EFFICIENT FEATURE EXTRACTION OF THE OPTIC CUP AND DISC FROM EYE IMAGES FOR OPTIMAL DETECTION OF GLAUCOMA

PERIYAKARUPPAN K[1*], KAVITHA M S[2], SHANTHINI J[3]
[1*]Professor, Department of Computer Science & Engineering,
SNS College of Engineering, Coimbatore-641107, TamilNadu, India
E-Mail: kperiyakaruppan@gmail.com

[2]Associate Professor, Department of Computer Science & Engineering,
SNS College of Technology, Coimbatore-641035, TamilNadu, India
E-Mail: drmskavitha@yahoo.com

[3]Professor, Department of Computer Science & Engineering,
Dr. N.G.P. Institute of Technology, Coimbatore-641048, TamilNadu, India
E-Mail: drjshanthini@gmail.com

Abstract

Glaucoma is a progressive eye disease that can cause irreversible vision loss if left untreated. Regular glaucoma screenings are important because the disorder is irreversible in its advanced stages. In this paper, we use a feature extraction process after pre-processing, where the feature extraction involves feature selection, ranking and sampling. The simulation is conducted on three different datasets to test the efficacy of the feature selection model. The results of simulation shows that the proposed weight based ranking provides an improved feature selection than the existing models.

Keywords: Glaucoma, Eye Disease, Screening, Optic Cup, Feature Extraction
DOI Number:10.14704/nq.2022.20.8.NQ44536 **NeuroQuantology 2022; 20(8): 5095-5105**

1. Introduction

Vision loss (often on both sides of the body) and eventual blindness are the end outcomes of glaucoma, a group of progressive neuropathies. Retinal ganglion cell death and optic nerve damage are hallmarks of glaucoma. What more, in a number of nations, it the leading reason why people sustain permanent eye damage or go blind [1].

Two common forms of glaucoma, each of which may be traced back to specific causes, are as follows:

Open-angle glaucoma (OAG): The vast majority of glaucoma cases are due to open-angle glaucoma (OAG). Blockage of the eye natural drainage channels are a common cause of elevated intraocular pressure. The iris and cornea form a healthy open-angle when the angle between them is wide and unobstructed.

Angle-closure glaucoma (ACG): Angle-closure glaucoma is a kind of glaucoma characterised by an abrupt increase in intraocular pressure. Due to blocked drains, this area has risen to this level. Acute glaucoma, also called narrow-angle glaucoma, differs from OAG in that the gap between the eye cornea and iris narrows rather than widens.

Open-angle glaucoma (OAG) and angle-closure glaucoma are the two most common forms of the disease (ACG). Congenital glaucoma occurs in infants due to the faulty eye drainage model, while normal-tension glaucoma occurs when the optic nerve is damaged despite the eye pressure not being extremely high; the reason of this damage is unknown. Most of them have not been investigated extensively, and their underlying causes remain a mystery.

Since OAG and ACG are often assessed independently, [2] reports that the average worldwide incidence of OAG is 1.96% and the average worldwide prevalence of ACG is 0.69%. Despite this, blindness after a diagnosis still occurs at a rate of over 1% annually [3].

A rise to 111.8 million glaucoma sufferers is predicted by 2040 [4] due to the ageing of the global population. Physical functioning, Quality of life and mental health all suffer as a result of glaucoma-related blindness and visual impairment, which is a leading cause of disability around the world [5] [6].

Examining medical images as part of a diagnostic strategy for glaucoma is currently making substantial headway against more conventional procedures. The optic nerve head (ONH), cup, peripapillary atrophy, retinal nerve fibre layer, and other aspects of retinal anatomy must all be checked in such cases. The optic nerve head (ONH) appears as a bright, rounded spot in fundus images; a smaller, spherical area inside the ONH is called a cup. The optic nerve and its surrounding structures form the complex [7]. Peripapillary atrophy is characterised by a crescent-shaped loss of tissue just outside the outer nuclear layer (ONH). The retina nerve fibre layer, easily recognised by its striated white coloration, extends beyond the optic nerve head (ONH). There are two types of computer vision algorithms that can be used to detect glaucoma diagnostic signals: segmentation and feature extraction [8].

