ISSN: 0369-8963

Genetic Algorithm Based Deep Learning Model Selection for Diabetic Retinopathy Classification

R S RAJKUMAR[1,*], A. GRACE SELVARANI[2]

[1,2]Sri Ramakrishan Engineering College, Coimbatore, Tamil Nadu, 641022, India

Abstract:

Background: Diabetic Retinopathy (DR) is a vision-threatening disease as a result of long-term diabetes. DR cannot be cured however it can be controlled if it is detected at an early stage. Hence the early diagnosis of DR is essential to prevent vision loss for diabetic patients. Deep learning models have their footprint in the early diagnosis of disease from medical images. Thus the best deep learning model is selected from various existing deep learning models for DR diagnosis.

Method: Genetic Algorithm (GA) is a nature-inspired algorithm based on Darwin's theory of evolution. Here GA is proposed to select the best deep learning model from various existing pre-trained deep learning models. Pre-trained models are the initial population of GA. During the GA evolution, the best model is identified or new populations are generated with the help of GA operations.

Result: Experiment results show that the GA-based model identification performs better for early diagnosis of DR. It is observed that ResNet50 outperforms various existing pre-trained models VGG16, Inception Net, and GoogleNet.

Introduction

In this technical era, everyone is running behind technology for their career growth. Unfortunately, they forget their health. This lifestyle change has brought a lot of health complications. One such life-threatening disease is diabetes. If the glucose level is in uncontrolled level for long-term, there is a more likely chance for diabetic patients to develop DR. Once DR developed is no medication to completely cure however it can be controlled if diagnosed at an early stage. Diagnosing DR at an early stage is very difficult for physicians and it is very time-consuming too. This requires regular eye screening which is costlier for diabetic patients.

Deep Learning is an artificial neural network with deep layers learning several parameters. Advanced technology like parallel computing and GPU has made deep learning popular as the computations are done in parallel. Deep convolution neural networks (DCNN) are widely used to classify and recognize images [1-3]. There exists numerous pre-trained DCNN for the classification task. However, they have been developed for various purposes and datasets [4]. Hence it is difficult to select the best model for the DR dataset.

 https://doi.org/10.37896/pd91.3/91316

ISSN: 0369-8963

Training the DCNN model from scratch for a particular dataset requires learning millions of parameters. This requires high computational power [5]. Hence transfer learning can be employed to reduce the computations required. AlexNet, GoogleNet, VGG16, Inception Net, ResNet, and MobileNet are some of the popular pre-trained models used for medical image classification.

A Genetic Algorithm (GA) is a nature-inspired algorithm based on Darwin's evolution theory. GA is extensively applied for searching and optimizing problems. The genetic operations of GA are selection, crossover, and Mutation [6]. Integration of GA with DCNN is now becoming popular as the DCNN hyper-parameters are learned automatically [7].

Differentiating from present work, we use GA to select the best pre-trained model for early diagnosis of DR. GA autonomously selects the best model from their population or produces new models as their offsprings. The selected model will have the best parameters representing the pre-trained architecture.

The rest of the paper is structured as follows. Section II describes the literature survey related to GA and DCNN. Section III describes the proposed work. Section IV describes the experimental setup and results.

Related Work

Various literature survey related to the detection of DR has been given in one of our previous work [8]. DR diagnosis with the help of ResNet with Fuzzy Rough C-Means Clustering (FRCM) has been proposed in [9]. The vagueness and uncertainty in the data are removed using a rough set and fuzzy set. The research for automatic modeling of Neural Network (NN) Structure has been carried out for many years [10] [11]. The explicit focus was given to GA for fine tuning the NN structure. In [12] NN structure and its hyper-parameters are learned together. Since GA has computational overload, they reduced it by hand-picking selected parameters. In [13] the hyper-parameters of NN were fine-tuned by GA. The test error was used as the evaluation function by them for fitness calculation. In [14] the weights of the deep NN were fine-tuned by GA. It was used for human gesture identification and GA helped them in getting rid of the NN getting stuck in the local optima. Binary Sequence of fixed piece encoding was proposed in [15]. This genetic convolutional neural network automatically structured the NN with the help of GA. G. Genetic algorithm technique to automatic optic disc detection and removal of false exudates, they used baseline method to evaluate the improvement in classification in [16]. A hybrid approach called Genetic algorithm and vertex chain code for blood vessel detection was proposed in [17]. Dual classification and Genetic Algorithm for the Automated detection of Proliferative Diabetic Retinopathy were used in [18].

