International Journal of Engineering & Technology, 7 (4) (2018) 4201-4206

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet. v7i4.21243
Research paper

A data mining framework for the classification of retinopathy images based on a new multistage prediction algorithm

Karim Hashim AL-Saedi [1] *, Herbert F. Jelinek [2]

[1] Department of Computer Science, College of Science, Mustansiriyah University, Iraq
[2] School of Community Health, Faculty of Science, Charles Sturt University, Australia
**Corresponding author E-mail:dr.karim@uomustansiriyah.edu.iq*

Abstract

Medical image processing, analysis and classification is a rapidly expanding field providing possible solutions for health care providers including ophthalmologists and optometrists. To be of value, image analysis and classification requires high accuracy and fast processing. Early detection of diabetic retinopathy can lead to better treatment outcomes, especially in rural and remote areas where there is a lack of specialists. In the current work we propose a highly accurate prediction model based on optic disc color characteristics. We propose a data mining algorithm based on a top-down processing framework. The framework involves a new Multistage Prediction algorithm (MSP1) consisting of segmentation of the optic disc, dilation, and color normalization, color histogram determination, and calculating the predicted classification score of each image. The final step carries out the process of classification of all images based on the Group Method of Data Handling(GMDH) application. One hundred and fifty seven images were available for classification. The results indicate that the proposed Multistage Prediction algorithm combined with the GMDH classification framework improved on previous results with an overall accuracy of 96.8%, and sensitivity of 95% and an F-measure for the classification performance of96%. MSP1 is easy to implement on any laptop and therefore provides a robust option for tele-ophthalmology diagnostics of retinopathy.

Keywords:*Method of Data Handling; GMDH; Diabetic Retinopathy; Optic Disc; Medical Image processing; Data Mining.*

1. Introduction

Data mining techniques and image processing are used extensively in engineering and medical sciences. The early detection of the any disease or disease process is a very significant step in providing timely treatment and optimal outcomes [1-2].Group Method of Data Handling (GMDH Shell) is GUI application which consists of a host application and plug-ins [3-4]. The GMDH Shell (https://gmdhsoftware.com/) includes seven template techniques that can be used by an end user to facilitate decision making. Those techniques include Classifying, Clustering, Curve fitting, Demand forecasting, Load forecasting, Regression, and Time series forecasting [5].

Data mining algorithms have been shown to be accurate and robust for identification of various diseases including retinopathy, kidney disease,diabetes, high blood pressure, and atherosclerosis [6]. Datamining can be usedforclassificationand involves utilizing variables or fields within the database to identify the unknown values of the variables of interest and include decision trees, neural networks, Bayesian, SVM. [1-2]

One of the most important advances in retinal vascular image analysis has been the clear demonstration that physiological and pathological alterations within the retinal vascular network are related to diabetes and associated diabetic retinopathy progression [3]. However, clinical assessment of the retinal vasculature and macula is time consuming for specialists who are often not available in rural and remote areas. Recent advances in image analysis provides specialises and rural health workers with objective and duplicable results for diagnostics of diabetic retinopathy [3], [4].

Figure 1 depicts an optic disc of a normal and an advanced stage of diabetic retinopathy.

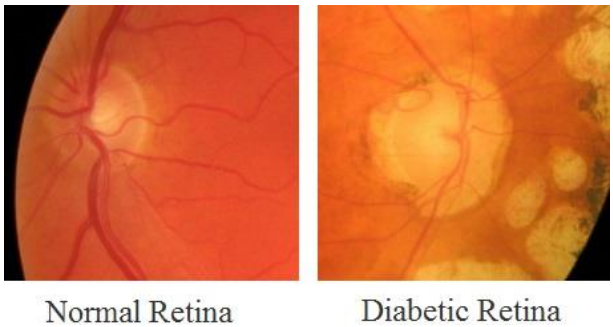

Fig. 1: Normal and Diabetic Retina Image.

