6CS012: Artificial Intelligence and Machine Learning

CATARACT DETECTION USING CNN

Student Id	: 2051891
Student Name	: Upshot Awal
Section	: L6CG7
Lecturer	: Siman Giri
Submitted on	: 2nd May 2022

Abstract

This project is based to save the item and money for people who are unable to seek medical supervision. Around 285 million people worldwide have a vision impairment, according to the World Health Organization. We used CNN because on bigger datasets, convolutional neural networks outperform other state of the art methods. They're made up of neurons with adjustable weights and biases.

Introduction

Brief Summary for CNN

Convolutional neural networks are regularized versions of multilayer perceptron (MLP) neurons with learnable weights and biases. Each neuron gets a large number of inputs and then computes a weighted sum of them before passing them through an activation function and responding with an output. CNNs employ visual recognition and classification to find objects and recognize faces among other things.

Red-blue-green (RGB) encoding is used in digital color pictures. They are a rectangular box with a width and height determined by the number of pixels in those dimensions. Such pictures are ingested by a convolutional network as three independent strata of color placed one on top of the other. Channels are the depth levels in the three layers of colors (RGB) interpreted by CNNs. (Bansari, 2019) The convolution layer is the initial layer of a CNN network, and it serves as the basis and does the majority of the computational heavy lifting. To convolve data or pictures, filters or kernels are utilized. Filters are little units that we use to apply to data via a sliding window. The result of a convolution with a 3d filter and color would be a 2d matrix. Now, envision a torch beaming over the top left corner of the image to demonstrate a convolutional layer. Consider shining a spotlight on the upper left corner of the image. Consider this flashlight moving over all of the regions of the input image. This filter is also a number array (the numbers are called weights or parameters). (Chatterjee, 2019)