Methods like thresholding and active contours are widely used due to their ability to localise and segment an image. Furthermore, novel techniques like fuzzy c-means have been developed. Misclassification of the ONH region is a common result of using these approaches, necessitating the employment of morphological techniques for correction. However, as the ONH texture is the most important property in glaucoma detection, many techniques have been devised to extract this feature [9]. As a result, many of these methods have been developed. Wavelet and higher-order spectral techniques are commonly used for feature extraction (HOS). Some examples of prior work utilising a support vector machine (SVM) classifier with discrete wavelet transform (DWT) are provided [10].

Principal Component Analysis (PCA) is used by certain authors, including [11], to reduce the number of variables in the characteristics produced by DTW. Only a small sample of the many studies that have led to major improvements in glaucoma detection methods in recent years are presented here. However, substantial progress has been made in the analysis of medical images generally thanks to the introduction of machine learning that can extract the features for accurate diagnosis [12]. To be effective, these strategies need to be applied to a dataset that contains a wide range of images taken by people with glaucoma and controls who do not have the illness.

To help in glaucoma diagnosis, neural network-based algorithms can automatically analyse the images and extract the relevant features. The functioning of such systems requires additional processes, such as the preprocessing phase, the careful selection of the network design, and the training phase (which may include monitoring). Despite this, research in [13] have shown that using neural networks with medical images can

*e*ISSN1303-5150

lead to conclusions that are superior to those provided by conventional diagnostic approaches (and not just for glaucoma diagnosis). Based on this premise, the current work entails using machine learning algorithms to fundus images collected from different hospitals in order to develop a glaucoma diagnostic.

2. Related works

People over the age of 40 who suffer from glaucoma, an eye illness, have a one in two chance of going blind permanently. Automatic glaucoma identification was proposed by Carrillo et al. [14]; this would be facilitated by a computer tool developed by the same authors. Fundus imaging has surpassed all other methods for diagnosing glaucoma because to its low price, small size, and portability. The authors uncovered a new level of progress in disc segmentation by comparing their findings to those in the scholarly literature. We used thresholding for disc segmentation, edge detection for vessel segmentation, and a method that combined vessel and cup intensities for cup segmentation. When looking at a collection of fundus images, 88.5% of the time glaucoma was correctly identified. The authors [14] argue that in order to provide a meaningful evaluation, future studies should employ a larger database of fundus images.

A method for detecting glaucoma was devised by Sengar et al. [15], which involves analysing fundus images for the presence or absence of haemorrhages in a predetermined region close to the optic disc. If glaucoma is suspected, this test could help confirm the diagnosis. It has been shown that this strategy is both accurate in identifying glaucoma cases and economical in terms of the amount of computing power [19][20], network bandwidth [21][22] needed to do so.

For their method of glaucoma diagnosis, Salam et al. [16] proposed an algorithm that integrates CDR with hybrid textual and multiple intensity features. The CDR and classifier results are incorporated into the refinement of the image arranger. The results demonstrate that the proposed fusion enhanced the sensitivity, specificity, and accuracy of the combined test.

Automatic glaucoma identification, proposed by Poshtyar et al. [17] would do away with the requirement for a trained professional or expensive machinery. Quantitative testing is essential for a correct diagnosis of glaucoma. Over all, the approach taken here was fruitful. The methods of detecting glaucoma in its earliest stages were proposed by Ahmed et al. [18] This study examine glaucoma through the lenses of the CDR and NRR classifications to see where it lies on the matrix.

3. Proposed Method

Using a system that integrates image-based features with clinical measurement characteristics driven by past information in the area, we offer a method for the automatic detection of glaucoma. A new method for extracting these features from images was developed using the IFOV feature model as its basis. Screening for glaucoma should account for even subtle differences in receptive fields, since this increases the usefulness of the recovered data and ensures more accurate diagnoses.