Automated funds analysis depends upon precise image segmentation and identification of hallmarks of the disease. These may include micro aneurysms or cotton wool spots. Identification of known lesions within images presents a bottom-up image analysis and classification approach [6]. Data mining algorithms can be incorporated to improve on classification if several features characterise the image but are also useful if features are not based on known or accepted structural characteristics seen in diabetic retinopathy progression [7], [8].

A survey of segmentation techniques for the cup boundaries and optic disc(OD) has been published and provides current algorithms and their accuracy [12].An automated optic disc pathology detectionapproach was proposed by Alghamdi et al., 2016 [13], which was based on a supervised model for detection of abnormal

features associated with the OD. Their model was divided into two learning parts: the multiple classifier, which is responsible for learning OD global attributes by applying the deep Convolution Neural Networks (CNNs) to extract OD regions. These extracted regions are then passed to the second layer of the optimized deep CNNs to learn abnormal attribute representations.

An approach that automatically detects exudates in diabetic retinopathy images based on a Fuzzy C-Means clustering approach was suggested by Sathiyamoorthy and Kulanthaivel [14]. The proposed approach consisted of a pre-processing phase, which is applied to improve the quality of the image; a second phase thatenhances the image based on adaptive histogram equation (AHE) and a third phase to detect exudates using binary operations. The limitation of this approach was that only hard exudates were detected from the retinal image.

A new methodology to extract the blood vessels from retinal images using a nonparametric method was presented in [15]. The method had three objectives: the first objective was to detect blood vessels; the second objective wasto detect hemorrhagesand the third objective was to classify the different stages of diabetic retinopathy into normal, moderate and no proliferative diabetic retinopathy (NPDR).The authors applied the random forests technique based on the area and perimeter of the blood vessels and presence of hemorrhages. However the method fails if there were rare outcomes or rare predictors present because of the algorithmic rule according to bootstrap sampling. This makes it not ideal for identification of mild retinopathy where there is a lack of abnormal features present.

An adaptive automated segmentation approach for the optic disc based on mathematical morphology was suggested in [16]. The segmentation approach had two phases.First the detection of the optic disc location using the retinal vessels arcade and the second phase of detecting the boundary of the optic discussing a watershed algorithm that is transformed from external markers.

Baisheng et al. [17] proposed a variational model to segment and detect the optic disc, using sparse coding. This approach consisted of determining the evolution curve to the OD boundary, including the one with a low contrast, by applying a phase-based boundary energy algorithm. This was followed by a principle component analysis (PCA)-based form energy thatconstrainedthe evolution curve to a typical OD form andsuppressedthe negative impact of bright pixel interference. The third phase wasthe region energy term that drives the evolution curve to the boundary of the consistent regions. This method improved the sensitivity of the model to boundary noise and the initial position of the evolution curves.The main drawbacks of the level set methods is that they often require specification of variant parameters and it is tedious to tune them, especially when the desired contour does not correspond to a local energy minimum.

We propose a new Multistage Prediction algorithm to analyseretinopathy images and predict classification score for each image based on optic disc characteristics., The algorithm is based on the determining the amount of Red, Green, and Blue present in the optic disc as well as Yellow color from the inside and outside of the OD. This innovation addresses the importance of changing of the color composition within and outside of the optic disc associated with aging and disease progression.

Proposed Framework

The Proposed Multistage Prediction AlgorithmFramework(MSP1) classifies imagesin JPEG format, which has the advantage of smaller size images for faster processing and for telephthalmology. The availableimages were divided into 1) No DR, 2)mild DR, 3) moderate DR and 4) severe DRfollowing the National Health and Medical Research Council guidelines [18].

The proposed framework consists of three modules.Module 1 includes three maincomponents: the segmentation process of theimage, the dilation process, and the color normalization process. Module 2 includes three components consisting of determining the Mean, Standard Deviation, Variance, the histogram for each color band (Red, Green, and Blue), and the class scoring process. Finally, Module 3 is designed for the purpose of classification using the GMDH application. Figure 2 depicts the proposed Framework for the architecture.

