JOURNAL ON COMMUNICATIONS

ISSN:1000-436X

REGISTERED

Scopus

www.jocs.review

An Amalgamate Learning Model for Detection of Retinal Diseases Using OCT Images

Rajababu Pukkala[1], *S. Jhansi Rani*[2]
[1]*Research Scholar, Department of Computer Science and Systems Engineering, Andhra University, India*
[2]*Associate Professor, Department of Computer Science and Systems Engineering, Andhra University, India .*

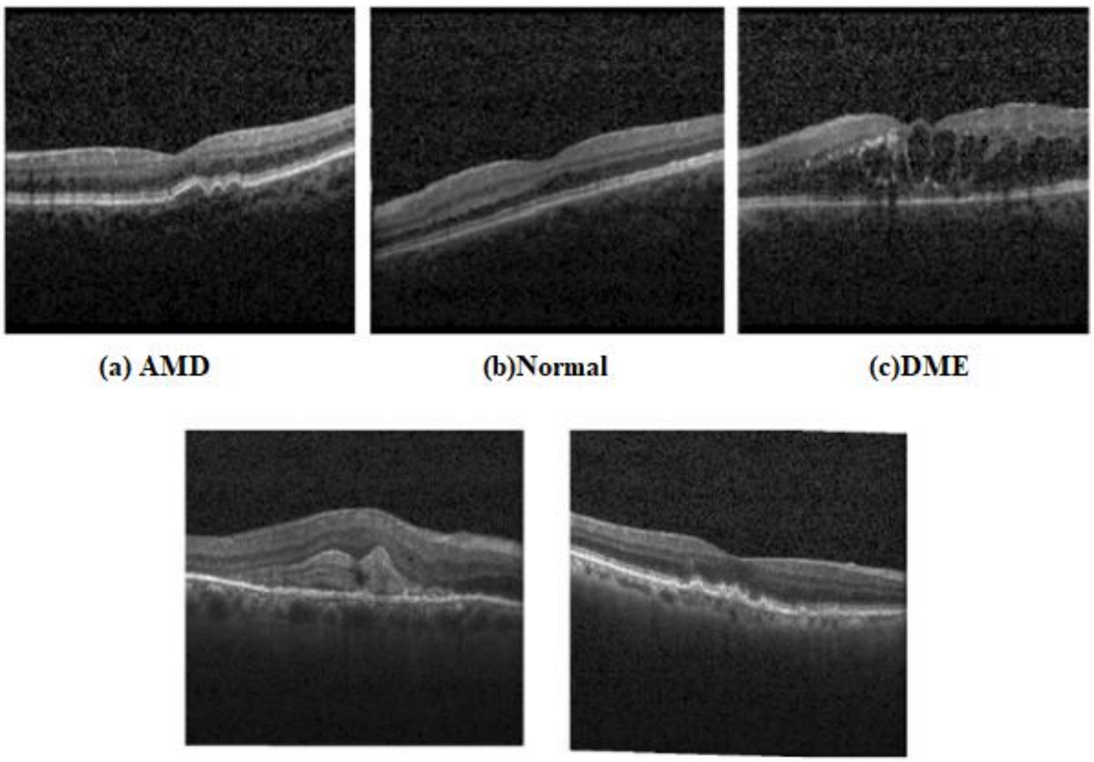
Abstract: Retinal diseases pose significant risks to human vision and require timely and accurate diagnosis for effective management. Optical Coherence Tomography (OCT) has emerged as a valuable imaging modality for the early detection and monitoring of various retinal pathologies. In this study, an Amalgamate Learning Model (ALM) introduced to detect and classify retinal diseases using OCT images automatically. The ALM is designed as a hybrid approach that combines the strengths of different noise removal techniques and Generative Adversarial Networks (GANs) + CNN algorithms. The model obtained the hierarchical features learned by the CNNs from OCT images and captured the sequential patterns within the OCT images using ALM. The pre-trained model DenseNet169 is fine tuned on the OCT images dataset. The DenseNet-169 assists the strategy to capture the features from the OCT images dataset and then transfer to the proposed model. Although the classification model diagnoses three typical types of retinal diseases, such as AMD (Age-related Macular Degeneration), DME (Diabetic Macular Edema), and DRUSEN, which can be diagnosed and monitored using OCT (Optical Coherence Tomography) imaging. We had ophthalmologists with expertise meticulously label the dataset to guarantee precision of diseased labeling. We have performed a bunch of experiments to tune the ALM hyperparameters and achieve good performance. The model demonstrated state-of-the-art accuracy for retinal disease detection and classification against solo CNN and RNN models. The ALM showed high sensitivity and specificity, and it is considered a reliable method for early diagnosis and treatment. In addition, an interpretability investigation was conducted to obtain insights into the reasoning behind the ALM. This further facilitates model interpretations and increases the confidence of clinicians in its predictions. The proposed Amalgamate Learning Model is a powerful and effective method for the detection and classification of retinal diseases with OCT images. The joint modelling of spatial and sequential information on OCT scans in POL-Net also contributes the excellent performance. The model's accuracy and interpretability can significantly enhance the clinical workflow and improve patient outcomes. Future work will focus on deploying the ALM in real-world clinical settings and expanding its application to other medical imaging tasks.