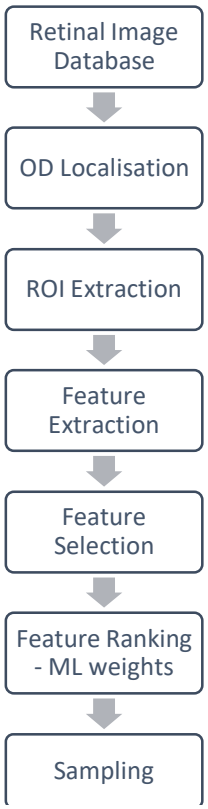

Figure 1: Proposed Model

Pre-Processing

When using the classification subsystem, we apply the same processing to the dataset images as we did when using the segmentation subsystem, but we scale them down to the standard resolution of 224×224 pixels.

Feature model

For the extraction of retinal images, we use Increasing Field of View (IFOV) Model. This model classifies the field into scale area that includes OD region, OC area, ROI and fundus image. The study uses GLCM and Gabor transform to extract the visual features. As the ophthalmologist moves clockwise around the lesion in the patient eye, the patient visual field expands, as seen in Figure 9. This enables feature extraction from images at a wide range of granularities.

Feature selection

We counted a grand total of 1600 individual characteristics using our method. The computational efficiency and quality of our classifiers may suffer, though, if we use too many of these features. This highlights the significance of the most relevant features and narrowing the focus to fewer dimensions. In order to find the optimal feature set, we employ a wide range of feature-selection methods in this study. This allows us to achieve high-quality classification results with a minimum of input features.

Step 1. Low-variance features are eliminated in the first stage using variance analysis. This is done due to the fact that such characteristics typically exhibit values that are quite consistent and, as a result, do not contribute much to the categorization of glaucoma. So as to boost computational efficiency without sacrificing performance, we lowered the variance

threshold of rejected features to 0.001. This allows us to filter out the vast majority of qualities with minimal volatility.

Step 2. We use Pearson correlation coefficient to examine the closeness of the two characteristics. Our investigation revealed that a high degree of association exists between two features if their Pearson correlation coefficient was more than 0.9. These unnecessary features will be eliminated because they slow down the categorization process and provide no useful purpose.

Step 3. Feature redundancy exists after correlation and variance analysis, which significantly affects the growth of the glaucoma classifier. In order to improve classification accuracy while reducing the total features, which is a crucial feature set for the select and choose relevant feature subsets.

Ranking – ML Weighting

The classifier efficacy will be drastically improved by combining feature data from many ranking techniques. We ranking techniques to rank the features, and then use only the highest-scoring features to train the classifier. Data-adaptive weighting of input attributes is also available to users. The contributions of SOM qualities and the expertise of interpreters are both necessary to calculate the weight matrix **W**. In a SOM model, the contribution of the n^{th} input attribute, denoted by the symbol ω_n, is calculated as follows, given *P* prototype vectors and *N* input attributes:

$$\omega_n = \sum_{p=1}^{P} d_p \frac{|\xi_{pn}|}{\sum_{n=1}^{N} |\xi_{pn}|}$$

and

$$d_p = \frac{h_m}{M}$$

where

h_m – multi-attribute samples for training nearer to the p^{th} vector,

M - multiattribute samples for training,

d_j - training sample density that gets allocated to the vector *p*, and

ξ_{pn} - p^{th} vector value along the n^{th} attribute dimension.

The contribution of the target attribute at a prototype vector is likely to be large if a large number of training samples cluster near that vector and that vector has a high value in the dimension of the target attribute. After that, we sum up all of the prototype vectors to get an overall contribution that the target attribute has made to the SOM model. We advise running a search that exhaustively explores all of the permutations that can occur between three or more qualities to achieve the most precise evaluation of the overall best contribution for a certain collection of attributes.

$$S = \sum_{i=3}^{N} \frac{N!}{i!\,(N-i)!},$$

Once given each characteristic combination a weight based on how often it appears, you may add them all together to get your final result.

$$\omega_i = \sum_{l=1}^{S} N_l \widetilde{\omega}_{il},$$

where

! - factorial operation,

S - total models for search,

N_l - l^{th} attribute combination,

$\widetilde{\omega}_{il}$ - i^{th} attribute contribution to the l^{th} model, and

ω_i - final i^{th} attribute contribution.