Though various works had been carried out for the optimization of neural network using genetic algorithm, automatic selection of neural network for a particular dataset and purpose is hardly carried out. Hence the automatic selection of neural network without manual intervention has been carried out.

Proposed Work

An automatic selection of CNN network structure from the pre-trained model with the help of GA is proposed. The list of pre-trained CNN architectures forms the initial population for GA. From the population, specific combinations of the models are selected and they form the

 https://doi.org/10.37896/pd91.3/91316

ISSN: 0369-8963

individual of the GA model. The selected individuals are trained further for deep feature extraction. The major drawback of employing GA is its huge computational cost. To reduce it LSTM classifier is employed to classify the extracted deep features. To fine-tune the algorithm the accuracy of the model is used for its evaluation as the fitness function. The accuracy of the testing data is to rank the individual models in the existing population. With several repetitions of genetic operations - selection, crossover, and mutation the final best fit model is evolved. This best fit model will have deep representations of features and are used for the final DR severity diagnosis. The overall architecture of the proposed work is shown in figure 1.

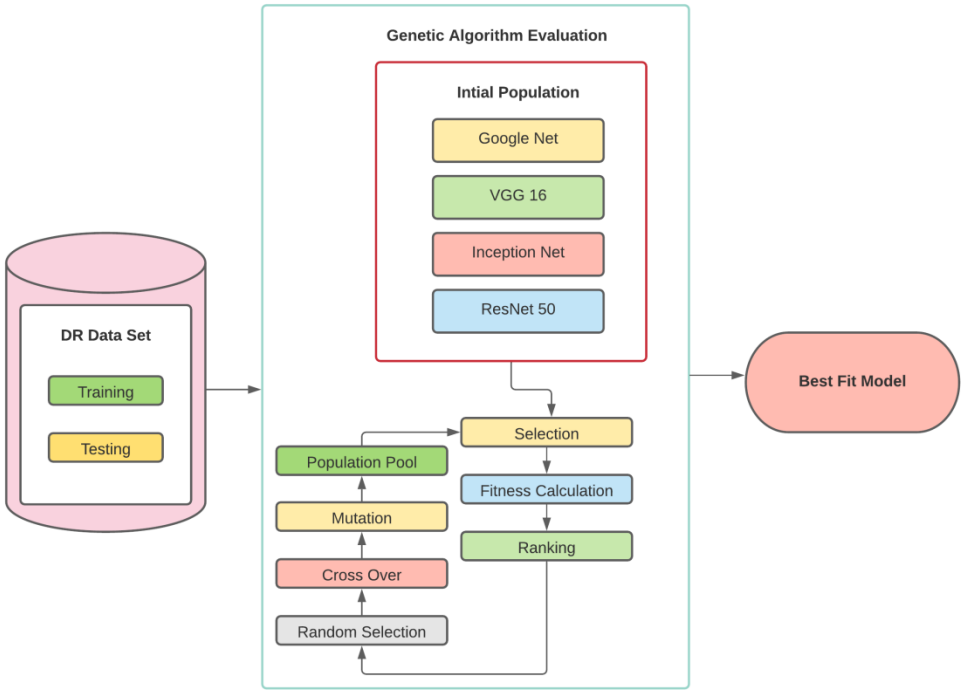

Figure 1. Model Selection using GA evolution

The detailed steps of GA evolution are given in Algorithm 1.

