ABSTRACT

Cancer has been distinguished as a dangerous disease that has many alternative sub categories. Machine learning is employed widely in detection and diagnosis of cancer among patients. It’s important to classify cancer patients into high or low risk groups, such a big amount of research teams from biomedical field have studied the application of machine learning methods. These techniques have been appointed as an objective to model the progression and treatment of cancer conditions. Additionally the power of Machine learning tools to trace key features from complex datasets reveals their importance. Here we are using convolutional neural networks to predict the cancer qualities. Convolutional neural network is beneficial in analysing visual imagery. In this project, we use Convolutional Neural Network models that take unstructured organic phenomenon inputs like images to classify tumours and non-tumour samples into their cog nominate cancer types or as normal.

Keywords: Cancer, Machine learning, CNN

**I. INTRODUCTION**

**Machine Learning** is the area of study that makes computers capable to learn without being programmed normally. Machine Learning is one of the most rousing technologies in the present generation. As it is obvious from the name, it makes the computer to do the same things as humans one of such thing is the ability to learn. Machine learning algorithms helps us in such a way that we can build a [mathematical based model](https://en.wikipedia.org/wiki/Mathematical_model)  on sample dataset, called as "[training data](https://en.wikipedia.org/wiki/Training_data)", in order to make predictions without providing any explicit program to perform the task.

Machine Learning has also been proven as a vital area in biomedical research with many applications where an acceptable generalization is obtained. Machine learning techniques mostly differ by the nature of the function they use for predicting by searching through an n-dimensional space for a given set of biological samples, using different techniques and algorithms. Machine learning mainly consists of two methods that are (i) supervised learning and the other is (ii) unsupervised learning. In supervised learning we need labelled data set for training to map the desired output from given input, in unsupervised learning methods there are no labelled data sets.

Cancer is a disease in which some of the body’s cell grow uncontrollably and spread to other parts of the body. A tumour is an abnormal mass of cells. It is of two types they are (i) Malignant and (ii) Benign. A Malignant tumor means it is made of cancer cells, and it can expand in nearby tissues. In contrast a benign tumor is a mass of [cells](https://en.wikipedia.org/wiki/Cell_(biology)) that doesn’t expand to neighbouring tissues. To predict the cancer, we are using Convolutional Neural Network method in machine learning.

CNN is a deep learning algorithm which can take in an input image, assign importance to various objects in the image and be able to differentiate from the other objects. Most commonly applied to analysing visual imagery.

It is used commonly in image classification, such as identifying the space satellite images which contains roads or in some examples the classification of hand written letters and digits. It has other mainstream tasks such as image segmentation and signal processing. It has been used for understanding in Natural Language Processing (NLP) and also in the recognition of speech even though in NLP Recurrent Neural Nets are used. Convolutional Neural Networks can also be implemented as U-Net architecture that are essential in two almost mirrored CNN’s resulting in a CNN whose architecture is in a U-shape.

**2. RELATED WORK**

Cancer is a dangerous disease which is contributing to the second leading reason for death in the world. Cancer is the reason for the death on an average of one in six people. Research efforts are being made to cancer diagnosis and treatment techniques to lessen its impact on human health. Prediction of cancer mainly focus on cancer susceptibility, recurrence and prognosis. The aim of cancer detection is based upon the classification of tumor types and identification of markers of each cancer such that the machine can be built to identify the particular metastatic tumour type and detect cancer at early stage before it become worse in human body. They have been several methods that helps to detect the cancer in the earliest stage and can help to reduce the growth in the body. The diagnosis can be done based on the level of the cancer present.

*Cancer diagnosis*.