Even though this approach requires running the model multiple times with different

*e*ISSN1303-5150

combinations of input attributes, it is fundamentally a parallel problem, and the additional calculation time that is required in comparison to traditional neural network model is negligible with a sufficient number of threads or processors.

In our method, we give more weight to characteristics whose distributions are less skewed in absolute terms (have high kurtosis values) than to those whose distributions are more skewed (low kurtosis). When discussing the skewness of a distribution, what is typically meant is the third instant of the *a*-standard variable score.

$$s(a) = \mathbb{E}\left[\left(\frac{a - \mathrm{a}}{\sigma_a}\right)^3\right]$$

where

$\bar{a}$ - variable *a* mean,

σ_a - standard deviation, and

$\mathbb{E}$ - Expectation.

This is similar to the definition of kurtosis, which states that it is the 4th moment of the standard score of *a* variable.

$$k(a) = \mathbb{E}\left[\left(\frac{a - \mathrm{a}}{\sigma_a}\right)^4\right]$$

After the attribute contribution **ω** has been calculated, the skewness and kurtosis measurements are actually taken. We then modify the skewness and kurtosis once the **ω** calculations are done to bring them into the range of zero to one. After re-normalizing the data with a z-score and adding skewness and kurtosis as weights to the previously established **ω**, 5100

$$w_i = \left(3 - \frac{|s_i| - \min\limits_{i=1,N}|s_i|}{\max\limits_{i=1,N}|s_i| - \min\limits_{i=1,N}|s_i|} - \frac{k_i - \min\limits_{i=1,N} k_i}{\max\limits_{i=1,N} k_i - \min\limits_{i=1,N} k_i}\right)\omega_i$$

$$\widehat{w}_i = \frac{w_i - \overline{w}}{\sigma_w}$$

Here,

w_i - i^{th} attribute weight prior normalization of *z*-score,

$\overline{w}$ - $w_i{}^{\text{th}}$ mean,

σ_w - standard deviation, and

$\widehat{w}_i$ - i^{th} attribute post the normalization of z-score.

We will assume that skewness and kurtosis are well balanced given that they have both been normalised to a value between zero and one. Similar to how $\widehat{w}_i$ was normalised, z-score is used to further normalise w_i, producing $\widehat{w}_i$. Therefore, it is not as important to know w absolute value. Finally, we restrict the weight to being a sigmoid function and define the members of the diagonal weight matrix **W** as a sigmoid function with values between zero and two.

$$W_{ii} = \frac{2}{1 + e^{-\widehat{w}_i}}$$

Sampling Approach

If there is a significant difference between the sample sizes of different groups within a dataset and the overall size of the dataset, we claim that the dataset is unbalanced. In this case, the representation of samples from minority groups is disproportionately low compared to the representation of samples from groups that are statistically more likely to be the norm. This causes the trained classifier to give preference to the category that includes the most examples, which always yields subpar performance. One of the disadvantages of this study is that the dataset utilised in the

experiment has an issue with uneven sample distribution. Using features extracted from the data of minority ethnic groups, our method generates new feature samples in a flexible manner suited to the minority class. The divide between the healthy and glaucoma groups is thus reduced.

4. Results and Discussions

We will evaluate the chi-square ([χ^2]), mutual information, and GBDT techniques for attribute ranking and draw comparisons between them. The chi-squared test utilises the statistic of interest, χ^2, to determine the significance of a feature. The degree to which two variables are interdependent can be quantified by computing their mutual information. The relevance of the features is weighed using the mutual information. Regression, classification, and feature ranking are just some of the many tasks that benefit greatly from GBDT unified model of decision trees. When weighing the significance of each feature, the GBDT takes into account the drop in pollution that follows node splitting. Because of this, the features can be ranked. Using a measure called mean square error (MSE) to ascertain data contamination.

Our ML-based disc and cup segmentation is compared to a previous result on the same fundus imaging data sets. There are two very important aspects of our work that you should know about. 75% of each dataset are used for training, while another 25% are used for validation.