Keywords: Retinal Diseases, Machine Learning, Deep Learning, OCT Images.

I. Introduction

Eyes are essential part of human life, each and every person rely on the eyes to see and sense the world around them. One of the most vital senses is sight because it accounts for 80% of all information, we take in. By taking proper care of eyes, we will lower the risk of becoming blind and losing vision, while also keeping an eye out for any developing eye-conditions like glaucoma and cataracts [1]. Most people experience eye issues at some point of time. Some of the eye issues are minor and simple to cure at home which will go away on their own, other major eye issues need assistance from the expert doctors. When these eye diseases are accurately diagnosed at an early stage, only then the progression of these eye diseases can be stopped. These eye diseases have a wide range of visually discernible symptoms. To accurately diagnose eye illnesses, it is required to analyse a wide range of symptoms. In this paper, our proposed model analyses and classifies eye diseases namely cataracts, crossed eyes, bulging eyes, uveitis and conjunctivitis [2].

Vision is one of the most important human senses, lack of which can affect productivity and independence of a person. Retinal diseases affect millions of people and may result in loss of vision if the disease is not diagnosed and treated early [3]. Example of retinal diseases include diabetic retinopathy, age related macular disorder, glaucoma etc. Early treatment options that are available may cure or slow the onset of the disease. Patients treated get several more years of vision in their life. In India, although there are a number of hospitals and eye clinics in the cities, the doctor to patient ratio is still low [4]. In rural areas, there is lack of both infrastructure and availability of ophthalmologists. Even community outreach programs are handicapped by the lack of trained personnel to effectively screen patients in rural areas. Remote image acquisition and diagnostic, require high cost and infrastructure [5]. With progress in technology and image analysis, it is possible to automate the process of disease detection and refer the patient to the doctor for further consultation [6]. A number of such clinical decision support systems have been developed especially to diagnose diabetic retinopathy, age related macular disorder using advances in digital image processing and machine learning.



(a) AMD (b)Normal (c)DME

(d)CNV (e)Drusen

Figure: 1 (a), (b), (c), (d) and (e) Showing the types of Retinal Diseases (a) AMD (b) Normal (c)DME (d) CNV and (d) Drusen

II. Literature Survey

Many eye conditions, including trachoma, cataracts, and corneal ulcers, can impair vision. Only when these eye illnesses are effectively diagnosed at an early stage can progression be stopped. These eye illnesses have a wide range of visually discernible symptoms. To accurately diagnose eye illnesses, it is required to analyze a wide range of symptoms. Therefore, utilizing machine learning techniques like deep convolution neural network (DCNN) and support vector machine, paper [7] suggested a novel strategy in to create an automated eye illness recognition system using visually observable symptoms, from experimental findings it is observed that the DCNN model performs better than SVM models. In paper [8] the author have used a deep neural network model to discriminate between different diseases like diabetic retinopathy, which aids in the early detection of glaucoma and diabetic retinopathy, and high resolution retina images taken under a variety of imaging settings. In terms of screening Eye Disease Identification using Deep Learning, it may prompt patients to contact an ophthalmologist. The created model has a lower level of complexity and achieved an accuracy of 80%. The author of the paper [9] developed a method for automatically classifying any retinal fundus image as healthy or sick using a deep learning model. They created a system named LCD Net using CNN that was able to do the binary classification.