1. *Laboratory tests***:**  In Laboratory tests, such as urine test and test on blood sample, may help the doctor detect the presence of cancer. For example, in people with leukemia disease, a blood test known as complete blood count may reveal an unusual number or type of white blood cells in the body.
2. *Imaging tests***:**  It allow the doctor to examine the bones and internal organs in a non-infective way. In the diagnosing of cancer the imaging tests that are used commonly are computerized tomography (CT) scan, bone scan, magnetic resonance imaging (MRI), positron emission tomography (PET) scan, ultrasound and X-ray, among others.
3. *Biopsy***:** In biopsy, the doctor will collect a sample of cells for testing in the laboratory. There can be several ways for collecting of sample bit the procedure that is more suitable for the patient depends upon the type of cancer and its location. In most of the cases biopsy is the only way to diagnose the cancer.
4. *Physical exam***:**  In the examining process the doctor may feel areas of your body for lumps that may indicate the presence of tumor. In the physical examination doctor may look for abnormalities such as change in skin color or enlargement of an organ, which are the signs of presence of cancer. In the laboratory, when the cell samples are examined under the microscope the cancer cells looks less orderly and varying in size when compared to the normal cells that are more uniform and well organized.

*Disadvantages:*

* In modern diagnosis there are high chances for false cancer detection.
* It is linked to a natural anxiety of specialists to avoid overlooking cancer at earlier stages.

**3. PROPOSED WORK**

Convolutional Neural Networks have an unlike design than that of the regular Neural Networks. Regular Neural Networks alter an input by putting it through a network of hidden layers. The layers are built from set of neurons, neurons within the previous layer are connected to the new layer the connection is fully connected. They need been different models that were developed for the prediction of the cancer which has specific aspect to the precise variety of cancer cells.

Commonly, to capture the Low-Level features like edges, color, gradient orientation the primary ConvLayer is helpful. For the extra added layers, the architecture is capable of adapting the High-Level features furthermore, which provides us a network which contains the wholesome understanding of images within the data-set. Based upon how we use them.

In CNN on the input image we perform the Convolution by the assistance of a filter or a kernel. to grasp the filtering and convolution we'd like to scan the screen ranging from top left to right so moving bottom a touch after covering the width of the screen and this could be repeated until the full screen is scanned perfectly. When a computer reads a picture, it'll have the array of pixel values. It’ll see a 32 x 32 x 3 array of numbers reckoning on the resolution and size of the image. To possess a sample explanation allow us to consider we've got a color image in JPG form and its size is 480 x 480.

The output array are going to be 480 x 480 x 3. These numbers are given a worth from 0 to 255 which is able to describe a pixel intensity at that time. These numbers aren't useful to us within the image classification these are the sole inputs that are available to the pc. The basic idea is that we submit the pc the array of numbers and it'll give the output numbers that describe the probability of the image of the certain class.

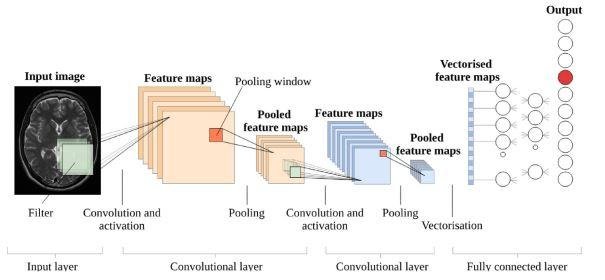
****

Fig 1. This figure illustrates the basic architecture of the Convolutional Neural Network.

The convolutional layer is that the first layer of the CNN which may extract features from the photographs by using kernel or filter. Because pixels are only associated with the adjacent and shut pixels, convolution operation allows us to preserve the connection between different parts of a picture. Convolution Operation is filtering the image with a smaller pixel filter or kernel to decrease the scale of the image without losing the link between pixels. after we apply convolution operation to the 5 x 5 image by employing a 3x3 filter with 1x1 stride (1-pixel shift at each step), we'll find yourself having a 3x3 output image.