Dataset:

An essential part of this procedure is being able to compare our results to those of the preceding study. These two collections of data feature images of people whose eyes have been annotated to show whether or not they have glaucoma. As part of the labelling procedure, samples from the dataset are inspected by an expert to ensure they meet quality standards. This expert verifies that all images in each dataset are of either a patient with glaucoma or a healthy person serving as a control. Ophthalmologist annotated images form the basis for the disc and cup.

The study has used the DRIONS dataset, however it is not applicable here because it lacks cup segmentation data. Because of this, it was excluded from the existing models. The DRISTI-GS collection provides 101 disc and cup tagged colour fundus images, whereas the RIM-ONE dataset contains 151 such images.

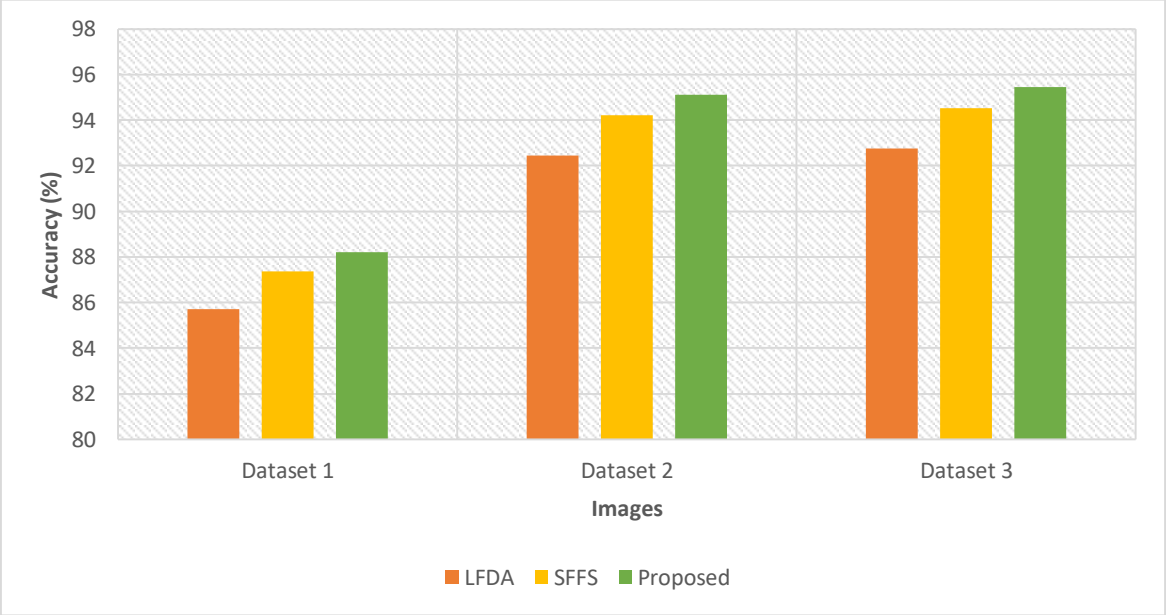

Figure 2: Accuracy of Feature selection in Training

Figure 3: Accuracy of Feature Selection in Testing

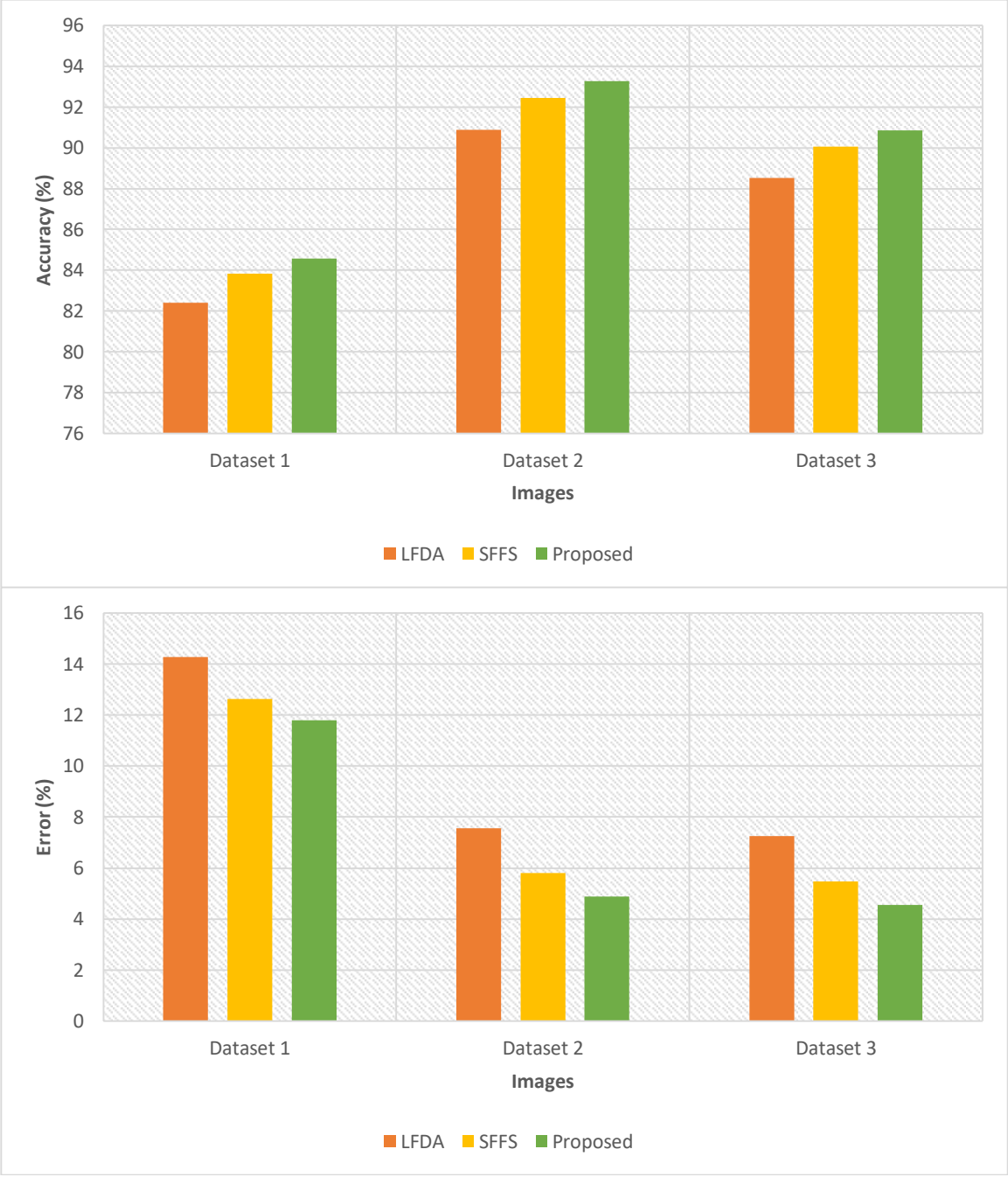

Figure 4: Error in Feature selection in Training

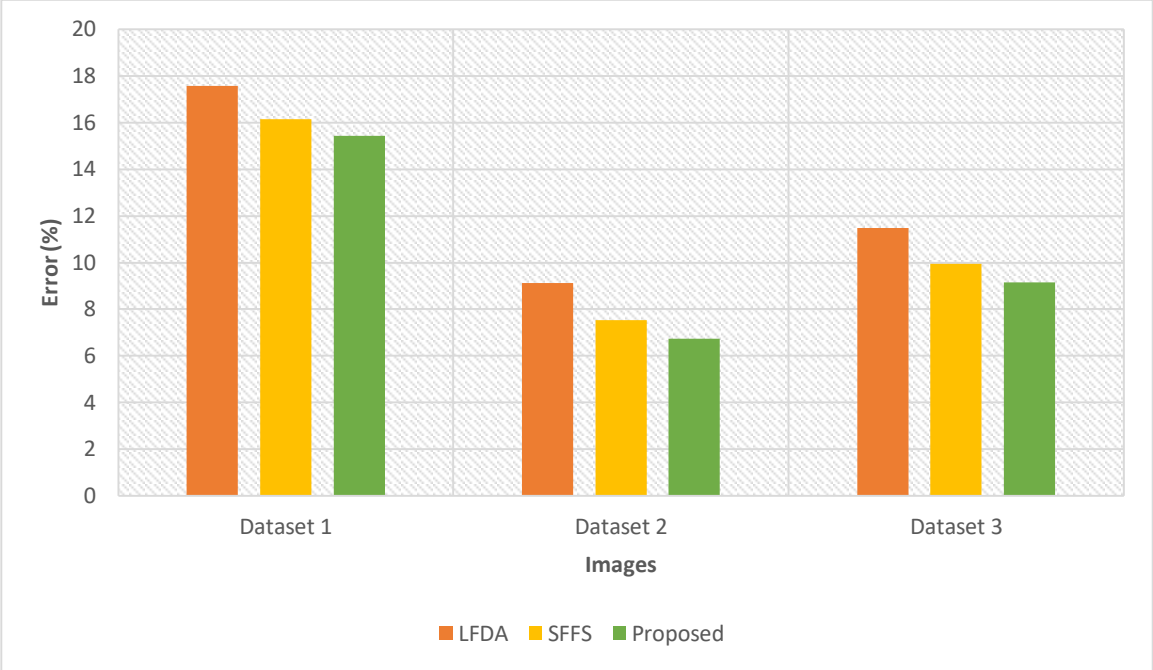

Figure 5: Error in Feature Selection in Testing