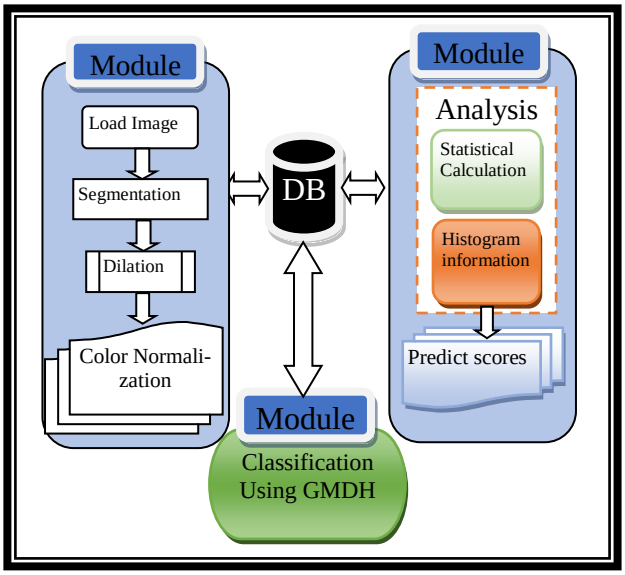

Fig. 2: Framework Proposed Architecture.

The implementation of the modules 1 and 2 is based on the MSP1 algorithm. The MSP1algorithm consists of 5 phases. From the first to the fourth phase, MSP1 implements Module 1, whereas the fifth phase is associated with Module 2, based on a new function named KHIndexing. KHIndexing is responsible for drawing the histogram for the main colors inside and outside of the OD, as well as calculatingthe Mean, Standard Deviation, and max colors.

Module 1

The main objective ofthis module is to prepare the image using sequence stages,carriedoutusing the fourth component (Optic disc segmentation, Dilation, Colorization index, and Statistical analysis).For details about this module, consider the previousresearch[19].

A 3*3 adaptive filterwas used when pre-processing the selected images. The threshold was adapted by pre-defining the value to be used when separating the optic disc from the rest of the image. To complete the optic disc and fill the gaps resultingfrom the adaptive shareholding step, a dilation process was carried out. To normalize the color, three steps were executed:First, grouping the neighbours with similar color by means of k-means clustering algorithm; second, encapsulating the blood vessels in the optic disc via two threshold values (495 and 640) determined through the processes of trial and error; and third, inversing the image for the purpose of obtaining the optic disc without any blood vessels.

Module 2

This module wasdeveloped from Section II of the previous paper [19]. Additions includedmaximum value, mean value, and the standard deviation value of the yellow color inside the OD, a histogram representationfor each band inside and outside of the OD, computing the yellow color ratio, computing OD indexing from formula 2 below, and the variancevalue of the imagecalculated by formula 3 to use the results as inputs in the ClassScore (formula no. 4). Finally, a score for each image is predicted.

$$Yratio=Y_i mean/Y_o mean \quad (1)$$

Where:

$Y_i$mean is a the mean of the yellow inside OD

$Y_o$mean is the mean of the yellow outside OD

$$ODIndex = \sum[\big((R+G+B)*0.3+(Region*0.1)\big)*Yratio)]/100 \quad (2)$$

$$\sigma^2 = \frac{\sum(X-\mu)^2}{N} \quad (3)$$

Where:
σ^2Is the variance

ClassScore=$\sigma^2 + Yratio$The processing steps are summarizedbelow for the yellow color (Algorithm 1) processing.

