Lung Cancer Detection using Bilinear Convolution Neural Network

Taimur Zahid Maiza Razzaq Simra Waheed Hunble Dhillon Maryam Bashir

Department of Computer Science
National University of Computer and Emerging Sciences
Lahore, Pakistan
Email: { l134191, l134124, l134008, l134118 }@lhr.nu.edu.pk , maryam.bashir@nu.edu.pk

Abstract

Lung cancer is the second most common cancer in both men and women. Approximately 433 people die every day due to lungs cancer. The survival rate for lung cancer is 55 per cent for cases detected at an earlier stage when the disease is still localized. This research investigates the utility of Bilinear Convolutional Neural Network (CNN) for lung cancer detection. We have used dataset provided by Data Science Bowl competition on Kaggle Website by the National Immunotherapy Coalition (NIC) to determine whether a tumour in CT scan is cancerous or not. The neural network is preferred due to its notable performance in image classification by preserving the spatial structure and relation of pixels with the surrounding pixels of the image. Bilinear CNN helps in fined grained classification of images. Experimental results demonstrate that there is a significant improvement in classification accuracy using bilinear CNN as compared to conventional linear CNN.

Keywords: Convolution Neural Network, Bilinear CNN, lung cancer detection

I. INTRODUCTION

Lung Cancer is by far the leading cause of cancer death among both men and women. Each year, more people die of lung cancer than of colon, breast, and prostate cancers combined. About 1 out of 4 cancer deaths are from Lung Cancer. Early detection is critical to give patients the best chance of recovery and survival. The computer-aided diagnosis has proven to be very helpful in early detection by increasing the diagnostic accuracy [2, 4]. About a year ago, the office of the U.S Vice President lead a new initiative, the Cancer 'Moonshot' Initiative with the goal of finding vaccine-based biological therapies against cancer by the year 2020. The initiative is being managed by National Immunotherapy Coalition (NIC). The organization attempts to redefine how cancer should be tackled, by utilizing the body's own immune system and re-programming it to detect and destroy the body's cancer cells.

In the year 2017, the Data Science Bowl at Kaggle [1] completed a critical milestone in the support of the Cancer Moonshot initiative by bringing together data science and medical communities to develop lung cancer detection algorithms. The dataset, provided at the Kaggle website consists of thousands of low-dose CT images of high-risk patients in DICOM (Digital Imaging and Communications in Medicine) format. DICOM is a standard for handling, storing, printing, and transmitting information in medical imaging. The Kaggle Data Science Bowl challenge aimed to improve the performance of the current algorithms that determine whether lesions in the lungs are cancerous or non-cancerous.

Neuroimaging data poses a new challenge because it requires interdisciplinary knowledge of data science and medical sciences [9]. Machine Learning techniques are applicable for the detection of lung cancer because they can be trained to perform pattern classification in medical images [3]. In this research, we have implemented bilinear CNN for lung cancer detection using dataset provided by Kaggle Data Science Bowl. It was a challenging task as the difference between cancerous and non-cancerous lesions is minor. Bilinear CNN [12] has been shown to improve classification accuracy between images with minor differences. We have also tried different preprocessing techniques before passing images to CNN and the results show significant improvement after preprocessing. Major contributions of this research are as follows:

- Application of advanced machine learning techniques for solving a real-world problem (lung cancer detection) which can have an impact on saving patients lives.
- Application of bilinear CNN for the first time (to the best of our knowledge) to the problem of lung cancer detection from CT scan images.
- Preprocessing and analysis of high dimensional complex data (real patient's CT scan images) for the purpose of improved classification accuracy in lung cancer detection.

II. RELATED WORK

Machine Learning techniques are effective for the detection of lung cancer because classifiers can be trained to perform pattern classification in medical images [3]. Support Vector Machine (SVM) classifier [2] utilizes the training dataset to construct a hyperplane that divides the classes linearly. SVM has good performance in text-based classification but it is not explored in depth in medical diagnosis domain [5]. Naive Bayes (NB) classifier is simple, computationally efficient, and requires relatively little data for training. The authors in [5] emphasize that NB is one of the most efficient and effective classification algorithms through an empirical comparison of NB with five popular classifiers on 15 medical data sets. Moreover, it is also possible to train a cost-sensitive NB classifier to minimize the sum of misclassification costs and test costs [6]. The disadvantage of NB classifiers is that all attributes are assumed to be independent and it also has a hypothesis that the given data belongs to a particular class [5].

