Name: Sanayya

Email address: sanayya1998@gmail.com

Contact number:

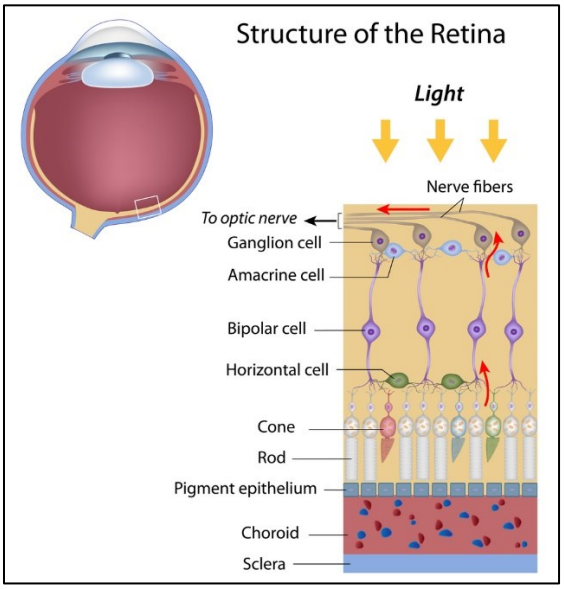
AnyDesk address:

Years of Work Experience: 0

Date: 04-04-2022

Self-Case Study -2: DL Based Approach for Diagnosis of Retinal Diseases from OCT Images

Overview

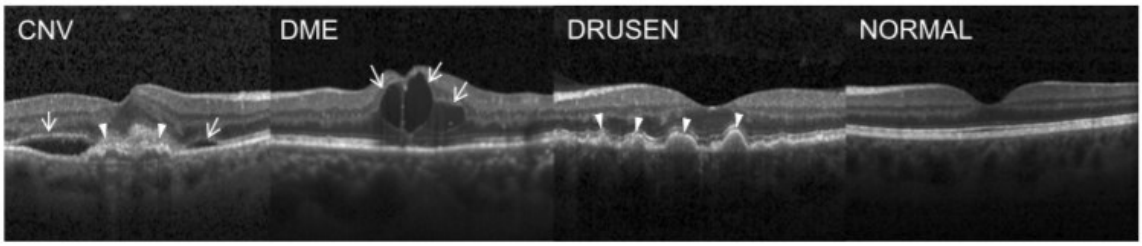


Layers of the Retina (https://discoveryeye.org/layers-of-the-retina/)

1. Retinal optical coherence tomography (OCT) is an imaging technique for capturing high-resolution cross sections of living patients' retinas.
2. OCT enables the ophthalmologist to visualise the retina's different layers. This lets the doctor map and measure their thickness. These analysis aid in the diagnosing process.
3. OCT testing has become the standard way to look at and treat most eye problems. OCT uses rays of light to measure the thickness of retina. This examination does not include the use of radiation or X-rays, and an OCT scan is neither painful nor uncomfortable.
4. Every year, around 30 million OCT scans are taken, and the analysis and interpretation of these pictures takes a lot of time (Swanson and Fujimoto, 2017).
5. OCT is beneficial for identifying a variety of eye disorders, including the following:
    - age-related macular degeneration (AMD)
    - diabetic retinopathy (diabetic eye disease)

Dataset Description

The dataset can be downloaded from https://www.kaggle.com/datasets/paultimothymooney/kermany2018?datasetId=17839&sortBy=voteCount.
The dataset is divided into three folders (train, test, and validation), each of which contains a subfolder for each image category (NORMAL, CNV, DME, DRUSEN). There are 84,495 X-Ray images (JPEG) and four classes (NORMAL, CNV, DME, DRUSEN).
Images are labeled as (disease)-(randomized patient ID)-(image number by this patient) and split into 4 directories: CNV, DME, DRUSEN, and NORMAL.