Algorithm 1: **Genetic based Neural Network Selection for DR Diagnosis**
Input: Labeled fundus image dataset
Output: Optimized Neural Network for DR diagnosis
Initialization:
Population, *P* = {InceptionV3, ResNet50, VGG16, GoogleNet}
With_hold = 0.3
Mutate = 0.3
Select_Random = 0.1
Steps:

i. For every, i ∈ *P* do
    a. Calculate fitness function, F(i)
    b. Grade, *G*[i] = F(i)
ii. Sort *P* with respect to non-increasing order of *G*[i]

https://doi.org/10.37896/pd91.3/91316

ISSN: 0369-8963

iii. For x∈(0, *with_hold* * *G*.size - 1) do
    a. *Parent*.append(*G*[*x*])
iv. For every *x* in *P* calculate a random number and do
    a. If random(*x*) > *Select_Random* then
        i. *Parent*.append(*G*[*x*])
v. For every *x* in *Parent* calculate a random number and do
    a. If random(*x*) > *Mutate* then
        i. *Parent*.append(*G*[*x*])
vi. *size* =*P*.size – *Parent*.size
vii. If *offsprings*.size < *size* then
    a. Select two parents, *P1* and *P2* from *Parent* at random and do
    b. If *P1* ≠ *P2* then
        i. *offspring* = (*P1*[0, *P1*.size/2-1] + *P2* [*P1*.size/2, *P1*.size-1)
        ii. *offsprings*.append(*offspring*)
        iii. goto step vii
viii. *Parent*.append(*offspring*)
ix. Return *parent*

The individual chromosome selection for GA is like a logic gate with four inputs and one output. The four inputs correspond to the individual model in the GA population. The output will be the new model which corresponds to the new offspring concerning the individual input selection. With the combination of four inputs 16 (2^4) individual selections are possible in the range of 0000 to 1111. However the input selection 0000 will not produce any offspring as it signifies no input model is selected. So the input 0000 is discarded from input selection. Thus there will be 15 ($2^4 - 1$) individuals with the combination of models in the population. The naïve search methodology for GA will incur huge computation costs. To overcome it, LSTM is used for classification. We aim to improve the accuracy of the model. Hence accuracy is taken for the fitness calculation and ranking of each offspring produced. Every iteration in algorithm 1, consists of initialization, fitness calculation, individual ranking, random selection, crossover and mutation, and best individual model selection.

The model list is generated and from the list, a fixed-length binary string like '1011' is generated for initialization of GA. The binary bit '1' corresponds to the particular model selection and '0' corresponds to the non-selection of the model for GA input. Initially, with the pre-trained models, the population is randomly generated. GA operations are performed on the selected input and new populations are successfully generated in each iteration of algorithm 1.

Fitness calculation will strictly allow only the fittest model to be included in the population. As described fitness is the objective or utility value to be maximized. This ensures the model with a higher fitness value to be included in the next generation of the population with greater probability. This is similar to Darwin's species evolution theory. As depicted earlier the deep features from the offspring (new model) are classified using LSTM. After classification, the accuracy of the model is evaluated as

$$f(x) = \frac{\mathrm{TP+TN}}{(\mathrm{TP+FP+FN+TN})} \quad (1)$$

 https://doi.org/10.37896/pd91.3/91316

ISSN: 0369-8963

Where TP represents the true positive, TN represents the true negative, FP represents the false positive and FN represents the false positive of the classification.

After the fitness calculation, each model is then ranked in the non-increasing order of their fitness value. The withheld rate of population is set as 0.3 in algorithm 1. This will enable only 3 individuals will survive out of 10 produced models. The other 7 individual models are considered to be non-fittest and are supposed to vanish or die.

Though only fittest models are retained for the next population, to introduce genetic variance, few models with low fitness scores are also included in the population. This introduction of genetic variance will help algorithm 1 from getting stuck at local maxima. This helps the optimal algorithm to attain the global maxima. To ensure it, low fitness models are randomly selected with a threshold of 0.1. For every new offspring generated, a random number [0,1] is assigned. The new model will be kept in the momentary pool if its random number is greater than the threshold. From the momentary pool, few models are selected at random for including genetic variance.

Genetic variance is further improved by randomly changing a small portion of every individual. This random change in an individual is coined as mutation. With higher mutation, algorithm 1 can be converged faster. The mutation probability is set to 0.3.