In the domain of computer-aided diagnosis, false positives and false negatives pose a great problem as the cost of misclassification can be both financial and human in nature [7]. Imaging science has grown primarily in distribution, listing, and judgment. The challenges associated with medical image processing includes the measurement and interpretation of the features that ultimately leads toward right or wrong classification. The most pervasive and productive machine learning model is the convolutional frameworks. It has found exceptional achievement in medical imaging applications due to good performance in image classification. Experiments have shown that they perform well in 2D as well as 3D segmentation. CNN minimize classification error by learning/extracting features while simultaneously training a classifier. In this research, we have used CNN for classification of medical images for lung cancer detection. Bilinear CNN [12] have been used in past research for fine-grained image classification of different bird species.

Deep convolutional neural networks (CNN) have proven to perform well in image classi cation [14, 20, 30], object detection [27], and other visual tasks. They have found great success in medical imaging applications [17], and are for example able to detect skin cancer metastases [34], achieving substantially better sensitivity performance than human pathologists. These methods all operate on two-dimensional images, typically a cross-sectional image of the affected body part.

III. METHODOLOGY

According to [8], there are four basic stages of the lung cancer detection system. In the first stage, image capture stage, the data i.e CT scans images of the lungs are collected. The second stage, **the** image enhancement stage, applies image preprocessing techniques to increase the quality of the images. In the third stage, Image thresholding and segmentation algorithms are applied. In the last stage, feature extraction stage, general features which indicate the normality or abnormality of lungs are extracted from the enhanced images. The two methods used for this purpose are:

- Binarization: This approach depends on the number of black versus white pixels. If black pixels in the segmented image are greater than white colour then the image is normal otherwise abnormal.
- Masking: This approach depends on the masses in lungs that appear as white connected areas inside the region of interest of the image. The blue colour of solid indicates normal case other than that indicates cancer.

Combining both the approaches it is concluded whether the case is normal or abnormal [8].

A. Architecture of CNN

A convolutional neural network is a classifier that takes an input object (image in our case) and outputs a label that best describes the object (image). It is a supervised machine learning algorithm. Figure 1 shows the architecture of CNN. The first layer of the CNN is the convolution layer. In this layer, features are detected from the input image by performing element-wise multiplication with the feature detector. The output of this layer is a set of feature maps. After convolution layer, the size of the image is reduced without any loss to the spatial relationships between pixels. The second layer is the ReLU layer which introduces non-linearity in the network. ReLU layer is followed by the pooling layer. In pooling layer, each region of the 3x3 pixel is represented by the max value (max pooling) in the region. This further reduces the image size, while preserving the features and reducing the number of parameters, which can help prevent overfitting. Pooling layer is followed by the flattening step which simply converts the 2-dimensional grid into a single dimensional array. The last layer of the network is the fully connected layer. In a fully connected layer, we attach an artificial neural network to the output of the flattening layer. This layer combines features to create more attributes that predict the classes. It works by giving weight to certain features that predict a certain class. Thus, these features have a higher vote for a certain class. Errors are back propagated to improve accuracy.

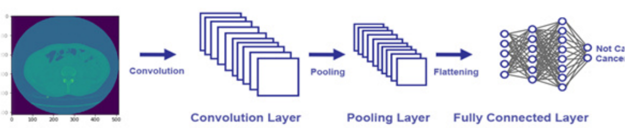

Figure 1: Convolutional Neural Network, [10] and [11]

B. Architecture of Bilinear CNN

The Bilinear Convolutional Neural Network, inspired by [12], used had similar preprocessing steps as mentioned in the Preprocessing Section. The additions in the preprocessing step were as follows:

- The dataset was resized to 64*64*64 dimensions instead of 32*32*32.
- By utilizing the Hounsfield Units of the CT scans, we were able to perform Lungs Segmentation by masking air, blood vessels, bones, and tissue. Hence we removed noise from the images.
- Resampling of the images was done using isomorphic resolution instead of random spacing, thus yielding a fixed pixel spacing of 1 mm.

The Bilinear CNN had a single input layer followed by two parallel architectures consisting of convolution, ReLU, and pooling layers followed by another convolution, ReLU, and pooling layers. The architecture is then merged at the outputs of the second convolution layer, where they are multiplied using the outer product at each location of the image and then pooled to obtain the bilinear vector before passing it to the ReLU activated fully connected layer. The details of the individual layer are similar to the CNN architecture. Figure 2 shows the architecture of bilinear CNN.