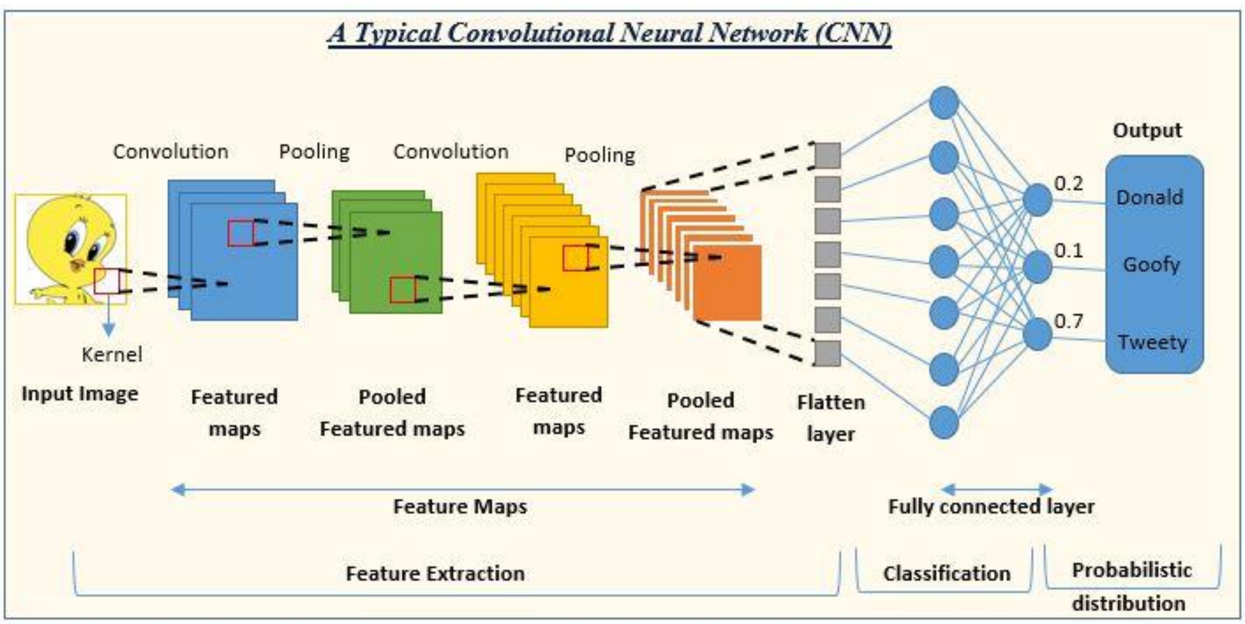


Figure 1 Work flow of Convolutional Neural Network

Different Layers of CNN

- Convolutional Layer
  The dot product is a convolutional mathematical operation performed between the input image and an MxM-sized filter. This is the initial layer that is used to extract different attributes from the input photographs. The resulting Feature Map comprises information about the image, such as its corners and edges. Later, this feature map is fed through other layers, which learn various image features. (Gurucharan, 2020)

- Pooling Layer
  The biggest piece from the feature map is used in Max Pooling. A Convolutional Layer is usually followed by a Pooling Layer. This layer's major goal is to lower the size of the convolved feature map in order to reduce computational expenses. Pooling activities are classified into numerous kinds based on the approach utilized. (Gurucharan, 2020)

- Fully Connected Layer
  The Fully Connected (FC) layer, which includes weights and biases as well as neurons, is used to connect neurons from different layers. These layers are often placed prior to the output layer and constitute the final few levels of a CNN Architecture. The preceding layers' input images are flattened and supplied to the FC layer in this step. (Gurucharan, 2020)

- Dropout
  Overfitting happens when a model performs so well on training data that it has a detrimental influence on the model's performance when applied to fresh data. To lower the size of the model, a dropout layer is used, in which a few neurons are removed from the neural network during the training process. When a dropout of 0.3 is reached, 30% of the nodes in the neural network are dropped out at random. (Gurucharan, 2020)

- Activation Functions
  The activation function is an important component of the CNN model. They are used to approximate any form of continuous and sophisticated network variable connection. Various activation functions, such as the Softmax and Sigmoid functions, are often used. For a binary classification CNN model, the sigmoid and softmax functions are used. (Gurucharan, 2020)