Two sources of retinal fundus pictures were used to construct a total of eight testing datasets. Using existing datasets, image preprocessing methods, deep learning models, and performance evaluation criteria, the author of the paper [10] developed a model for the automated identification of diabetic eye illness. It includes works that used TL, built DL network architecture, and used a combined DL and ML approach in terms of classification algorithms. From medical photos, we may deduce that CNN is now the most popular deep neural network, especially for the identification of diabetic eye illness and the diagnosis of other pathological indications. The effectiveness of different current models, including neural networks and deep learning algorithms, in detecting eye disease has been examined in the research work [11]. The process of identifying eye diseases using retinal images is broken down into several smaller processes, including feature extraction, classification, and picture pre-processing. This study provides an overview of deep learning, its algorithms, the operation of convolution neural networks, and its applications to image processing, machine learning, and deep learning techniques that are utilized for retinal image-based eye disease identification [12].

III.Proposed Methodology:

Neural Network:

Artificial Neural Networks (ANN) are mathematical models that copy the neural structure of the mammalian cerebral cortex but on much smaller scale [13]. Neural networks are neurons organized in layers. Layers of the ANN are made up of a number of fully connected 'nodes' which contain a non-linear 'activation function' that can be used to reduce error using back propagation while performing gradient descent. Patterns are recognized by the network via the 'input layer', which is followed by one or more 'hidden layers' where the pattern processing is done via a system of weighted 'connections'. The hidden layers then link to an 'output layer' which realizes patterns across retinal images in this case. Most ANNs contain some form of 'learning rule' which alters the weights of the neurons according to the input patterns that it is presented with. However a traditional neural network fails to analyze patterns in different positions.
A Convolutional Neural Network (CNN) is a type of neural network that can identify structural features in an image. The CNN is able to capture the pattern at any location across the retina, by allowing a filter to slide through the entire image to perform pattern matching [14]. The filter moves across the image using a stride which decides how much the filter must move as it matches the image pattern. CNNs are made up of processing units that have self-learning weights and biases. Each neuron receives some inputs, performs a dot product of these with weights and biases and optionally follows it with an activation function. The whole network uses a single differentiable score function, from the raw image pixels on one end to class scores at the other. The CNN architecture takes advantage of the fact that the inputs are images, which allows to encode certain properties into the architecture. This makes the forward function more efficient to implement and vastly reduces the amount of parameters in the network. In particular, unlike a regular ANN, the layers of a CNN have neurons arranged in 3 dimensions: width, height, depth. A CNN is made up of layers (minimum 5). Every layer transforms 3d input volume of data to a 3d output volume with some differentiable function that may or may not have parameters. The CNN consists of the three components. The first one is the convolutional layer which consists of filters moving by a parameter called stride to identify patterns across the image. The second component is the max pool layer which performs down sampling to ignore unnecessary features and reduce computation. The third component is the fully connected normal dense layer like a traditional neural network required to output the result. Deep CNN is different from a neural network in the sense that a neural network is fed all the pixels of an image to a single layer and then interconnected with another dense layer, this however tends to lead to an over fitting model if the abnormal pattern occurs at different locations/positions in the retina. The edges may be detected by certain neurons and the core by others. Using by a particular stride (stride can be thought of as steps) and

ensure that different neurons obtain different data of localization of patterns (in this case abnormal patterns in the retina). This way the network learns more about the pattern than figuring out where it occurs in the image as opposed to an ordinary neural network.