Algorithm(1) MSP1 Algorithm
Input : JPG Image
Output: Index Image, Variance, ClassScore
Begin
Phase 1 :Load JPG Image// load the JPG Image from DBwith pre-process

Image= $[Imagearray] * \begin{vmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{vmatrix}$ // pre-processed using a 3*3 adaptive filter

Phase 2 :Segmentation

$$\sum_{i=1}^{i=x-1} \sum_{j=1}^{j=y-1} if R(i,j) + G(i,j) + B(i,j) > 500, Image2 = RGB(R(i,j), G(i,j), B(i,j))$$

Mask(i,j)=1, Else Image(i,j)=RGB(0,0,0), Mask(i,j)=0,

Phase3 : Dilation

$$\sum_{i=1}^{i=x-1} \sum_{j=1}^{j=y-1} \sum_{ii=i-10}^{ii=i+10} \sum_{jj=jj-10}^{jj=jj+10} Sum = sum + mask(ii,jj),$$

If sum>threshold1 $\sum_{ii=i}^{ii=i+3} mask(ii,j) = 1, \sum_{jj=j}^{jj=j+3} mask(i,jj) = 1$,Else mask1(i,j)=0

Phase4 :Color Normalization

$$\sum_{i=1}^{i=x-1} \sum_{j=1}^{j=y-1} \text{if} mask1(i,j) = 1 \text{ thens} = 32, k = 0, found = 0, rc\ (i,j) = R(i,j)$$

While found=0 & k<256

{If R(i,j)>=k &R(i,j)<(k+s) then R(i,j)=k, If G(i,j)>=k &G(i,j)<(k+s) then G (i,j)=k

If B(i,j)>=k &B(i,j)<(k+s) then B (i,j)=k, found=1, k=k+s}

co=R(i,j)+G(i,j)+B(i,j), Image2(i,j)=RGB(R(i,j),G(i,j), B(i,j))*mask1(i,j)

If co>threshold2 and co<=threshold3 then Image5(i,j)= RGB(R(i,j),G(i,j), B(i,j))

Else Image6(i , j)= RGB(R(i,j),G(i,j), B(i,j))

Phase5 :Analysis and ClassScore

Call KHIndexing (image of in optic disk) , Call KHIndexing (image of output optic disk)

YRatio=meanYinOD/meanYoutOD , IndexImage=Ratio RGB*YRatio

$\sigma^2 = \frac{\sum(X-\mu)^2}{N}$, ClassScore=$\sigma^2$ + YRatio

End

Algorithm 1: A New Multistage Prediction Algorithm (MSP1).

The KHIndexing function is the main process in phase 5 of applying the MSP1 algorithmdepicted in Figure 3.

KHIndexing function

$$\sum_{k=0}^{k=255} histoR(k), histoG(k), histoB(k) = 0$$

$$\sum_{i=0}^{i=x} \sum_{j=0}^{j=y} if mask1(i,j) = 1\ then image1(i,j) = RGB\big(R(i,j), G(i,j), B(i,j)\big), histoR\big(R(i,j)\big) = histoR\big(R(i,j)\big) + 1, histoG\big(G(i,j)\big) = histoG\big(G(i,j)\big) + 1, histoB\big(B(i,j)\big) = histoB\big(B(i,j)\big) + 1, sumR = sumR + R(i,j), sumG = sumG + G(i,j), sumB = sumB + B(i,j) endif$$

sumY=(0.375*sumR+0.5*sumG+0.125*sumB)/3, maxR&minR=histoR(0), maxG&minG=histoG(0),

maxB&minB=histoB(0)$\sum_{k=1}^{k=255} if histoR(k) > maxR then maxR = histoR)(k), if histoG(k) > maxG then maxG = histoG,$ $if histoB(k) > maxB then maxB = histoB$

*maxY=(maxR+maxG)/2, meanR=sumR/(x*y), meanG=sumG/(x*y), meanB=sumB/(x*y), meanY=sumY/(x*y),*

$$\sum_{i=0}^{i=x} \sum_{j=0}^{j=y} if mask1(i,j) = 1\ then sR = sR + (R(i,j) - meanR)\text{^}2, sG = sG + (G(i,j) - meanG)\text{^}2,$$

sB=sB+(B(I,j)-meanB)^2, sY=sY+((sR+sG)/2)-meanY)^2 end if

*sdR=(sR/(x*y))^0.5, sdG=(sG/(x*y))^0.5, sdB=(sB/(x*y))^0.5, sdY=(sY/(x*y))^0.5*

$$\sum_{i=0}^{i=255} draw\ \ line(i, 130) - (i, 130 - histoR(i) mod\ 135 - 1), Red, Green, Blue$$

Fig. 3:KHIndexing Function.