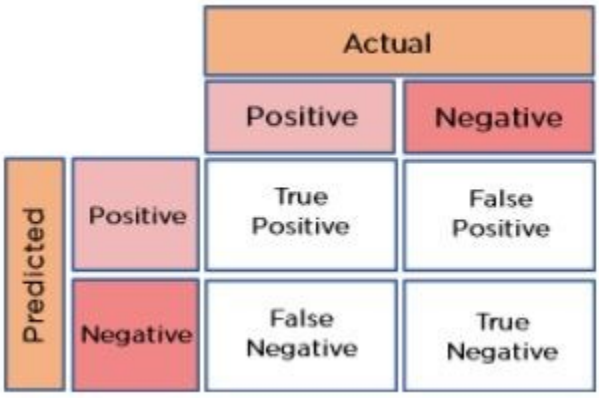
Representative Optical Coherence Tomography Images and the Workflow Diagram [Kermany et. al. 2018] http://www.cell.com/cell/fulltext/S0092-8674(18)30154-5

(Far left) choroidal neovascularization (CNV) with neovascular membrane (white arrowheads) and associated subretinal fluid (arrows).
Choroidal neovascularization (CNV) is the formation of new blood vessels from the choroid into the subretinal pigment epithelium (sub-RPE) or subretinal space via a rupture in the bruch membrane. CNV is a major factor of vision loss.
(Middle left) Diabetic macular edema (DME) with retinal-thickening-associated intraretinal fluid (arrows).
DME is a complication of diabetes. Individuals with type 1 or type 2 diabetes are at risk of developing DME. DME occurs when excess fluid starts to build up in the macula of the eye. The macula is responsible for our ability to focus and see small details. It is placed in the middle of the retina, the blood vessel-filled lining in the back of the eye. When excess fluid builds in the macula, visual problems happen.
(Middle right) Multiple drusen (arrowheads) present in early AMD.

Drusen are yellow deposits that develop behind the retina. Drusen are composed of fatty acids and proteins. There are different kinds of drusen. Drusen that are larger in size increase the likelihood of developing advanced AMD, which can result in vision loss.
(Far right) Normal retina with preserved foveal contour and absence of any retinal fluid/edema.
When light is focused directly on the retina rather than in front of or behind it, normal vision occurs. A person with normal eyesight is capable of seeing objects both near and far.

Problem Statement
The objective is given retinal OCT Images (optical coherence tomography) classify a new retinal image belonging to one of the four categories (CNV, DME, DRUNSEN, NORMAL).

Evaluation Metric
Confusion Matrix:



Confusion Matrix (https://www.simplilearn.com/tutorials/machine-learning-tutorial/confusion-matrix-machine-learning)
Our model's performance can be diagnosed more accurately with the help of confusion matrix. A confusion matrix is a summary of classification problem prediction outcomes. The number of correct and incorrect predictions is summarized with count values and broken down by each class. The confusion matrix depicts the various ways in which our classification model is perplexed when making predictions. This technique sheds light not only on the errors made by our classifier, but also on the types of errors caused by the models. The confusion matrix can be used to calculate the values of true positive (TP), true negative (TN), false positive (FP) and false-negative (FN).
TP stands for correctly predicted positive class;
FP stands for incorrectly predicted positive class;
FN stands for incorrectly predicted negative class; and
TN stands for correctly predicted negative class.

Research-Papers/Solutions/Architectures/Kernels

[1]. https://www.intechopen.com/chapters/26340

- The development of OCT in the 1990s represents a significant advancement in ocular imaging, as it is a non-invasive, radiation-free examination approach capable of precisely seeing the retinal layers.
- OCT is a technique for high-resolution imaging that is classified into SD-OCT and TD-OCT. SD-OCT images produce a high-resolution cross-sectional and volumetric view of the retina. TD-OCT provides a two-dimensional image of the internal anatomy of the retina.
- The study discovered that OCT is an effective approach for studying, monitoring, and assessing the various stages of AMD.
- Additionally, drusen can be evaluated based on a variety of structural properties.

[2]. https://opg.optica.org/boe/fulltext.cfm?uri=boe-8-2-579&id=357053

- A transfer learning method was presented for identifying retinal diseases utilizing retinal OCT images and the inception network.
- The dataset included OCT images of patients with dry AMD, DME, and healthy participants.
- Their finding revealed that when compared to classical learning approaches, the fine-tuned CNN was effective at identifying abnormalities.
- The classifying OCT method demonstrated with limited training data and trained on non-medical pictures can be fine-tuned.
- The average prediction accuracies for normal, AMD, and DME were 99%, 89%, and 86%, respectively.