Median Filers for removing the noise from OCT images

In this scenario, the median filter is used to reduce the noise from the OCT input image and this works like mean filter. This filter initializes the every pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. Instead of simply replacing the pixel value with the mean of neighboring pixel values, it replaces it with the median of those values. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value. (If the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is used.)

The median filter is non-linear. This means that for two images P(a) and Q(a):

$$\mathrm{median}[\mathrm{P(a)}+\mathrm{Q(a)}] \neq \mathrm{median}[\mathrm{P(a)}] + \mathrm{median}\,[\mathrm{Q(a)}] \qquad (1)$$

Ensemble Edge Detection Approach:

The Marr-Hildreth edge detector was a very popular edge operator. Previously canny edge detection is proposed the algorithm to find the accurate edges for OCT input images. It is a gradient based operator which uses the Laplacian to take the second derivative of an image. It works on zero crossing method and uses both Laplacian and Gaussian (LoG) operator to reduce the noise and finds the sharp edges.

The Gaussian function is defined by the formula

$$\mathrm{G(a,b)} = \frac{1}{\sqrt{2\prod \sigma^2}} \exp - \left(\frac{\mathrm{a}^2+\mathrm{b}^2}{2\sigma^2}\right) \qquad (2)$$

Where, σ is the standard deviation. And the LoG operator is calculated from

$$\mathrm{LoG} = \frac{\partial^2}{\partial \mathrm{a}^2}\,\mathrm{G(a,b)} + \frac{\partial^2}{\partial \mathrm{b}^2}\mathrm{G(a,b)} = \frac{\mathrm{a}^2+\mathrm{b}^2-2\sigma^2}{\sigma^4}\exp\left(-\frac{\mathrm{a}^2+\mathrm{b}^2}{2\sigma^2}\right) \qquad (3)$$

DenseNet (Densely Connected Convolutional Networks)

DenseNet (Densely Connected Convolutional Networks) is a deep learning architecture that has gained popularity due to its effectiveness in image classification and object detection tasks. The DenseNet architecture addresses some difficulties associated with training intense neural networks by encouraging feature reuse and addressing the vanishing gradient problem.

DenseNet employs so called "dense connections" which connect all previous layers to each layer in a feedforward fashion. Such dense connections facilitate flow of information and gradients in the network, leading to improved features reuse with a more concise architecture. The total of DenseNet architecture are set into several 'Dense Blocks' and each dense block contains several densely connected layers. The Dense Blocks are connected by transition layers that reduce the spatial dimensions of feature maps as well as number of channels to help controlling model complexity. DenseNet models are equally identified by number of layers and growth rate.

DenseNet-169 to retinal disease detection:

Detection of retinal disease involves examining the retina images for diabetic retinopathy, age-related macular degeneration and glaucoma. The detection of such diseases is challenging due to the subtle and complicated accounting patterns in the retinal images. Deep learning algorithms, like DenseNet-169, have shown encouraging results in automating this detection. denseNet-169 is able to capture intricate patterns and structures in retinal images, because of its deep architecture and dense connections between the layers. The profound layers with high connectivity facilitate the information to propagate through the network, enabling the model learning and modeling global-local features of images. As for such ability of remembering the hierarchical representations, since DenseNet-169 contains multiple layers as deep its structure, it could be beneficial to detect the subtle abnormalities.

When retinal disease is detected, DenseNet-169 model is usually fine-tuned on a labeled retinal image dataset. Learning the patterns of different diseases, the model learns to classify retinal images into normal and abnormal ones. dense-net-169 has large enough capacity to mine the texture, structure, and color differences that are related with retinal diseases. In addition, owing to that it can learn fine points from images and complex patterns as well as the hierarchy relationship of its architecture benefit for precision promotion, it is a powerful structure making it easy to be used in retinal disease detection tasks.