Module 3

The major objective of this module is to classify the images using the class codes with GMDH. This method permits two-class and multi-class classification. GMDH usesmultiple inputs as a subset of components in a base function and a type of feedback loop or self-organization model that takes the optimum values at each step using a number of different algorithms until it reaches the final output using the Kolmogorov-Gabor polynomial. The classification methods are applied to the selected data to determine the accuracy ratio of each method. The highest accuracy ratio is then chosen automatically for classification. The results are then presented in the form of a Confusion matrix when the model is ready to be tested on all data. Such a step can be carried out by selecting the "Apply model" option.

2. Results and discussion

The dataset used in this research consisted of 157 images. Images were captured as JPEG format using the Canon CR6-45NM Fundus camera with a Canon 30D back and an IBM T40 computer for processing. The dataset was divided into no DR (48), mild (35) moderate (25) and severeDR (49) images. Figure 4 shows examples of the four types of images. Table 1 depicts the achieved prediction scores for each of the four classes.

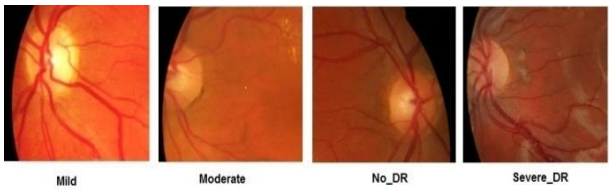

Fig. 4: Four Types Images [18].

Table 1: Classification of Images Based on the Class Score

No.	Types	Minimum Class Score	Maximum Class Score
3	No_DR	24.05	24.38
1	Mild	17.25	17.52
2	Moderate	21.11	21.31
4	Severe_DR	55.75	56.34

The range of scores of each retinopathy type is different from the others with no overlap identified using the minimum to maximum range of scores for each class.

The classification step applied theGMDHShellthree times; each time, the percentage of the DB used for training and classification waschanged using the same input variables. The three percentages used were 20%, 25%, and 30%. The next section explains the details of these experiments.

Experiment 1: This experiment was appliedon the DB to predict the available target, (Type Image) with four input (Table 2)variables.

Table 2: Variables Used with Experiment No. (1)

No.	Types	Variables
1	Prediction Variable	Type Image
2	Input Variable	GSD, YinSD, IndexOD, YrVariance

Where GSD is Green Standard Deviation color, YinSD is Yellow inside the OD Standard Deviation color, and YrVariance is Yellow ratio Variance.

Table 3 shows the confusion matrix and Table 4 depicts the result of this experiment with DB set at 20%. The correctly classified instances were30 with 96.8% accuracy whereas 1 image was incorrectly classified in the training set(3.2% accuracy).

Table 3: Experiment 1

Predicted class							
		1	2	3	4	Total	Recall
Actual class	1	35	0	0	0	35	1.000
	2	1	14	10	0	25	0.560
	3	0	0	48	0	48	1.000
	4	0	0	0	49	49	1.000
	Total	36	14	58	49	157	
	Precision	0.972	1.000	0.828	1.000		
	F-measure	0.986	0.718	0.906	1.000		
	Baseline	0.777	0.841	0.694	0.688	0.312	

	Accuracy	0.994	0.930	0.936	1.000	0.965

Table 4: Results of 3 Repeats of Experiment 1

Attempts	The percentages	Results
1	The percentage 20% of DB	96.5%
2	The percentage 25% of DB	87.2%
3	The percentage 30% of DB	87.2%

From Table 4, the best results are obtained when using 20% from DB with 96.5%.