[3]. https://opg.optica.org/boe/fulltext.cfm?uri=boe-5-10-3568&id=301172

- This research described a classification strategy based on support vector machine (SVM) classifiers and Histogram of Oriented Gradients (HOG) descriptors, which was used to detect dry AMD and DME using OCT imaging.
- The inner retinal layers were not segmented in the proposed strategy.
- SD-OCT datasets included 45 volumetric images, 15 normal, 15 AMD, and 15 DME.
- The algorithm has the best specificity and sensitivity, detecting 100% of AMD cases, 100% of DME cases, and 86.67% of normal cases.

[4]. https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/s12938-017-0352-9

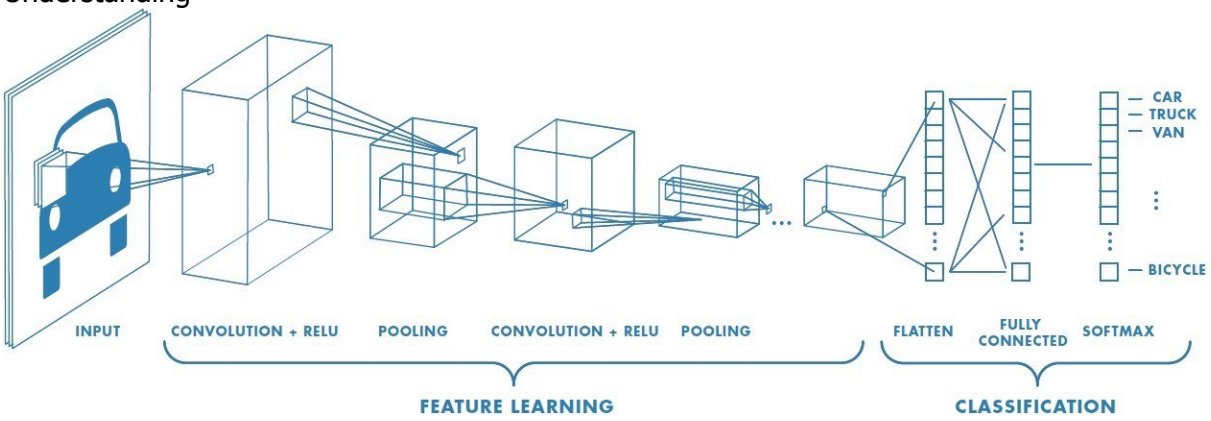
- This paper depicts an automated classification framework for detecting DME in large amounts of SD-OCT imaging data.
- Their procedure consisted of the following steps: preprocessing, feature detection, feature representation, and classification.
- The SVM and principal component analysis achieved an 87.5% sensitivity and 87.5% specificity.
- LBP-ri vectors were important in achieving the most successful result in illness classification.

[5]. https://www.aaojournal.org/article/S0161-6420(17)31424-0/fulltext

- Using OCT images, this research offered a DL-based technique for detecting distinct forms of fluids in the retina across various macular disorders.
- Their dataset included OCT scans from 1200 individuals, 400 of whom had AMD, 400 of whom had DME, and 400 of whom had RVO (retinal vein occlusion).
- This fully automated method was created to quantify and detect subretinal fuid and intraretinal cystoid and attained a mean accuracy of 0.94 with 0.91 precision and 0.94 recall value.

[6] Convolutional Neural Network

Understanding-



CNN Architecture (https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53)