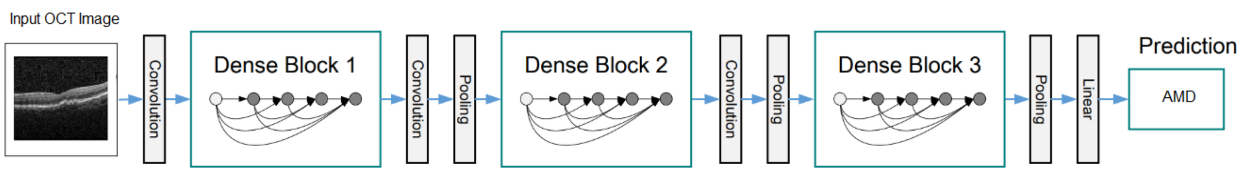


Figure 2 Architecture of Densenet-169

An Amalgamate Learning Model (ALM)

In the last decade, GANs and CNNs have demonstrated great potential for a variety of tasks, such as retinal disease recognition. This powerful synergy leverages the benefits of GANs and CNNs to enhance both accuracy and robustness for disease detection models.

The working of GANs and CNNs can be used together for retinal disease detection:

Data Augmentation Using GANs: One of the main challenges to train CNNs for medical image analysis is the lack of labeled data, especially when it comes to rare diseases. The GAN offers the capabilities to generate artificial retinal images that is indistinguishable from real ones, automatically enhancing the training set. Such generated images can be used to augment the training data, and therefore might help improve generalization of CNN model.

Image Enhancement Using GANs: GANs could additionally be implemented to enhance the resolution of retinal images. Given a dataset of lower-resolution and high-resolution images, a GAN can be trained to generate better versions of the input images. This is useful when dealing with noise corrupted and low quality retinal images, since the CNN architecture model performance can be increased.

Disease Detection Using CNNs: CNNs outperform manual feature extraction in the medical image analysis. A CNN can be trained to recognize patterns and deviations relevant to particular diseases such as diabetic retinopathy, macular degeneration or glaucoma. This task is well-suited to CNNs, as they have an inherent ability to discover hierarchical representations.

Adversarial Training: In GAN and CNN training, typically the CNN component is trained in a supervised fashion to force it classify retinal images correctly. The GAN part on the other hand, it is trying to generate believable enough images that can trick CNN into making false predictions. This competitive framework compels CNN to enhance its generalization, and ensure better discrimination between real and generated samples.

Transfer Learning: In this context, transfer learning can also be used. Pre-trained CNNs, such as those trained on large-scale natural image datasets, can be fine-tuned for retinal disease detection. This method enables the CNN to apply knowledge learned from generic image features to the specific retinal domain.

IV.Experimental Results

The confusion matrix is used to analyze the classification algorithm performance. Measuring the performance with a confusion matrix gives a better for find the accurate errors. It is also used to solve several classification issues. This can be applied for classification of binary issues and also for multiclass classification issues. The count values are based on various attributes such as

TP: In this the actual value of true (disease present) and predicted value is true (disease present).

TN: The actual value is true (disease present) and predicted value is false (No disease).

FP: The actual value is false (No disease) and predicted value is true (disease present).

FN: The actual value is false (No disease) and predicted value is false (No disease).

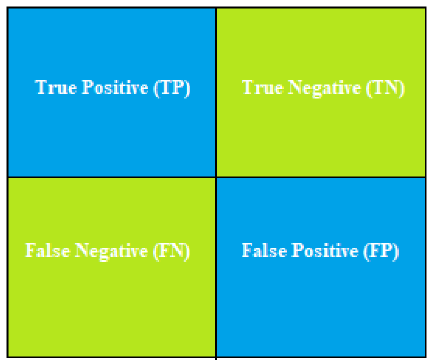


Figure 3: Confusion Matrix

Sensitivity (S_n): Sensitivity, recall, or the TP rate (TPR) is the fraction of positive values out of the total actual positive instances (i.e., the proportion of actual positive cases that are correctly identified):

$$S_n = \frac{TP}{TP + FN}$$

Specificity (S_p): Specificity gives the fraction of negative values out of the total actual negative instances. In other words, it is the proportion of actual negative cases that are correctly identified. The FP rate is given by (1 – specificity):

$$S_p = \frac{TN}{TN + FP}$$

Precision (P): Precision or the positive predictive value, is the fraction of positive values out of the total predicted positive instances. In other words, precision is the proportion of positive values that were correctly identified:

$$P = \frac{TP}{TP + FP}$$

Accuracy (Acc): Accuracy shows the total number of prediction that is correct. Actual and predicted values are correct. It is represented with below formula.