Experiment 2: This experiment was applied on the DB to predict the available target, (Type Image) with four input variables. The input variables of Experiment 2 differed from the input variables of experiment 1. Table 5 shows the details of the input variables.

Table 5: Variables Used in Experiment 2

No.	Types	Variables
1	Predict Variable	Type Image
2	Input Variable	Gmax, Yinmax, IndexOD, YrVariance

Where Gmax is Green max color and Yinmax is Yellow inside the OD max color. Table 6 shows the confusion matrix attempt and Table 7 illustrates the result of this experiment. The correctly classified instances are 28 with 90.3% whereas the incorrectly classified instances are 3 with 9.7%.

Table 6: Experiment 2

Predicted class							
		1	2	3	4	Total	Recall
Actual class	1	35	0	0	0	35	1.00
	2	1	10	14	0	25	0.400
	3	0	0	48	0	48	1.000
	4	0	0	0	49	49	1.000
	Total	36	10	62	49	157	
	Precision	0.972	1.000	0.774	1.000		
	F-measure	0.986	0.571	0.873	1.000		
	Baseline	0.777	0.841	0.694	0.688	0.312	
	Accuracy	0.994	0.904	0.911	1.000	0.904	

Table 7: Results of Experiment 2

Attempt	The percentages	Results
1	The percentage 20% of DB	90.4%
2	The percentage 25% of DB	87.2%
3	The percentage 30% of DB	87.2%

In this experiment, results in Table 7 indicate that the best results from 3 attempts was 90.3% when using 20% from DB.

Experiment 3: This experiment was applied on the DB to predict the available target, (Type Image) with four input variables. The input variables of Experiment 3 differ from the input variables of Experiment 1and Experiment 2. Table 8 shows the details of the input variables.

Table 8: Variables Used In Experiment 3

No.	Types	Variables
1	Predict Variable	Type Image
2	Input Variable	Rmax, Bmean, BSD, Gmax, Yinmax, Yomax, IndexOD, YrVariance

Where Rmax is Red max color, Bmean is Blue mean color, BSD is Blue Standard Deviation color, Gmax is Green max color, Yinmax is Yellow inside the OD max color, and Yomax is Yellow outside OD max color.

Table 9 shows the confusion matrix and Table 10 illustrates the result of this experiment. The correctly classified instances were 154 with 98.1% accuracy whereas the incorrectly classified instances were 3 with 1.9%.

Table 9: Attempt 1 of Experiment 3

Predicted class							
		1	2	3	4	Total	Recall
Actual class	1	35	0	0	0	35	1.000
	2	1	24	0	0	25	0.960
	3	0	2	46	0	48	0.958
	4	0	0	0	49	49	1.000
	Total	36	26	46	49	157	
	Precision	0.972	0.923	1.000	1.000		
	F-measure	0.986	0.941	0.979	1.000		
	Baseline	0.777	0.841	0.694	0.688	0.312	
	Accuracy	0.994	0.981	0.987	1.000	0.981	

Table 10: 3 Attempts Results of Experiment 3

Attempt	The percentages	Sensitivity	Accuracy
1	The percentage 20% of DB	95%	96.8%
2	The percentage 25% of DB	97%	98.1%
3	The percentage 30% of DB	96%	97.5%

In this experiment, results in Table 10 indicate that the best result from 3 attempts was 98.1% when using 25% from DB. Table 11 depicts the comparison with others datamining approaches.

Table 11: Comparison of Results with Existing Approaches in Sensitivity and Accuracy

No	Approaches	Sensitivity	Accuracy
1	Verma K. [15] 2011	87.5%	88.46%
2	Sujith Kumar S B [20] 2012	94.44%	Not mentioned
3	Thammastitkul A. [21] 2016	85%	87%
4	Kamil R. [4] 2018	91%	94%
5	MSP1	97%	98.1%

Table 11 illustrates that the obtained result from the proposed MSP1 framework is significantly better than those obtained from previous work using different methods.