5. Conclusions

Since it provides a visual image of the human body, medical imaging technology is crucial to the operation of public healthcare systems, where it may be used to monitor and track a wide range of ailments. Multiple healthcare systems use digital image processing and computer vision technology to diagnose patients. Glaucoma is a degenerative and chronic eye disease that causes irreversible damage to the optic nerve and, eventually, blindness. In this paper, we use a feature extraction process after pre-processing, where the feature extraction involves feature selection, ranking and sampling. The results of simulation shows that the proposed weight-based ranking provides an improved feature selection than the existing models. It is seen from the results that the proposed model achieves higher degree of feature selection accuracy than the existing models.

References

[1] Betzler, B. K., Young, S. M., & Sundar, G. (2022). Intraocular Pressure and Glaucoma in Thyroid Eye Disease. *Ophthalmic Plastic and Reconstructive Surgery*, *38*(3), 219-225.

[2] Baudouin, C., Kolko, M., Melik-Parsadaniantz, S., & Messmer, E. M. (2021). Inflammation in Glaucoma: From the back to the front of the eye, and beyond. *Progress in Retinal and Eye Research*, *83*, 100916.

[3] Iraha, S., Takihara, Y., Urahashi, Y., Watanabe, T., Nakamura, K., Urahashi, M., ... & Inoue, T. (2022). Factors associated with the surgical outcomes of Baerveldt glaucoma implant for open-angle glaucoma, an age-related eye disease. *Scientific reports*, *12*(1), 1-9.

[4] Sahlu, M., & Giorgis, A. T. (2021). Dry eye disease among Glaucoma patients on topical hypotensive medications, in a tertiary hospital, Ethiopia. *BMC ophthalmology*, *21*(1), 1-8.

[5] Voicu, L., & Salim, S. (2021). New strategies for the management of ocular surface disease in glaucoma

patients. *Current opinion in ophthalmology*, *32*(2), 134-140.

[6] Patil, N., Patil, P. N., & Rao, P. V. (2021). Convolution neural network and deep-belief network (DBN) based automatic detection and diagnosis of Glaucoma. *Multimedia Tools and Applications*, *80*(19), 29481-29495.

[7] Shin, Y., Cho, H., Jeong, H. C., Seong, M., Choi, J. W., & Lee, W. J. (2021). Deep learning-based diagnosis of glaucoma using wide-field optical coherence tomography images. *Journal of Glaucoma*, *30*(9), 803-812.