$$Acc = \frac{TP + TN}{TP + FP + TN + FN}$$

F1-Score (F1S): The F1-score combines the precision and recall of a classifier into a single metric by taking their harmonic mean.

$$F1S = 2 * \frac{P * S_n}{P + S_n}$$

Table 1: Confusion Matrix count values

Count Values	CNN	RESNET	ALM
TP	2378	2675	2879
TN	2213	2087	1856
FP	245	125	165
FN	114	64	98

Table 2 Shows the Performance of Existing and Proposed Algorithms for detection of CNV

	CNN	ResNet	ALM
Sensitivity (SE)	77.22	86.76	98.98
Specificity (SP)	81.54	85.76	98.56
Precision (PE)	80.65	85.43	98.34
Accuracy (ACC)	81.66	87.52	97.76
F1-Score	82.12	87.53	97.58

Table 3 Shows the Performance of Existing and Proposed Algorithms for detection of DME

	CNN	ResNet	ALM
Sensitivity (SE)	78.32	86.78	98.96
Specificity (SP)	81.34	87.34	98.12
Precision (PE)	82.34	88.43	98.44
Accuracy (ACC)	82.45	87.12	98.12
F1-Score	83.12	88.43	98.56

Table 4 Shows the Performance of Existing and Proposed Algorithms for detection of Drusen

	CNN	ResNet	ALM
Sensitivity (SE)	79.32	87.64	98.89
Specificity (SP)	82.34	87.87	97.12
Precision (PE)	82.34	88.53	97.34

Accuracy (ACC)	83.45	86.12	98.12
F1-Score	84.12	87.43	98.56

Table 5 Shows the Performance of Existing and Proposed Algorithms for detection of AMD

	CNN	ResNet	ALM
Sensitivity (SE)	78.12	85.34	98.56
Specificity (SP)	80.34	86.34	97.34
Precision (PE)	81.34	87.43	99.34
Accuracy (ACC)	82.45	86.12	99.45
F1-Score	82.12	86.43	98.56

Table 6 Shows the Performance of Existing and Proposed Algorithms for detection of Normal

	CNN	ResNet	ALM
Sensitivity (SE)	71.12	84.34	99.56
Specificity (SP)	81.34	87.12	98.34
Precision (PE)	82.34	87.98	99.67
Accuracy (ACC)	83.45	85.89	99.78
F1-Score	84.12	87.09	99.89

Table 7 Shows the Performance of Existing and Proposed Algorithms for detection of Overall performances

	CNN	ResNet	ALM
Sensitivity (SE)	73.32	83.34	99.23
Specificity (SP)	81.64	88.12	98.34
Precision (PE)	82.74	88.98	99.45
Accuracy (ACC)	84.85	89.89	99.76
F1-Score	84.78	88.19	99.67