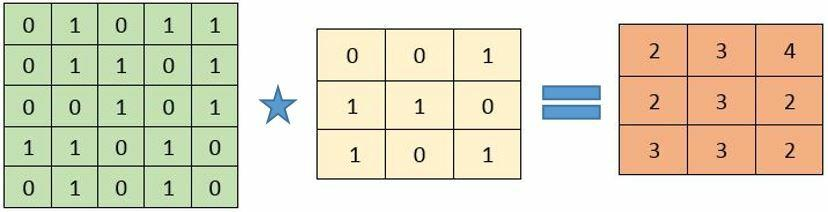
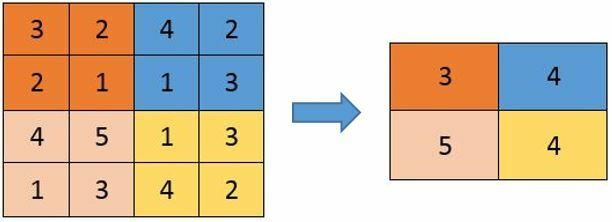


Fig 2. This figure illustrates the Convolution

The biological cortical region are the inspiration for the CNNs. its small regions of cells that are sensitive to the precise areas of the sight view. This concept was enlarged by an experiment team lead by Hubel and Wiesel in 1962. It had been showed by the researchers that some volunteer neurons within the brain activated or fired only within the presence of a selected orientation edges like vertical or horizontal edges. For instance, a number of the neurons fired when exposed to vertical sides and a few when shown a horizontal edge. They found that neurons were in well-ordered in an exceedingly columnar fashion and that they were capable to provide beholding. this concept of getting specialized components inside a system having specific tasks is one machines use yet and it may be find in CNNs.  
Convolution consist the property of being translational invariant.

  
Fig 3. This figure illustrates the Pooling

There are four steps in convolution:

* align the feature and therefore the image
* Perform multiplication with each image pixel by corresponding feature pixel
* Add up the values and calculate the sum
* Perform division on the sum by whole number of pixels within the feature.

Once the convolution layer is completed you get the feature maps, it's often to try and do pooling or a sub-sampling layer in CNN layers. The same as the Convolutional Layer, the Pooling layer is employed to scale back the spatial size of the Convolved Feature. It helps to cut back the computational power required to process the info through dimensionality reduction. It’s useful for uprooting the dominant features of the image which are rotational and positional invariant, thus maintaining the method of effectively training of the model. Pooling helps to cut back the training time and controls over-fitting.

There are two styles of pooling,

Max Pooling and Average Pooling, Max Pooling gives the best value from the part of the image covered by Kernel Max Pooling is additionally used as Noise Reducer. It eliminates the noisy activation altogether from the image and also it executes the de-noising together with dimensionality reduction within the image.

Average Pooling gives the typical of all the values from the part of the image covered by the Kernel. It executes dimensionality reduction as a noise reducing mechanism. Hence, we are able to say that Max Pooling performance is best than that of average pooling.

I-th layer of a Convolutional Neural Network is created by the convolutional layer and also the poling layer. Based upon the high differences within the images, the layers can also be able to increase for getting low-levels details even further, but the price of power for computation is high. After completing above process, we've got successfully empower the model to grasp the features. So, we are visiting compress the ultimate output and supply it to the regular Neural Network for classification.

**4. CONCLUSION**

In this article, CNN architecture can take high dimension cancer scanned image inputs and perform cancer type prediction while considering their tissue of origin. Our model achieved an equivalent 95.7% prediction accuracy. Whereas the earlier published studies were unable to provide such accuracy. The convolutional Neural Networks and the reducing of the Image Noise has help to achieve the accurate result. This allows us to implement a normal interpretation of our Convolutional Neural Network model to alert the cancer markers for each cancer type, with hope in future development that will lead to markers for earlier cancer detection.

**5. REFERENCES**

1. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2018. CA Cancer J Clin 2018, 68(1):7-30.

2. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, Douville C, Javed AA, Wong F, Mattox A et al: Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 2018, 359(6378):926-930.

3. Haendel MA, Chute CG, Robinson PN: Classification, Ontology, and Precision Medicine. N Engl J Med 2018, 379(15):1452-1462.

4. Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, Anagnostou V, Fiksel J, Cristiano S, Papp E et al: Direct detection of early-stage cancers using circulating tumor DNA. Sci Transl Med 2017, 9(403):eaan2415.

5. Schiffman JD, Fisher PG, and Gibbs P: Early detection of cancer: past, present, and future. Am Soc Clin Oncol Educ Book 2015:57-65.