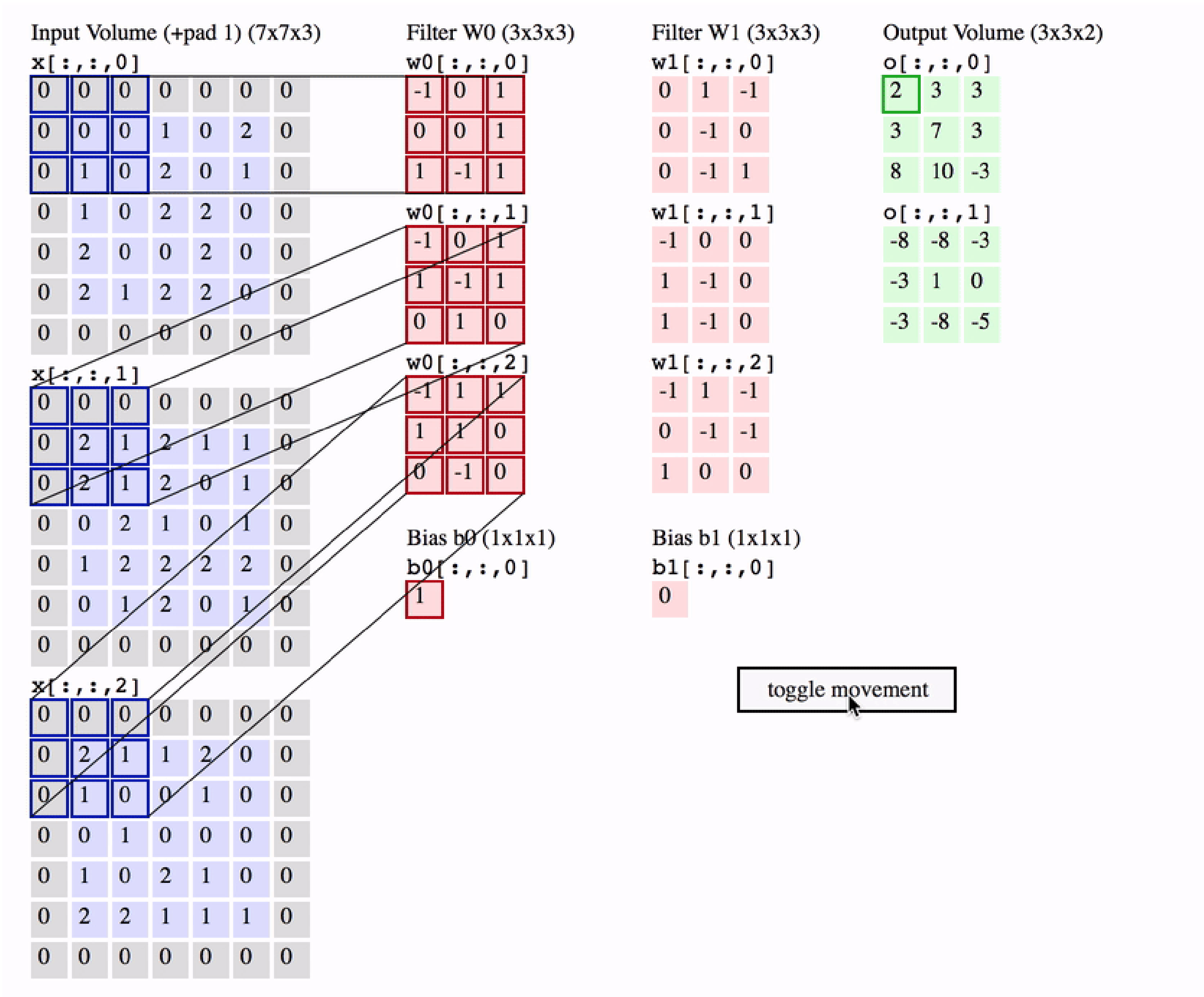


Figure 2 Convolutional Layer Filtering

Use of CNN for Cataract Detection

Computational intelligence is one of the most powerful methods for handling complicated challenges in image processing. There are various image classification algorithms for detecting and evaluating cataracts. In this paper, we examine the key properties of several grading and classification systems. These approaches are discussed in a straightforward manner, with the goal of providing a means of understanding the working principles and summarizing their pertinent applications. (Hans Morales-Lopez, 2020) Cataract is one of the most frequent vision-distorting eye conditions. The greatest strategy to control the danger and avert blindness is to identify cataracts accurately and on time. Artificial intelligence-based cataract detection systems have recently gotten a lot of interest in the scientific world. Cataract is a lenticular opacity that obscures the clear lens of the human eye. Cataracts cause light to get obstructed and not reach the lens, resulting in blurred vision. It is the most common eye illness in the world, yet it does not damage vision until later in life. After a time, it can obstruct eyesight and possibly cause vision loss in persons over the age of 40. Early identification of cataracts may help to avoid unpleasant and costly operations. Cataracts were responsible for 33% of vision impairment and 51% of blindness. According to the World Health Organization, around 285 million individuals worldwide have a visual impairment. (MASUM SHAH JUNAYED, 2021) In summary the main aim of this project is to identify the cataract disease to help the people who seek medical attention, or eye treatment in low cost and without full supervisor of doctors.

