Diagnosing COVID 19 using AI based medical image analyses

Problem Statement

Radiologists have watched the coronavirus disease 2019 (COVID-19) pandemic unfold. Radiology literature suggests a pivotal role for CT/X-RAY as CT/X-Ray findings in COVID-19 patient has pneumonia, and hence medical imaging has high sensitivity for diagnosis of COVID-19. The objective of this statement is to figure out usage of AI in medical imaging modalities for fast and accurate diagnosis of COVID-19.

Solution

The problem statement can be modeled using two approaches

1. Deep Learning(DL) driven approach
2. Natural Language Processing(NLP) driven approach

Environment and tools

1. scikit-learn
2. keras
3. numpy
4. pandas
5. Matplotlib
6. Tensorflow
7. nltk

Deep Learning(DL) driven approach

Dataset expected to be chest X-Rays of patient with few other features like fatalities, recovered, confirmed details.

A Convolutional Neural Network (CNN) is a deep learning algorithm that can recognize and classify features in images for computer vision. It is a multi-layer neural network designed to analyze visual inputs and perform tasks such as image classification, segmentation and object detection, which can be useful for autonomous vehicles. CNNs can also be used for deep learning applications in healthcare, such as medical imaging.
There are two main parts to a CNN:

- A convolution tool that splits the various features of the image for analysis
- A fully connected layer that uses the output of the convolution layer to predict the best description for the image.

CNN architecture is inspired by the organization and functionality of the visual cortex and designed to mimic the connectivity pattern of neurons within the human brain.

The neurons within a CNN are split into a three-dimensional structure, with each set of neurons analyzing a small region or feature of the image. In other words, each group of neurons specializes in identifying one part of the image. CNNs use the predictions from the layers to produce a final output that presents a vector of probability scores to represent the likelihood that a specific feature belongs to a certain class.

A CNN is composed of several kinds of layers:

- **Convolutional layer—**creates a feature map to predict the class probabilities for each feature by applying a filter that scans the whole image, few pixels at a time.
- **Pooling layer (downsampling)—**scales down the amount of information the convolutional layer generated for each feature and maintains the most essential information (the process of the convolutional and pooling layers usually repeats several times).
- **Fully connected input layer**—"flattens" the outputs generated by previous layers to turn them into a single vector that can be used as an input for the next layer.
- **Fully connected layer**—applies weights over the input generated by the feature analysis to predict an accurate label.
- **Fully connected output layer—** generates the final probabilities to determine a class for the image.

The architecture of a CNN is a key factor in determining its performance and efficiency. The way in which the layers are structured, which elements are used in each layer and how they are designed will often affect the speed and accuracy with which it can perform various tasks.

1. ImageNet
2. LeNet-5
3. GoogleNet
4. AlexNet
5. VGGNet
6. Inception v3
7. MobileNet are few popular CNN architecture which could be potentially used

Thus, this approach uses CNN based architecture to diagnose the COVID-19. Since, CNN has the ability to intuitively extract features from the image, it would be more useful to diagnose the needful.

Natural Language Processing(NLP) driven approach

The dataset expected to be in the form [image → captions]. The dataset consists of input images and their corresponding output captions.

Image Captioning is the process of generating textual description of an image. It uses both Natural Language Processing and Computer Vision to generate the captions.

Encoder

The Convolutional Neural Network(CNN) can be thought of as an encoder. The input image is given to CNN to extract the features. The last hidden state of the CNN is connected to the Decoder.

Decoder

The Decoder is a Recurrent Neural Network(RNN) which does language modelling up to the word level. The first time step receives the encoded output from the encoder and also the <START> vector.

Training

The output from the last hidden state of the CNN(Encoder) is given to the first time step of the decoder. We set x1 =<START> vector and the desired label y1 = first word in the sequence. Analogously, we set x2 =word vector of the first word and expect the network to predict the second word. Finally, on the last step, xT = last word, the target label yT =<END> token.

During training, the correct input is given to the decoder at every time-step, even if the decoder made a mistake before.

Testing

The image representation is provided to the first time step of the decoder. Set x1 =<START> vector and compute the distribution over the first word y1. We sample a word from the distribution (or pick the argmax), set its embedding vector as x2, and repeat this process until the <END> token is generated.

During Testing, the output of the decoder at time t is fed back and becomes the input of the decoder at time t+1

Thus, when the captions are generated, using those captions as features, convert the text to bow vectors or tf idf vectors by using preprocessing techniques. Then, use respective machine learning model on it to diagnose the needful.

With this link we have attached sample solution of many diseases using Inception V3 and MobileNET. Deep learning driven approach.