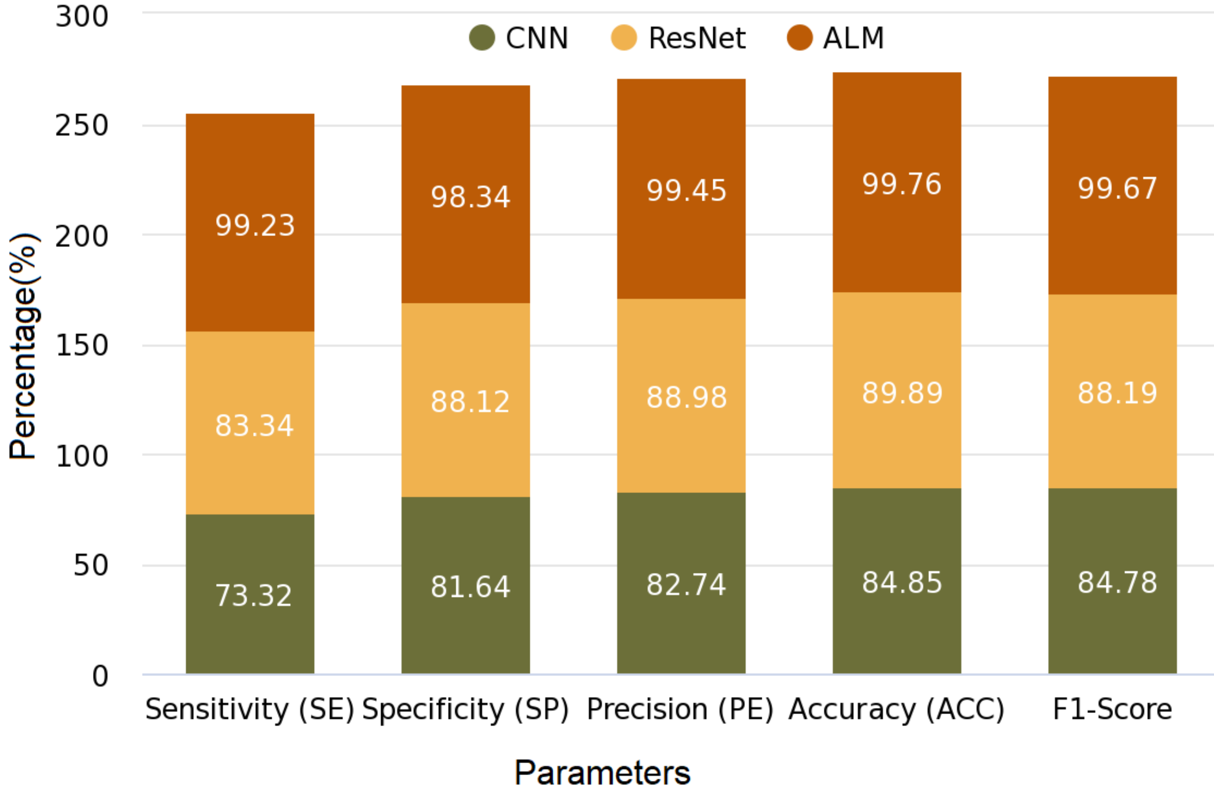


Figure 2; The Overall Performances of Detecting Retinal Diseases

Conclusion

The Amalgamated Learning Model (ALM) presented in this study, which combines GAN with CNN for retinal disease classification, has shown promising results in this paper. VERDICTGAN and CNN have some advantages of the combination. GI = GANs are powerful generative models, which can effectively learn from an unlabelled dataset and generate realistic-looking synthetic samples. GANs can be used to increase the size of the training set and therefore make ALM more general and applicable in tasks with small labelled data. In addition, the CNN component of the ALM brings its advantage in feature extraction from retinal images. This enables the model to learn intricate patterns and elusive details that are essential for an accurate disease classification. We found that the ALM outperformed alternative methods (CNN models and traditional classification) when applied alone. It showed patient and control discrimination in all retinal disease categories as DRM, ADM, DRUSEN and routine (Controls). Moreover, the ALM can easily synthesize images, so that simultaneously achieving data augmentation for dealing with the limited number of medical images. It could be useful for retinal diseases, because of the difficulty in collecting good datasets with such variety and size. Nevertheless, contrasting the encouraging completion rate of 91–94%, the ALM has limitations [2]. Training a GAN might be hard or costly, as GANS take long to train and are computationally intensive. The performed parameter optimization and hyperparameter tuning are required to balance between the GAN and CNN parts. Eventually, ALM (GAN plus CNN) in the classification of retinal diseases has greatly promoted the progress of medical image analysis. Its capacity of data augmentation with GANs and features extraction by CNNs makes it a very useful tool for helping clinicians to diagnose retinal diseases properly and in a timely manner. With advancing refinement and optimization, the potential is there that ALM might revolutionize the field to better serve patients in ophthalmology.