Similar Working Application

Cataract Net is a system for automatically detecting cataracts in fundus pictures. In this paper, Cataract Net, a lightweight deep learning-based automated cataract detection system, was developed. A cataract dataset of fundus images was rearranged, pre-processed, and expanded to improve the dataset for feeding the deep network. In order to minimize processing costs while retaining model accuracy, the Cataract Net project studied multiple layers, activation functions, loss functions, and optimization strategies. Cataract Net beat all five pre-trained CNN models when compared to them. The activation and loss functions are modified to train the network using tiny kernels, fewer training parameters, and layers. As a result, Cataract Net’s computational cost and average operating time are much lower when compared to other pre-trained Convolutional Neural Network (CNN) models. (MASUM SHAH JUNAYED, 2021)

Methodology

Model Summary

```
model.summary()

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 192, 256, 16)      448

 conv2d_1 (Conv2D)           (None, 192, 256, 16)      2320

 batch_normalization (BatchN  (None, 192, 256, 16)     64
 ormalization)

 max_pooling2d (MaxPooling2D  (None, 64, 85, 16)       0
 )

 dropout (Dropout)           (None, 64, 85, 16)        0

 conv2d_2 (Conv2D)           (None, 64, 85, 16)        2320

 conv2d_3 (Conv2D)           (None, 64, 85, 16)        2320

 batch_normalization_1 (Batc  (None, 64, 85, 16)       64
 hNormalization)

 max_pooling2d_1 (MaxPooling  (None, 21, 28, 16)       0
 2D)

 dropout_1 (Dropout)         (None, 21, 28, 16)        0

 flatten (Flatten)           (None, 9408)              0

 dense (Dense)               (None, 2)                 18818

=================================================================
Total params: 26,354
Trainable params: 26,290
Non-trainable params: 64
_________________________________________________________________
```

Figure 3 Model Summary

Our model has 12 layers, each layer serves a purpose and acts as a filter for classification of cataract disease. There are 4 convolutional layers,2 max pooling layers, 2 dropout layers, 2 batch normalization layer and an each flatten and dense layers. The parameters used in this models are epochs and SEEDS and for sequence models we can use models like activation, padding, and kernel size.

Training of model

The loss function is used to determine the quantity that the model should strive to decrease during training. The mean squared error function is a regularly used loss function in regression models. Cross entropy is the most often used loss function in classification models that predict probability. The lost function used in this model is categorical-cross entropy which is a part of cross-entropy.

The cross-entropy loss function is an optimization function used for training machine learning classification models that classifies data by calculating the likelihood (value between 0 and 1) that the data belongs to one of two classes. When the anticipated likelihood of class differs greatly from the actual class label, the degree of cross-entropy loss is considerable (0 or 1). It is most typically used as a cost function in logistic regression models or softmax models (multinomial logistic regression or neural network). When there are two or more output labels in a multi-class classification model, this function is used as a loss function. (Kumar, 2020)

The optimizer used in this model is Adam. Adam optimization approach is a stochastic gradient descent variant that has lately gained popularity for deep learning applications in computer vision and natural language processing. The key advantages of the Adam optimizer are that it is simple to build, computationally efficient, requires little memory, and is insensitive to gradient diagonal rescale.

An epoch is made up of one or more batches in which we train the neural network using a portion of the dataset. During an epoch, we only utilize all of the data once. A forward pass and a backward pass are combined to count as one pass. Iteration refers to the process of going through the training examples in a batch. (baeldung, 2021) We have used 10 epochs for training the model.

