**Using artificial intelligence for analyzing the retinal images (OCT) in people with diabetes**

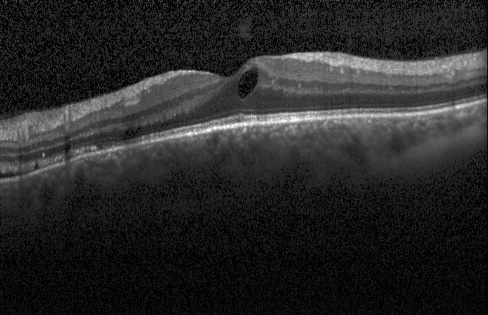
**Abstract**

Diabetic macular edema (DME) is a common disease of diabetic retinopathy (DR). Due to the infection of DME disease, many patients’ vision is lost. To cure DME eye disease, early detection and treatment are very important and vital steps. Medical imaging developed rapidly to play a central role in medicine today by supporting diagnosis and treatment of a disease. Furthermore, plays an important role in initiatives to improve public health for all population. To automatically diagnosis DEM disease, strategies involving Artificial Intelligence (AI) could provide a solution. Among the most promising clinical applications of AI is diagnostic imaging, and mounting attention is being directed at establishing and fine-tuning its performance to facilitate detection and quantification of a wide array of clinical conditions. The main goal of this project is develop an Artificial Intelligence solution that can help to classify OCT. To achieve our goal, Deep-Learning (DL) model will be used to predict the risk of patients with diabetes developing diabetic retinopathy. The outcomes would improve patient access to treatment and ease pressures on time and resources in ophthalmology clinics.

**Design**

This work shows how machine learning can be effectively adopted in the health field to derive models that use patient data to predict an outcome of interest. Artificial Intelligence may be applied to the construction of models for the prediction of patients at high risk of DME, which – once evaluated and tested– may be embedded within health care systems. Patients having high risk of DME can be predicted to set appropriate proactive interventions that reduce the negative impact and provides insightful implications for decision-making by management. The steps of the modeling process described in the following subsections with details of each step.

**Data**

The dataset downloaded from the Kaggle <https://data.mendeley.com/datasets/rscbjbr9sj/2>. A total of (3,500) OCT images included in this dataset. The dataset is organized into 2 folders (train, test) and contains subfolders for each image category (NORMAL and DME). There are 84,495 X-Ray images (JPEG) and 2 categories (NORMAL and DME). Images are labeled as (disease)-(randomized patient ID)-(image number by this patient). Optical coherence tomography (OCT) images (Spectralis OCT, Heidelberg Engineering, Germany) were selected from retrospective cohorts of adult patients from the Shiley Eye Institute of the University of California San Diego, the California Retinal Research Foundation, Medical Center Ophthalmology Associates, the Shanghai First People’s Hospital, and Beijing Tongren Eye Center between July 1, 2013 and March 1, 2017. The ratio of splitting the original dataset (3,500 images) is 50/20/30. Meaning training data made up of 50% training (1,960 images), then devote 20% to the validation set (490 images) to provide an unbiased evaluation of a model fit on the training dataset while tuning model hyperparameters. Test dataset made up of remaining 30% (1,050 images) to evaluate the accuracy of the model on the data it has never seen.

**Feature Engineering**

Resize image to establish a base size for all images fed into our machine learning algorithms. Classic augmentation techniques like flips and rotations will applied to each image in the training set without manually processing each image. Finally normalize all input images.

**Models**

Classification is the most widely used technique for machine learning, especially in prediction models. The machine learning algorithm will be Convolutional Neural Networks (CNN). It used for solving any image data challenge. CNN is a powerful neural network that uses filters to extract features from images. It also does so in such a way that position information of pixels is retained. All popular frameworks support Convolutional Neural Networks like Tensorflow-Keras.

Convolutional Neural Network (CNN) is a deep learning algorithm for images classification. CNN takes an input image, process it and classify it under certain categories (Eg., Normal, Abnormal). Computers sees an input image as array of pixels h x w x d ( h = Height, w = Width, d = Dimension), and it depends on the image resolution. Technically, to train and test CNN model, each input image will pass it through a series of convolution layers with filters (Kernals), Pooling, fully connected layers and apply activation functions to classify an object with probabilistic values between 0 and 1. To demonstrate the impact of convolutional, five models with different Convolutional layers built then the best one selected based on evaluation metrics. Convolutional layers can perfectly handle this type of data Conv2D for images. Dense layers require data to be one-dimensional at all times. For this, after finishing the previous step, pooled feature map applied then flatten our pooled feature map into one-dimensional. An epoch is how many times the model trains on our whole data set. Batch can be explained as taking in small amounts, train and take some more. Each epoch must finish all batch before moving to the next epoch. Early Stopping used to reduce the computer calculation, program running time and avoid overfitting. the number of training epochs treated as a hyperparameter and train the model multiple times with different values, then select the number of epochs that result in the best performance on the train. The model at the time that training is stopped is then used and is known to have good generalization performance.

**Model Evaluation and Selection**

The dataset will divide into three sets: the training set for the development of the model, validation set to fine-tune your model until satisfied with its performance and the test set for evaluate the model. Matrices that used to select the best model are:

* Confusion Matrix
  + True Positive =True Positive/(True Positive+False Negative)
  + False Positive (FPR)=False Positive/(False Positive+True Negative)
  + Precision=True Positive/(True Positive+False Positive)
  + Recall=True Positive/(True Positive+False Negative)
  + F-measure: F\_score =2\*(precision\*recall)/(precision+recall)
* Area Under the ROC curve (AUROC)
* Loss to evaluate and diagnose model optimization.