References

[1] Mateen, M.; Wen, J.; Hassan, M.; Nasrullah, N.; Sun, S.; Hayat, S. Automatic Detection of Diabetic Retinopathy: A Review on Datasets, Methods and Evaluation Metrics. IEEE Access 2020, 8, 48784–48811.

[2] Rekhi, R.S.; Issac, A.; Dutta, M.K. Automated detection and grading of diabetic macular edema from digital colour fundus images. In Proceedings of the 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON), Mathura, India, 26–28 October 2017; pp. 482–486.

[3] Al-Bander, B.; Al-Nuaimy, W.; Al-Taee, M.A.; Zheng, Y. Automated glaucoma diagnosis using deep learning approach. In Proceedings of the 2017 14th International Multi-Conference on Systems, Signals & Devices (SSD), Marrakech, Morocco, 28–31 March 2017; pp. 207–210.

[4] Zago, G.; Andreão, R.V.; Dorizzi, B.; Salles, E.O.T. Diabetic retinopathy detection using red lesion localization and convolutional neural networks. Comput. Biol. Med. 2020, 116, 103537.

[5] Kunwar, A.; Magotra, S.; Sarathi, M.P. Detection of high-risk macular edema using texture features and classification using SVM classifier. In Proceedings of the 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, India, 10–13 August 2015; pp. 2285–2289.

[6] Quigley, H.; Broman, A.T. The number of people with glaucoma worldwide in 2010 and 2020. Br. J. Ophthalmol. 2006, 90, 262–267.

[7] Bourne, R.R.; Flaxman, S.R.; Braithwaite, T.; Cicinelli, M.V.; Das, A.; Jonas, J.B.; Keeffe, J.; Kempen, J.H.; Leasher, J.; Limburg, H.; et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: A systematic review and meta-analysis. Lancet Glob. Health 2017, 5, e888–e897.

[8] Al-Bander, B.; Williams, B.M.; Al-Nuaimy, W.; Al-Taee, M.A.; Pratt, H.; Zheng, Y. Dense Fully Convolutional Segmentation of the Optic Disc and Cup in Colour Fundus for Glaucoma Diagnosis. Symmetry 2018, 10, 87.

[9] Chen, X.; Xu, Y.; Wong, D.W.K.; Wong, T.Y.; Liu, J. Glaucoma detection based on deep convolutional neural network. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; Volume 2015, pp. 715–718.

[10] Mary, M.C.V.S.; Rajsingh, E.B.; Naik, G.R. Retinal Fundus Image Analysis for Diagnosis of Glaucoma: A Comprehensive Survey. IEEE Access 2016, 4, 4327–4354.

Sinthanayothin, C. Image Analysis for Automatic Diagnosis of Diabetic Retinopathy; University of London: London, UK, 1999.

[11] Krishna, B.S.; Gnanasekaran, T. Unsupervised Automated Retinal Vessel Extraction Framework Using Enhanced Filtering and Hessian Based Method with Hysteresis Thresholding. J. Med. Imaging Health Inf. 2019, 9, 1000–1010.

[12] Khansari, M.M.; O'Neill, W.D.; Penn, R.D.; Blair, N.P.; Shahidi, M. Detection of Subclinical Diabetic Retinopathy by Fine Structure Analysis of Retinal Images. J. Ophthalmol. 2019, 2019, 5171965.

[13] Qummar, S.; Khan, F.G.; Shah, S.; Khan, A.; Shamshirband, S.; Rehman, Z.U.; Khan, I.A.; Jadoon, W. A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection. IEEE Access 2019, 7, 150530–150539.

[14] Pan, X.; Jin, K.; Cao, J.; Liu, Z.; Wu, J.; You, K.; Lu, Y.; Xu, Y.; Su, Z.; Jiang, J.; et al. Multi-label classification of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography based on deep learning. Graefe's Arch. Clin. Exp. Ophthalmol. 2020, 258, 779–785.