If any more individuals can be added further after population retainment, the crossover of individuals is performed to fill the population. For crossover to happen, two individual models are selected at random as parents. The parent models are allowed to crossover with the crossover point being mid of the parent length. There are no restrictions on the individual for their participation in the crossover. However only restriction is that both the randomly selected individual selected as parents should not be the same. If both the parents are supposed to be the same, then crossover will not introduce new models into the population.

With 30% individual retainment and 30% mutation, the utmost 3 generations are required for the evolution to get the best model as a solution. Two or more individual models may compete for their survival during the last generation. If so, the clash with them is resolved by simply performing an 'AND' logic operation on them. For instance if '1011' and '1101' compete each other with same fitness score, '1001' (1011 && 1101 = 1001) . Individual 1001 will incorporate both the features of 1011 and 1101. This will also reduce the computational overhead by reducing the redundant features of individuals.

Experimental Setup and Result Analysis:

A. Dataset

There exist several real-time datasets for DR diagnosis. For reducing the overfit of a deep learning model, it is necessary to feed them with a huge amount of data. This will help in producing a more generalized model. Hence among the various publically available datasets, Diabetic Retinopathy Detection [19] was used for algorithm 1. It is publically accessible from the Kaggle forum. It contains more than 35k images. The images were taken by EyePacs foundation under various conditions. The images in the dataset vary in the aspect of their

 https://doi.org/10.37896/pd91.3/91316

ISSN: 0369-8963

resolution, shape, contrast, and brightness. Every image in the dataset was labeled from 0-1 according to their DR development stage by ophthalmologists. This dataset is highly imbalanced.

B. Experimental Setup

The proposed system was implemented using python with TensorFlow which was built over Keras. The algorithm was run on a server equipped with Tesla V100 card. The 35, 126 retinal images in the dataset were divided into 60:20:20 ratio for train data, validation data, and test data. The retinal images are pre-processed for better results. Pre-processing of images are required as

i. the images are to be equally sized by re-sizing and
ii. the images in each DR development class have to be balanced by augmentation.

C. Result Analysis

The system performance is analyzed by accuracy and weighted kappa score. Since the dataset is of with high imbalance f1-score is also used to measure the performance of the system. The GA evolution for best deep learning model selection for DR diagnosis has a good ability to identify the best model. The confusion matrix and the inference from it for the model selected by our algorithm are shown in figure 2 and figure 3.