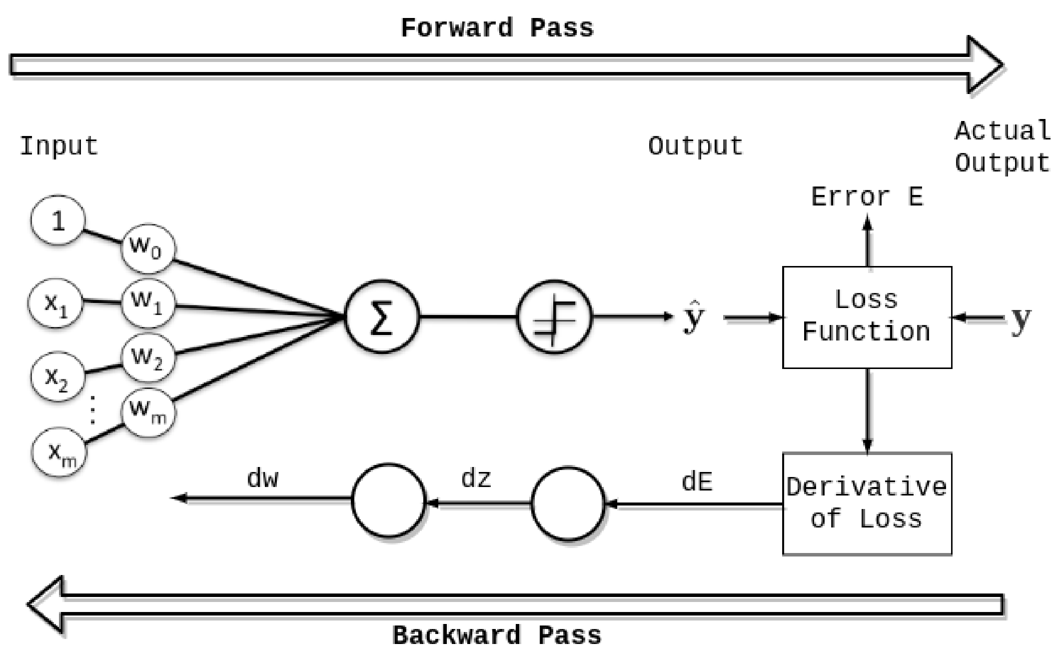
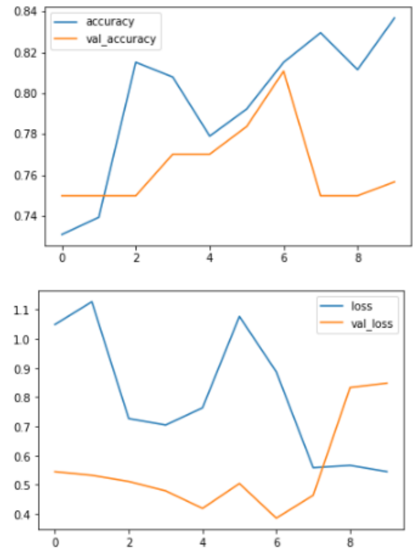


Figure 4 Epochs Work Flow

After training the model in both training and validation dataset. We have plotted two histogram plots for accuracy and loss for both train and validation dataset.




```
In [23]: model.evaluate(test_imgs, test_df)
8/8 [==============================] - 7s 851ms/step - loss: 0.9058 - accuracy: 0.7439
Out[23]: [0.905763566493988, 0.7439024448394775]
```

Figure 5 Validation score of Accuracy and Loss

Findings and Discussion

Convolutional neural networks outperform other state-of-the-art algorithms on larger datasets. CNNs, like artificial neural networks, are biologically inspired (ANN). They are composed of neurons that have programmable weights and biases. When a neuron receives an input, the information is multiplied by the weight that has been allocated to it. Convolutional neural networks outperform other state-of-the-art algorithms on larger datasets. They are composed of neurons that have programmable weights and biases. When a neuron receives an input, the information is multiplied by the weight that has been allocated to it.

References

baeldung. (2021). Epoch in Neural Networks.

Bansari, S. (2019). Introduction to how CNNs Work.

Chatterjee, C. C. (2019). How a classic CNN (Convolutional Neural Network) work? *Basics of the Classic CNN* .

Gurucharan, M. (2020). Basic CNN Architecture: Explaining 5 Layers of Convolutional Neural Network. *Artificial Intelligence* .

Hans Morales-Lopez, I. C.-V.-M. (2020). *Cataract Detection and Classification Systems Using Computational Intelligence.* SpringerLink.

Kumar, A. (2020). Keras – Categorical Cross Entropy Loss Function.

MASUM SHAH JUNAYED, M. B. (2021). *CataractNet: An Automated Cataract Detection.* IEEE Access.