

Convolution Neural Network Approach to Diagnosing Diabetic Retinopathy

José Solomon *

Abstract

This is the first draft of the project report of the diabetic retinopathy diagnosis using convolution neural networks (ConvNN). It is derived from the Kaggle Diabetic Retinopathy [1] aimed at creating a robust algorithm to automate the process of classifying the level of diabetic retinopathy found in retinal scans. the purpose of this initial draft is primarily to describe how the ConvNN work, and the inter-dependencies of its various functional elements. In the final draft, details of the implementation will be given and various resulting performance metrics will be reviewed.

Contents

1	Introduction	1
2	Diabetic Retinopathy	1
2.1	The Kaggle Grand Challenge	2
3	Convolution Neural Networks	3
3.1	The Convolution Layer	3

1 Introduction

There are two key facets to the project: first is the concept of the diabetic retinopathy and it's diagnosis; the second is ConvNN and it general functional principles as applied to digital imaging processing and categorization. We begin with a cursory description of the former, and then focus the bulk of the theoretical discussion on the latter.

2 Diabetic Retinopathy

Diabetic retinopathy (DR) is a disease that generally afflicts those who have dealt with diabetes for a period of 5 years or more [2]. It is defined specifically as the deterioration of blood vessels that feed the retina of the human eye, as illustrated in Figure 1 below. The primary cause of DR is that as a byproduct of the condition of diabetes, thinner blood

*jose.e.solomon@gmail.com

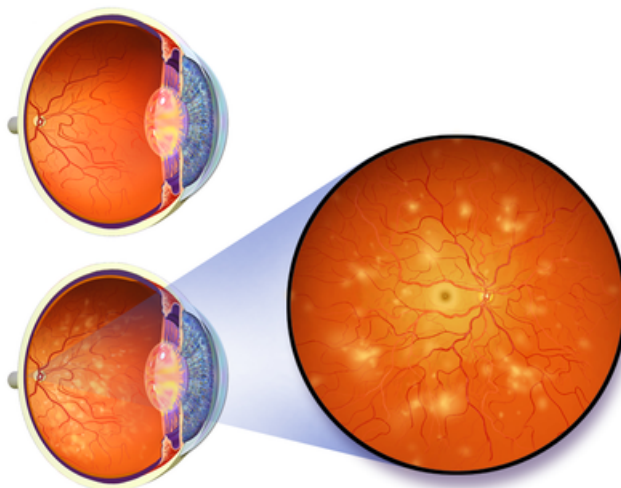
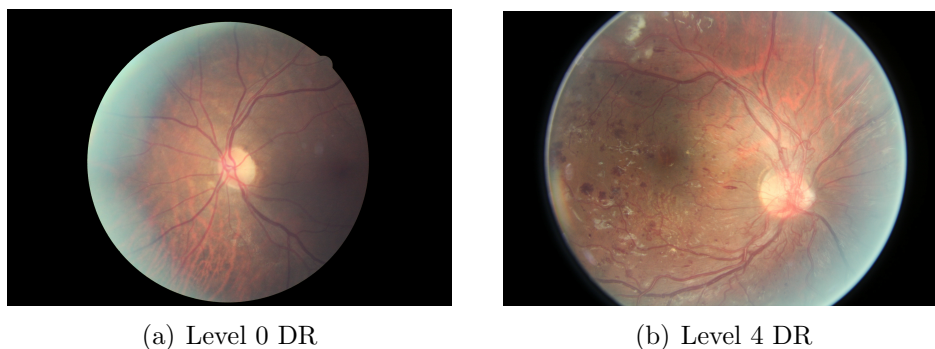


Figure 1: Diabetic retinopathy in comparison to a normal eye [3]

vessels in the human body tend to form abnormal branching structures and thinning walls. This leads to either deteriorated blood supply to the effected areas or actual hemorrhaging. In terms of the retina, this leads to blind spots forming in the field of vision of the afflicted individual. To diagnose the condition, a retinal scan of the eye is taken and a classification is assigned depending on the number of abnormal vein branching, wall thickness of the blood vessels and blood stains observed. DR is categorized on a scale from 0 to 4, and examples of level 0 and 4 DR are given below. It should be noted that this method of categorizing DR is not fundamentally rigorous in nature and is more of a qualitative scale.



(a) Level 0 DR

(b) Level 4 DR

Figure 2: Examples of level 0 and level 4 DR

2.1 The Kaggle Grand Challenge

Kaggle has presented the automated DR diagnosis as open grand challenge. In doing so, Kaggle provided 35 gigabytes of retinal scan data, over 35,000 images, thus each image is roughly 1.5 megabytes in size. To facilitate download and processing of the data set, Kaggle

divided the image set into a training group, and a testing group, and the subsequently divided each group into smaller subsets to facilitate file transfer. All the results presented here and in subsequent discussions are based on the first training set of images.

It is noted that the aim of the current work is to understand ConvNN from a fundamental perspective, and not to actually create a competitive solution for the grand challenge. Due to the size of the reference data set, and the limitations imposed by Python in terms of loading elements into shared memory, it would be very difficult to create a code base that could compete directly with other implementations in CuDa, (a GPU-level system language), which is defacto favorite language for image processing.

3 Convolution Neural Networks

ConvNN are a powerful machine learning technique that fall under the of general premise of deep learning. There are number of flavors of ConvNN, but the specific implementation presented here is one of the first implementations of the technique, the LeNet-5 [4], and is especially adept at processing digital imaging.

The LeNet-5 consists of 3 primary component layers: the convolution layer, the pooling layer, the conventional fully connected layer [5].

3.1 The Convolution Layer

The convolution layer is the work horse of the ConvNN, at is what makes it such an indispensable tool for image processin. Mathematically, the concept of convolution is somewhat straightforward, and plays a key role in spectral methods and Fourier space treatments. In terms of digital imaging, convolution can be expressed as [5] the product of the pixel intensity of a given image with that of a kernel, and is stated as

$$H[m, n] * G[m, n] = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} H[u, v] G[m - u, n - v] \quad (1)$$

This may seem somewhat convoluted at first, but consider that an image is an array of pixels, each with an intensity that ranges from 0 up to a given bit depth, (e.g. 255 for 8-bit images), than we can see that an image can be translated as matrix, where each element is a given pixel's image intensity. The concept is illustrated in Figure ?? A kernel, in the context of digital imaging, is itself a matrix of a prescribed dimension, such as

$$H_{3 \times 3}^L = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (2)$$

which is applied, via convolution across the matrix of the image.



(a) original image (b) Region of red rectangle (c) Gray-scale value

Figure 3: Digital image as a matrix [?]

References

- [1] Kaggle. <https://www.kaggle.com/>.
- [2] Nih national eye institute. <https://nei.nih.gov/health/diabetic/retinopathy>.
- [3] Diabetic retinopathy. http://en.wikipedia.org/wiki/Diabetic_retinopathy.
- [4]
- [5] Theano convolution network. <http://deeplearning.net/tutorial/lenet.html>.