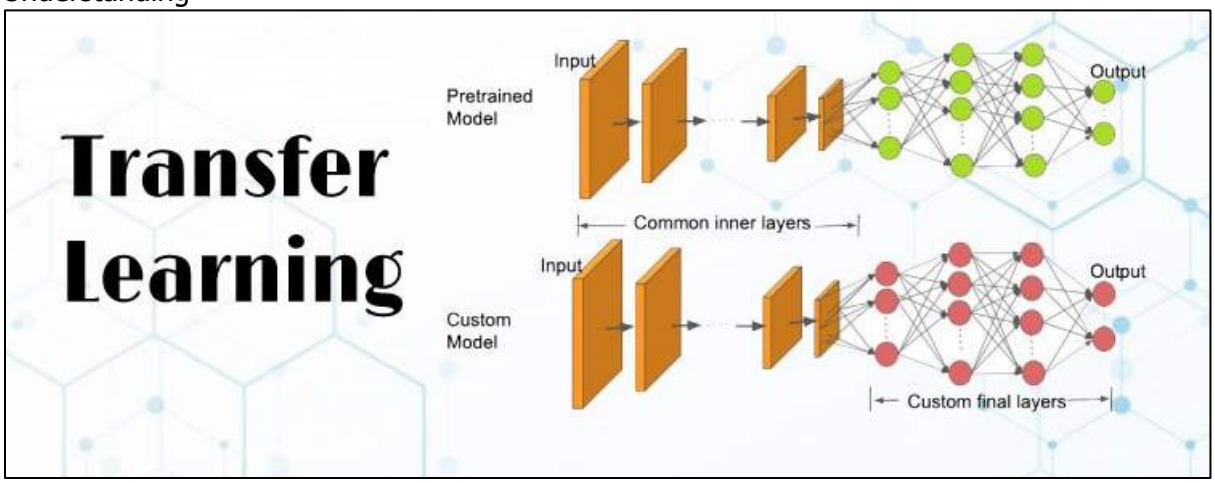
- A convolutional neural network is a feed-forward neural network that is commonly used to evaluate visual pictures by processing data in a grid-like manner. It's also referred to as a ConvNet.
- CNN takes an image as input and classifies and processes it into one of several categories. The computer reads an image as an array of pixels, which is dependent on the image's resolution. It will appear as h * w * d, where h = height, w = width, and d = dimension, depending on image resolution.
- Each input image in CNN will go through a series of convolution layers, as well as pooling, fully connected layers, and filters (also known as kernels). The Soft-max function will then be used to classify an item with probabilistic values of 0 and 1.
- In a nutshell, the ConvNet's job is to compress the images into a format that is easier to handle while preserving elements that are important for making a decent prediction.

References-

- https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/
- https://www.simplilearn.com/tutorials/deep-learning-tutorial/convolutional-neural-network

[7]. Transfer Learning

Understanding-



Transfer Learning Architecture (https://www.analyticsvidhya.com/blog/2021/10/understanding-transfer-learning-for-deep-learning/#h2_3)

- Transfer learning refers to the use of a previously trained model on a new task.
- In transfer learning, a machine uses knowledge from a previous assignment to improve prediction about a new task.
- During transfer learning, the information of an already trained machine learning model is transferred to a separate but closely related task.
- For example, if we trained a simple classifier to predict whether an image contains a backpack, you could use the model's training knowledge to identify other objects such as sunglasses.
- In a nutshell, transfer learning is the process through which we attempt to apply what we've learnt from one task to another in order to gain a deeper understanding of the underlying principles.

References-

- https://www.analyticsvidhya.com/blog/2021/10/understanding-transfer-learning-for-deep-learning/#h2_3
- https://towardsdatascience.com/what-is-deep-transfer-learning-and-why-is-it-becoming-so-popular-91acdcc2717a

---

First Cut Approach

Based on the research and readings that we have done. We will follow the below steps-

1. Using EDA, we'll build summary statistics for the dataset and construct multiple graphical representations to understand the data better.
2. Data preprocessing is one of the most critical aspects of this Kaggle problem. We will perform image augmentation i.e., we will take existing photos from our training dataset and apply image transformation operations to them, such as rotation, shearing, translation, zooming, and so on, to create new, altered versions of those images.
3. Since we have 84,495 images in this dataset, and we can't fit them all in primary memory, so we will load them and convert them into arrays batch by batch.
4. We'll apply various deep learning models, including 3 CNN (convolutional layer) with 3*3 kernel, 'adam' optimizer, and "relu" activations, 7 CNN (convolutional layer) with 3*3 kernel, 'adam' optimizer, and "relu" activations, custom VGG16 model, custom ResNet50 model, and custom DenseNet121 model.
5. In the end, we will create a simple GUI for the model that performed the best, using a relatively new but popular open-source app framework called Streamlit and deploy it.