**Tools**

* The Jupyter Notebook: it is an open-source web application include data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.
* Python programming language: the most popular language for machine learning and artificial intelligence.
* Libraries: A library is a collection of pre-combined codes that can be used iteratively to reduce the time required to code.
  + Tensorflow-Keras for developing deep learning models.
  + Numpy and pandas for data manipulation
  + Matplotlib and seaborn for visualization

**Communication:**

Five CNN models developed with different convolutional layers to classify OCT images to normal or DME. Table1 shows the performance of the CNN models on an unseen testing dataset. The best performing model was the CNN model with5-Convolutional Layer. It achieved better among all metrics to the CNN with fewer layers. Figure1 Shows the summary of the best CNN model builds by using Tensorflow implementing CNN with 5-Convolutional Layer. Model Consists of input layers, two hidden layers, and one output layer with sigmoid as activation function.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table1: The Performance of the Models on the Test Dataset | | | | | |
| CNN Model/Metrics | Accuracy | precision | recall | f1-score | AUC |
| 5-Convolutional Layer | 0.82 | 0.82 | 0.82 | 0.81 | 0.82 |
| 4-Convolutional Layer | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 |
| 3-Convolutional Layer | 0.74 | 0.75 | 0.74 | 0.74 | 0.75 |
| 2-Convolutional Layer | 0.74 | 0.74 | 0.74 | 0.74 | 0.74 |
| 1-Convolutional Layer | 0.63 | 0.70 | 0.63 | 0.60 | 0.63 |

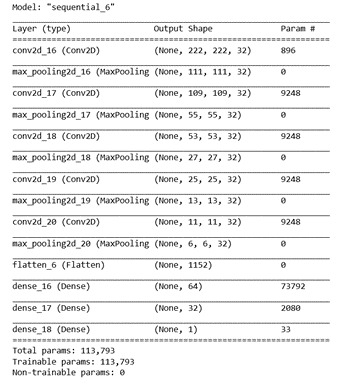


Figure1: The summary of Convolutional Neural Network (CNN) Model

The below two plots, present the accuracy and loss of training and validation among epochs for the best model. The training accuracy was linearly increasing with decreasing loss. As well increasing the accuracy of validation and decreasing loss.

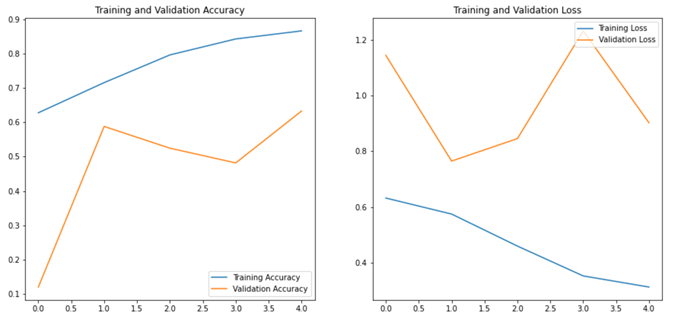


Figure2: Accuracy and loss for training and validation sets

Figure3 shows the improvement in accuracy among five model with increasing the number of convolution layers. Figure4 shows the five AUC for predicting the development of Diabetic Macular Edema were 0.63, 0.74, 0.75, and 0.79. The best AUC was0.82 for the best model.

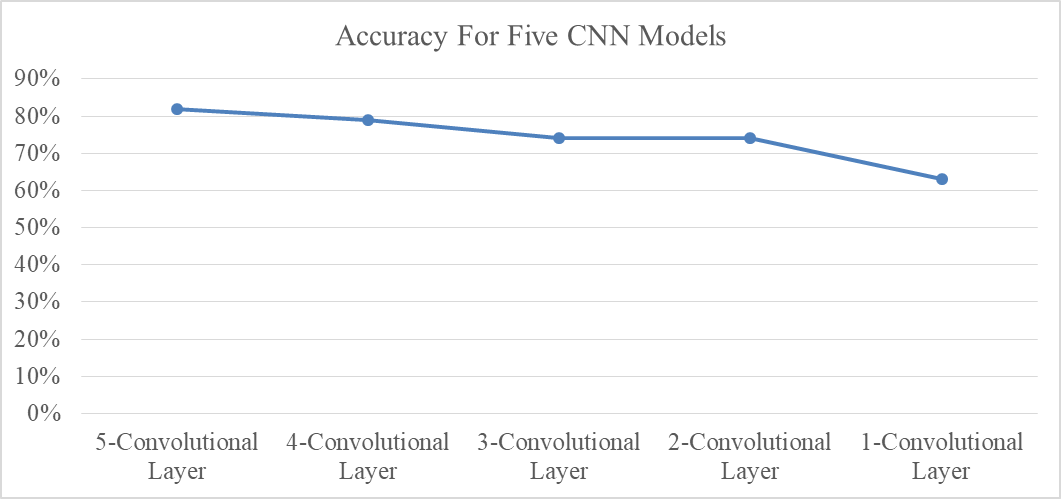


Figure3: The Accuracy achieved for five CNN models

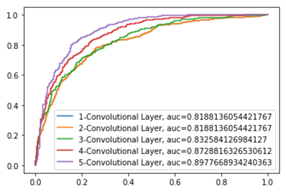


Figure4: The AUC achieved for five CNN models

**Conclusion**

We propose a deep learning based method to solve the classification problem in OCT images. Validation results show that the proposed method has been successful in predicting the presence of DME in OCT images.