

3.4 SVM

SVM is an algorithm for machine learning for classification and regression issues that is built on kernels. In recent years, SVM has gained a disproportionate amount of attention from researchers in the fields of data mining, pattern recognition, and machine learning due to its exceptional generalization efficiency, optimal outcome, and differentiation power. The main goal of SVM is to divide the training set's various classes into as many as possible using a surface that represents the proportion between the classes. With the objective of minimizing misclassifications, each training sample is represented by a piece of evidence in a two-dimensional region that can be divided by a line. More precisely, there are countless lines that can help create a data set that is adequately divided into acceptable parts. The goal of SVM classification is to choose the line that minimizes the maximum margin as the most suitable one [15].

4 Experimental Parameters

The best feature selection technique is tested based on the performance of the identification of KC. This method plays a vital role in the removal of redundant and irrelevant data, and increases the classification performance. The dataset is split into training and testing set. The classification performance of the classifier by using different feature selection techniques are shown in Fig. 5, 6, and 7. In feature selection technique considered, the generation of feature subset is crucial for the determination of the efficiency.

In this work, three different feature extraction techniques were implemented, of which, PSO performed well. Despite the difference in the classification network, the forward and backward selection algorithm yielded an average accuracy of 88%, GA with 90% accuracy, and finally PSO with 92% accuracy. The comparison of feature selection models with different classifiers is shown in Table 3. The NC-DEN classifier produced better results than all of the feature selection techniques used in the investigation.

Sensitivity is the ability of the network to identify the class of the disease correctly. Specificity is the ability of the network to identify people who do not have the disease class. The formula for the calculation of accuracy, sensitivity and specificity are mentioned in Eqs. 2–4.

Table 3. Classifier Vs Feature Extraction performance

Classifier	Feature selection model	Accuracy	Sensitivity	Specificity
NC-DEN	SF-SB	**0.93**	**0.93**	**0.92**
NN		0.85	0.91	0.89
SVM		0.86	0.84	0.93
NC-DEN	GA	**0.95**	**0.94**	**0.95**
NN		0.87	0.92	0.92
SVM		0.90	0.86	0.92
NC-DEN	PSO	**0.96**	**0.97**	**0.97**
NN		0.89	0.94	0.95
SVM		0.92	0.93	0.94