The three above-mentioned experiments, were executed using different features (see Table 2, 5, and 8) with each experiment repeated three times; each with a different rate of DB (20%, 25%, & 30%), the third experiment with (25% rate of DB) achieved the best results with accuracy reaching 98.1%. This was due to choosing suitable features in the third experiment. Such features affected effectively the process of obtaining high accuracy. By that, two objectives were achieved: the first was to determine the features that have the highest impact. The second objective was to obtain high accuracy in the results. Moreover, high percentages of sensitivity, precision, and F-measure were obtained reaching 97%, 97.3%, and 97.6%, respectively.

3. Conclusion

A recent model proposed by Alghamdi et al., 2016 [13] for optic disc segmentation resulted in an accuracy of 86.71% however these authors did not include detection and classification of diabetic retinopathy or DR progression. Verma et al (2011) [15] have applied a random forests technique classification for diabetic retinopathy based on the area and perimeter of the blood vessels and hemorrhages, and reported a sensitivity of 87.5% and accuracy 88.46%, using 65 images. While Kumar(2012) [20] used a 100 fundus images. The fundus images were classified into normal or retinopathy. Kumar obtained a sensitivity of 94.44% but accuracy was not reported. The main drawback of this methodarose from the method being dependent on vessel detection only.Rasha et al (2018) tested their system on a dataset consisting of 80 images, and obtained a 91% and 94% sensitivity and specificity respectively. The weakness of this approach was that images were classified into normal and abnormal DR in the first step, without passing the images through the all system processes, which leads to a lower accuracy when there is an incorrect classification in the initial step. The proposed MPS1 framework achieved an accuracy of 98.1% for DR classification. This proposed framework was based on eight features (Rmax, Bmean, BSD, Gmax, Yinmax, Yomax, IndexOD, and YrVariance) obtained from the optic disc image. The framework is robust for DR detection using a laptop or via the internet and therefore is of use in rural and remote areas for diabetic retinopathy screening.

Acknowledgement

This work is supported by the school of community health at Charles Sturt University as well as Endeavour fellowship program in Australia. Authors are thankful Mustansiriyah University in Iraq, the school of community health at Charles Sturt University for providing the facility to carry out the research work.

References

[1] KantardzicM.,"Data Mining: Concepts, Models, Methods, and Algorithms". 2nd Edition, Hardcover, 552 pages, Wiley-IEEE Press, ISBN: 978-0-470-89045-5. 2011.https://doi.org/10.1002/9781118029145.

[2] Han J., KamberM, and Pei J., "Data Mining Concepts and Techniques". Third edition, Morgan Kaufmann publishers as anreprint by Elsevier, USA, 2012.

[3] GMDH website, "GMDH algorithms and the algorithms of GMDH type". http://www.gmdh.net/GMDH_typ.htm. 2018.

[4] Mingzhu Z., Changzheng H., and Panos L., "A D-GMDH Model for Time Series Forecasting". Expert Systems with Applications Journal,Vol. 39, Issue 5. (April 2012), 5711-5716. DOI=http: // dx.doi.org /10.1016 / j.eswa. 2011.11.100. 2012.

[5] GMDH website, "GMDH Shell Documentation",http://d.gmdhshell.com/docs/concepts. 2018.

[6] Abdul Aziz A.andRehman A., "Detection of Cardiac Disease using Data Mining Classification Techniques". International Journal of Advanced Computer Science and Applications, Vol. 8, Issue 7, Pages 256-259. 2017.

[7] Kavakiotis I., Tsave O., Salifoglou A., Maglaveras N., Vlahavas I., ChouvardaI.,"Machine Learning and Data Mining Methods in Diabetes Research". Computational and Structural Biotechnology Journal. 15:104-116. doi:10.1016/j.csbj.2016.12.005.2017.

[8] KamilR., Al-Saedi K. H., Al-Azawi R., "An Accurate System for Measure the Diabetic Retinopathy by Using SVM Classifier".Ciência e TécnicaVitivinícola Journal, 2018. Vol. 33, Issue 1, Pages 135-139. 2018.