[8] Vyas, A. H., & Khanduja, V. (2021, December). A Survey on Automated Eye Disease Detection using Computer Vision Based Techniques. In *2021 IEEE Pune Section International Conference (PuneCon)* (pp. 1-6). IEEE.

[9] Singh, L. K., Khanna, M., & Thawkar, S. (2022). A novel hybrid robust architecture for automatic screening of glaucoma using fundus photos, built on feature selection and machine learning-nature driven computing. *Expert Systems*, e13069.

[10] Perez-Sanchez, A. V., Valtierra-Rodriguez, M. A. R. T. I. N., Perez-Ramirez, C. A., De-Santiago-Perez, J. J., & Amezquita-Sanchez, J. P. (2022). Epileptic Seizure Prediction Using Wavelet Transform, Fractal Dimension, Support Vector Machine, and EEG signals. *Fractals*, 2250154.

[11] Weiwei, H. (2022). Classification of sport actions using principal component analysis and random forest based on three-dimensional data. *Displays*, *72*, 102135.

[12] Walter, W., Haferlach, C., Nadarajah, N., Schmidts, I., Kühn, C., Kern, W., & Haferlach, T. (2021). How artificial intelligence might disrupt diagnostics in hematology in the near future. *Oncogene*, *40*(25), 4271-4280.

[13] Saravanan, V., Samuel, R., Krishnamoorthy, S., & Manickam, A. (2022). Deep learning assisted convolutional auto-encoders framework for glaucoma detection and anterior visual pathway recognition from retinal fundus images. *Journal of Ambient Intelligence and Humanized Computing*, 1-11.

[14] Carrillo, J., Bautista, L., Villamizar, J., Rueda, J., & Sanchez, M. (2019, April). Glaucoma detection using fundus images of the eye. In *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)* (pp. 1-4). IEEE.

[15] Sengar, N., Dutta, M. K., Burget, R., & Ranjoha, M. (2017, July). Automated detection of suspected glaucoma in digital fundus images. In *2017 40th International Conference on Telecommunications and Signal Processing (TSP)* (pp. 749-752). IEEE.

[16] Salam, A. A., Akram, M. U., Wazir, K., Anwar, S. M., & Majid, M. (2015, December). Autonomous glaucoma detection from fundus image using cup to disc ratio and hybrid features. In *2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)* (pp. 370-374). IEEE.

[17] Poshtyar, A., Shanbehzadeh, J., & Ahmadieh, H. (2013, December). Automatic measurement of cup to disc ratio for diagnosis of glaucoma on retinal fundus images. In *2013 6th International Conference on Biomedical Engineering and Informatics* (pp. 24-27). IEEE.

[18] Ahmad, H., Yamin, A., Shakeel, A., Gillani, S. O., & Ansari, U. (2014, April). Detection of glaucoma using retinal fundus images. In *2014 International conference on robotics and emerging allied technologies in engineering (iCREATE)* (pp. 321-324). IEEE.

*e*ISSN1303-5150 www.neuroquantology.com

[19] Sumathi A, Saravanan V, “Bandwidth based vertical handoff for tightly coupled WiMAX/WLAN overlay networks”, Journal of Scientific & Industrial Research, vol. 74, pp. 560-566, 2015.

[20] Saravanan V & Sumathi A, “Dynamic handoff decision based on current traffic level and neighbor information in wireless data networks”, Fourth International Conference on Advanced Computing (ICoAC), IEEE, pp. 1-5, 2012.

[21] Shakir Khan, V. Saravanan*, Gnanaprakasam C. N, T. Jaya Lakshmi, Nabamita Deb, Nashwan Adnan Othman, "Privacy Protection of Healthcare Data over Social Networks Using Machine Learning Algorithms", Computational Intelligence and Neuroscience, vol. 2022, 8 pages, 2022.

[22] Dr. Grandhi Suresh Kumar, Dr. Pratik Gite, Dr. M. Prasad, Dr. Madiajagan M, Dr. S. Selvakanmani, Dr. V. Saravanan, Design of Autonomous Production Using Deep Neural Network for Complex Job-Shops, International Journal of Grid and Distributed Computing, Vol. 14, No. 1, pp. 813-824, 2021.