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} * 100 \tag{2}$$

$$Sensitivity = \frac{TP}{(TP + FN)} * 100 \tag{3}$$

$$Specificity = \frac{TN}{(TN + FP)} * 100 \tag{4}$$

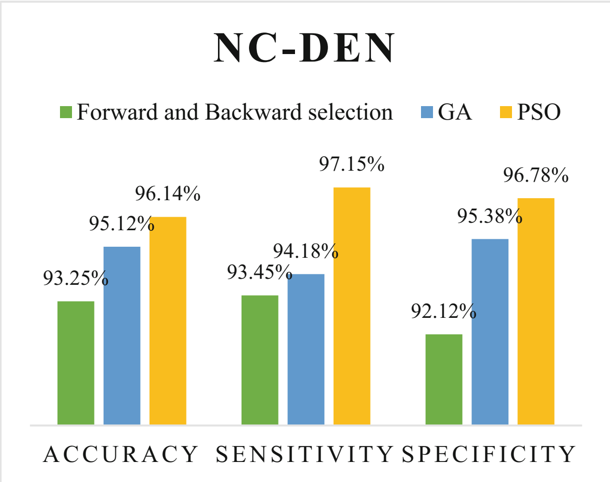

Fig. 5. Performance of Ensemble Network

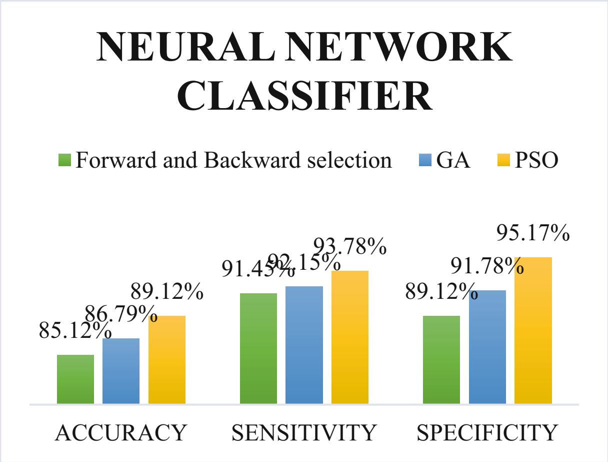

Fig. 6. Performance of Neural Network based Classifier

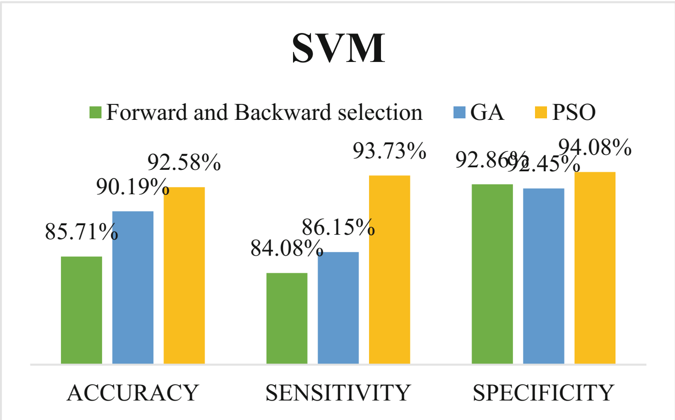

Fig. 7. Performance of SVM

From the Table 3 and Fig. 5 it is inferred that the NC-DEN classifier showed better results for all the feature selection methods utilized in the study. However, of all the feature extraction methods, PSO performed well for the entire classification network. This work highlights that PSO in combination with NC-DEN works efficiently in the identification of classes of KC.

5 Conclusion

It is hard to categorize the severity of keratoconus since the duration for the onset of the illness's signs and symptoms and their correlation with pathogenicity are highly diverse. The precision of the created algorithm enables the proper treatment to be administered when the sickness is still in its early stages, considerably improving quality of life. The acquired results indicate the validation as well as performance assessment of the created algorithm, which may be simply integrated in an ophthalmologist's instrument and help with the early detection of keratoconus. The proposed method is efficient in the identification of the severity of KC using the implemented feature selection and classification technique. PSO exhibited better performance and a higher convergence rate compared to SF-SB and GA. The execution time required for SF-SB, GA and PSO methods are 270 s, 240 s and 150 s respectively.

The feature selection technique is important in machine learning approach to obtain accurate results. They help in increasing the prediction accuracy of the classifiers. The proposed technique using PSO and NC-DEN provided promising results in the classification and identification of KC at an early stage. The network provided an average accuracy of 95%, sensitivity of 95%, and specificity of 94%. GA also showed better results compared with SF-SB technique. By employing the feature selection technique, the most redundant and the irrelevant data are removed, thus enhancing the performance in the classification of Keratoconus Disease.

References

1. Hallett, N., et al.: Deep learning based unsupervised and semi-supervised classification for keratoconus. In: 2020 International Joint Conference on Neural Networks (IJCNN), pp. 1–7 (2020)
2. Shahida, Qadir, M.: Keratoconous: an eye disease. Biomed. J. Sci. Tech. Res. (2020)
3. Gordon-Shaag, A., Millodot, M., Shneor, E., Liu, Y.: The genetic and environmental factors for keratoconus. BioMed Res. Int. (2015)
4. Al-timemy, A.H., Ghaeb, N.H., Mosa, Z.M., Escudero, J.: Deep transfer learning for improved detection of keratoconus using corneal topographic maps. Cogn. Comput. **14**, 1627–1642 (2022)
5. Santodomingo-Rubido, J., Carracedo, G., Suzaki, A., Villa-Collar, C.C., Vincent, S.J., Wolffsohn, J.S.: Keratoconus: An updated review. Contact Lens and Anterior Eye (2022)
6. Subramanian, P., Ramesh, G.P.: Keratoconus classification with convolutional neural networks using segmentation and index quantification of eye topography images by particle swarm optimisation. BioMed Res. Int. (2022)
7. Aatila, M., Lachgar, M., Hamid, H., Kartit, A.: Keratoconus severity classification using features selection and machine learning algorithms. Comput. Math. Methods Med. (2021)
8. Daud, M.M., Zaki, W.M., Hussain, A., Mutalib, H.A.: Keratoconus detection using the fusion features of anterior and lateral segment photographed images. IEEE Access **8**, 142282–142294 (2020)
9. Shanthi, S., Nirmaladevi, K., Pyingkodi, M., Dharanesh, K., Gowthaman, T., Harsavardan, B.: Machine learning approach for detection of keratoconus. IOP Conf. Ser. Mater. Sci. Eng., 1055 (2021)
10. Lavric, A., Popa, V., Takahashi, H., Yousefi, S.: Detecting keratoconus from corneal imaging data using machine learning. IEEE Access **8**, 149113–149121 (2020)
11. Tiwari, S.M., et al.: Applications of machine learning approaches to combat COVID-19: a survey. Lessons from COVID-19, 263–287 (2022)
12. Gaurav, D., Rodriguez, F.O., Tiwari, S.M., & Jabbar, M.A. (2021). Review of Machine Learning Approach for Drug Development Process. Deep Learning in Biomedical and Health Informatics

13. Al-timemy, A.H., et al.: A hybrid deep learning construct for detecting keratoconus from corneal maps. Transl. Vision Sci. Technol. **10** (2021)
14. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:https://doi.org/10.7910/DVN/G2CRMO
15. Liu, Y., Yao, X.: Ensemble learning via negative correlation. Neural Netw. **12**, 1399–1404 (1999). https://doi.org/10.1016/S0893-6080(99)00073-8