[9] Gulshan V, Peng L, Coram M, et al.,"Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs". JAMA,Vol. 316, Issue 22, Pages 2402–2410.https://doi.org/10.1001/jama.2016.17216.

[10] Mahendran G., Dhanasekaran R., "Investigation of the Severity Level of Diabetic Retinopathy using Supervised Classifier Algorithms". Computers & Electrical Engineering Journal, Vol. 45, Pages 312-323, https://doi.org/10.1016/j.compeleceng.2015.01.013.

[11] KamilR, Al-Saedi K. H., Al-Azawi R., "A Novel Approach for Optic Disc Detection in RGB Retinal Fundus Images". International Journal of Science and Research, Vol. 6, Issue 8, Pages 1263 – 1268, http://www.ijsrpublications.com/ijsr.net/archive/v6i8/v6i8.php, 2017.

[12] Almazroa A., Burman R., Raahemifar K., Lakshminarayanan V., "Optic Disc and Optic Cup Segmentation Methodologies for Glaucoma Image Detection: A Survey". Journal of Ophthalmology, Vol. 2015, Article ID 180972, Pages 1-28.

[13] Alghamdi, H., Tang H., Waheeb, S., and Tunde, P, "Automatic Optic Disc Abnormality Detection in Fundus Images: A Deep Learning Approach". In Proceedings of the Ophthalmic Medical Image Analysis Third International Workshop, Vol: 3, Pages 17-24, Athens, Greece.

[14] Sathiyamoorthy C. A. and Kulanthaivel, G., "Design and Develop a Computer Aided Design for Automatic Exudates Detection for Diabetic Retinopathy Screening". Journal of Engineering Science and Technology Vol. 11, Issue 4, Pages 605 – 618, 2016.

[15] Verma, K.; Deep, P.; Ramakrishnan, A. G., "Detection and Classification of Diabetic Retinopathy using Retinal Images". India Conference (INDICON), 2011 Annual IEEE, Vol., Pages 1-6, 16-18 Dec. 2011, URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6139346&isnumber=6139311https://doi.org/10.1109/INDCON.2011.6139346.

[16] Welfer D., Scharcanski J., Cleyson M. K., Dal Pizzol M., LudwigW.B., Marinho R., "Segmentation of the Optic Disk in Color Eye Fundus Images using an Adaptive Morphological Approach". Computers in Biology and Medicine Journal, Volume 40, Issue 2, Pages 124-137.

[17] DaiB. Xiangqian Wu, Wei Bu, "Optic Disc Segmentation Based on VariationalModel with Multiple Energies". Pattern Recognition Journal, Volume 64, Pages 226-235.

[18] Mitchell P., Foran, S.Australian Diabetes Society, National Health, and Medical Research Council, "Guidelines for the management of diabetic retinopathy", publisher National Health and Medical Research Council" ISBN 1741866723- 1741866715, https://nla.gov.au/nla.cat-vn4470187, 2008.
JelinekH.F, Al-SaediK. H. and BäcklundL.B. "Computer Assisted "Top-Down" Assessment of Diabetic Retinopathy". World Congress on Medical Physics and Biomedical Engineering, September 7 - 12, 2009, Munich, Germany, IFMBE Proceedings Springer Berlin Heidelberg, ISSN 1680-0737, Vol. 25, Issue 11, Pages 127-130. 2009.

[19] Kumar SB and Vipula S., "Automatic Detection of Diabetic Retinopathy in Non-dilated RGB Retinal Fundus Images". International Journal of Computer Applications, Vol. 47, Issue 19, Pages 26–32, 2012.

[20] Thammastitkul A. and UyyanonvaraB, "Diabetic Retinopathy Stages Identification Using Retinal Images". Annual Internatoinal Conference on Intelligent Computing, Computer Science & Information Systems, April 28-29 (2016), Pattaya, Thailand Pages 20–23, 2016.