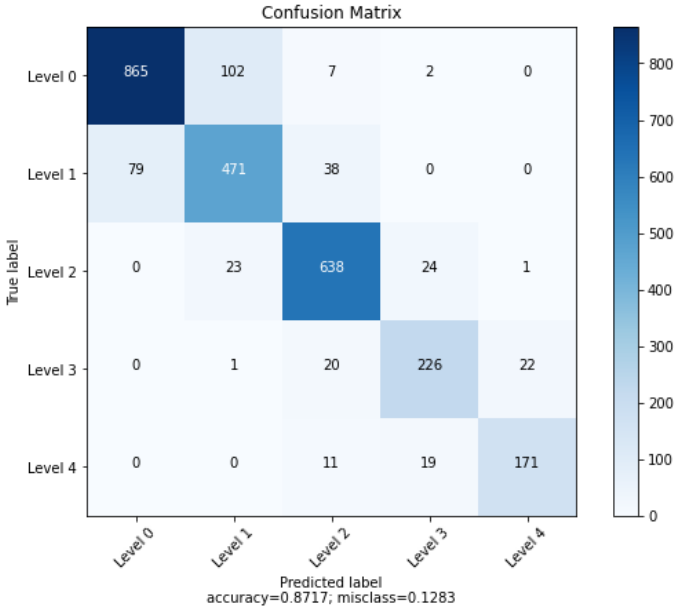

Figure 2: Confusion Matrix for Model Selection by GA Evolution

 https://doi.org/10.37896/pd91.3/91316

ISSN: 0369-8963

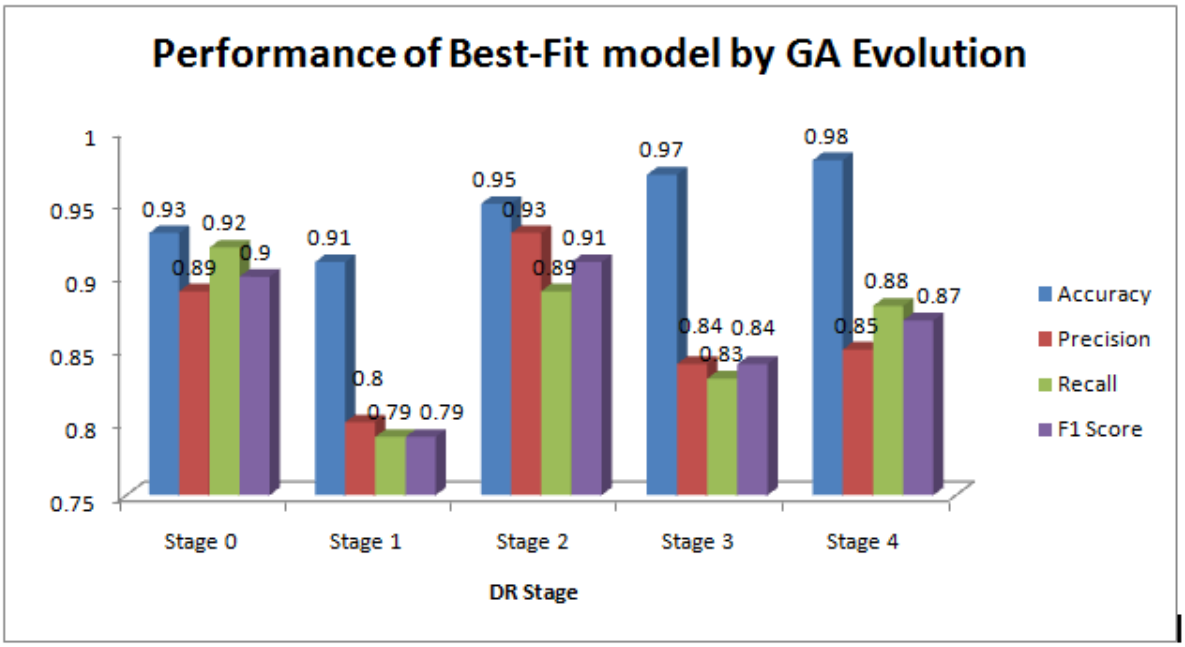

Figure 3: Performance Measures for Model Selection by GA Evolution

The selected best fit model by algorithm 1 is compared with the base pre-trained models InceptionV3, ResNet50, VGG16, GoogleNet. For the chosen dataset it is observed that the best fit model selected by the proposed algorithm is very much similar to the ResNet50 Model. Hence it is noted that combination of other deep learning models with ResNet50 does not show greater improvement in their DR diagnosis results. The performance of our GA model and baseline models are compared and revealed in Table 1.

Table 1: Evaluation Results of DR Detection Dataset on Various Models

Deep Learning Model	Overall Accuracy	Kappa Score	Average F1 Score
VGG16	0.8106	0.72	0.803
GoogleNet	0.8451	0.772	0.83
Inception V3	0.8587	0.786	0.85
ResNet50	0.8702	0.794	0.86
GA model with LSTM (Proposed)	0.8717	0.82	0.862

Conclusion

There exist many deep learning models sketched by hand. They are designed for specific tasks. There exists no method to automatically select the deep learning model for DR diagnosis. This confronts is addressed by genetic algorithm with its efficient search for deep learning model with optimized structure and hyper-parameters. The proposed genetic algorithm evolution for deep learning model optimization selects the best-fit model from the model population pool. The experiment results and its analysis shows that the proposed algorithm is more efficient than the base models. The Diabetic Retinopathy Detection dataset from the Kaggle contains noise and

 https://doi.org/10.37896/pd91.3/91316

ISSN: 0369-8963

uncertainty. In future model can be fine-tuned to address the noise and uncertainty to improve the classification results.

References:

[1]A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

[2] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[4] S. Pouyanfar and S.-C. Chen, "Automatic video event detection for imbalance data using enhanced ensemble deep learning," International Journal of Semantic Computing, vol. 11, no. 01, pp. 85–109, 2017.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[6] C. M. Anderson-Cook, Practical genetic algorithms. Taylor & Francis, 2005

[7] D. Hossain, G. Capi, and M. Jindai, "Optimizing deep learning parameters using genetic algorithm for object recognition and robot grasping," Journal of Electronic Science and Technology, vol. 16, no. 1, pp. 11–15, 2018.

[8] R.S. Rajkumar, A.G. Selvarani and S. Ranjithkumar, "Comprehensive study on diabetic retinopathy" in *Soft Computing for Problem Solving. Advances in Intelligent Systems and Computing*, Springer, Singapore , vol. 1057, pp. 155 - 164, 2020.

[9] R.S. Rajkumar and A. Grace Selvarani, "Diabetic Retinopathy Diagnosis Using ResNet with Fuzzy Rough C-Means Clustering" in Computer Systems Science and Engineering, vol. 42, no. 2, PP 509-521., 2022, DOI: 10.32604/csse.2022.021909.

[10] J.-T. Tsai, J.-H. Chou, and T.-K. Liu, "Tuning the structure and parameters of a neural network by using hybrid taguchigenetic algorithm," IEEE Transactions on Neural Networks, vol. 17, no. 1, pp. 69–80, 2006.

[11] S. R. Young, D. C. Rose, T. P. Karnowski, S.-H. Lim, and R. M. Patton, "Optimizing deep learning hyper-parameters through an evolutionary algorithm," in Workshop on Machine Learning in High-Performance Computing Environments. ACM, 2015, pp. 4:1–4:5.

[12] F. H.-F. Leung, H.-K. Lam, S.-H. Ling, and P. K.-S. Tam, "Tuning of the structure and parameters of a neural network using an improved genetic algorithm," IEEE Transactions on Neural networks, vol. 14, no. 1, pp. 79–88, 2003.

 https://doi.org/10.37896/pd91.3/91316

ISSN: 0369-8963

[13] S. R. Young, D. C. Rose, T. P. Karnowski, S.-H. Lim, and R. M. Patton, “Optimizing deep learning hyper-parameters through an evolutionary algorithm,” in Workshop on Machine Learning in High-Performance Computing Environments, 2015, p. 4.
[14] E. P. Ijjina and K. M. Chalavadi, “Human action recognition using genetic algorithms and convolutional neural networks,” Pattern Recognition, vol. 59, pp. 199–212, 2016.
[15] L. Xie and A. Yuille, “Genetic CNN,” in IEEE International Conference on Computer Vision, 2017, pp. 1388–1397.
[16] G.Ferdic ,Mashak Ponnaiah , Capt.Dr.S.Santhosh Baboo: ”Automatic optic disc detection and remival of false exudates for improving retinopathy classification accuracy”,, International Journal of Scientific and Research Publication , Volume 3, Issue 3, ISSN 2250-3153International Journal of Scientific and Research Publications, Volume 3, Issue 3, March 2013 1 ISSN 2250-3153
[17]Romany Fouad Mansour: “Using genetic algorithm for identification of diabetic retinal exudates in digital color images” , Journal of Intelligent Learning Systems and Applications, 2012, 4, 188-198 http://dx.doi.org/10.4236/jilsa.2012.43019 Published Online August 2012
[18]R.A Welikala, M.M. Fraza , J. Dehmeshkia , A. Hoppea , V. Tahb , S. Mannc , T.H. Williamsonc , S.A. Barman: “Genetic algorithm based feature selection combined with dual classification for the automated detection for proliferative dabetic retinopathy” , Digital Imaging Research Centre, Faculty of Science, Engineering and Computing, Kingston University, London, United Kingdom.2015, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives -4.0 International http://creativecommons.org/licenses/by-nc-nd/4.0/
[19] Diabetic Retinopathy Detection. Accessed: Sep. 1, 2018. [Online]. Available: https://www.kaggle.com/c/diabetic-retinopathy-detection/data

 https://doi.org/10.37896/pd91.3/91316