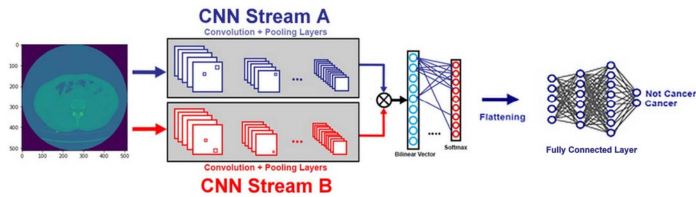

Figure 2: Bilinear Convolutional neural network

C. Preprocessing

The preprocessing phase is common among both the approaches with a few minor differences, so it is mentioned as a separate section preceding the approaches used. We used three datasets originating from three different sources. When combined, we had 3493 patients, where each patient has 2-dimensional slices of their chest cavity in DICOM format. Since the number of slices was different for all the patients, so the first step of preprocessing was to standardize the depth of the 3-dimensional images, from a depth of roughly 200 to 32. Next step was to resize the individual images, to obtain a smaller dataset. The individual slices were resized from 512x512 to 32x32. This was a mandatory step in order to reduce the processing overhead and to prevent overfitting. The final dataset composed of the CT scans with a dimension of 32x32x32 The preprocessing was concluded by creating an (numpy) array to send to the models for training.

IV. EXPERIMENTS AND RESULTS

A. Dataset

We have used dataset provided in the first competition of the Data Science Bowl, 2017 at Kaggle[1]. This dataset contains 1595 patients, where each patient has a different number of 2-dimensional slices of the chest cavity which, when combined, forms 3-dimensional images of the complete chest cavity of a particular patient.

We also worked with the LUNA'16 [14] dataset which is comprised of 888 patients and the LIDC/IDRI[13] dataset which is comprised of 1010 patients. In contrast with the Kaggle dataset which only provides whether or not a patient has cancer, these datasets provide information about the location and size of the nodules as separate annotation files along with the dicom images. However, these datasets didn't provide any information on whether or not a nodule was cancerous. Like the Kaggle[1] dataset, these datasets also posed the problem of having a different number of slices. This was taken care of in the preprocessing step.

B. *Parameters of CNN*

The Convolutional Neural Network used, had an input layer of 32*32*32 dimensions, following by three convolutions, pooling layers followed by a fully connected layer. The activation function used was ReLU after each convolution layer and in the fully connected layer. The information is summarized in Table 1 and Table 2.

Table 1: Convolution Layer Architecture of CNN

Layer Architecture	**First Convolution layer**	**Second Convolution layer**	**Third Convolution Layer**
Input	Single	32	64
Output	32	64	128
Sized window	3*3*3	3*3*3	3*3*3
Stride	1*1*1	1*1*1	1*1*1
Padding	same	same	same

Table 2: Pooling Layer Architecture of CNN

Layer Architecture	**First Pooling layer**	**Second Pooling layer**	**Third Pooling Layer**
Pooling limit	Max	Max	Max
Window	2*2*2	2*2*2	2*2*2
Stride	2*2*2	2*2*2	2*2*2
Padding	same	same	same

Table 3: Shape of the Tensor after the convolution and pooling operations

Convolution Function	**The shape of the tensor**
Convolution 1	(?, 31, 31, 31, 32)
Convolution 2	(?, 29, 29, 29, 64)
Convolution 3	(?, 12, 12, 12, 128)

Table 4: Shape of the Tensors after the convolution and pooling operations

Convolution Function	**Shape of the tensor**
Convolution 1_1	(?, 28, 28, 28, 32)
Convolution 2_1	(?, 28, 28, 28, 32)
Convolution 1_2	(?, 14, 14, 14, 64)
Convolution 2_2	(?, 14, 14, 14, 64)
Convolution 1_2	(?, 14, 14, 14, 64)

Table 5: Shape of the Tensors after matrix multiplication functions

Matrix Multiplication Function	**The shape of the tensor**
Convolution 1_2 transposed	(?, 64, 14, 14, 14)
Convolution 1_2_A after reshape	(?, 64, 175616)
Convolution 2_2 transposed	(?, 64, 14, 14, 14)
Convolution 2_2 after reshape	(?, 64, 175616)
Convolution 2_2_T transposed	(?, 175616, 64)
The shape of the bilinear vector after vector multiplication of the tensors 1_2_A and 2_2_T	(?, 64, 64)

We performed two experiments with the CNN architecture. There were no changes made to the preprocessing techniques, nor to the architecture itself. The only difference is that the first experiment only utilized the dataset available at Kaggle[1], whereas the second experiment couldn't work due to incompatibility issues. It utilized the datasets available at Kaggle [1], LIDC [13], and LUNA'16[14] datasets. The first experiment resulted in an accuracy of 55%.

Similar to the experiments with CNN architectures, we also performed two experiments with the Bilinear CNN. The first experiment had two different CNN architectures while the second had two similar architectures. Since in order to obtain the bilinear vector, we needed compatible vectors, so the experiment with different architectures was unsuccessful while the second approach resulted in an accuracy of 66%.

V. CONCLUSION

Automatic lung cancer detection at an early stage is an important area of research as lung cancer is one of the common causes of death among cancer patients. This paper investigated the utility of bilinear CNN for lung cancer detection. Bilinear CNN is popular for fine-grained image classification. They are appropriate for this task since there is a minor difference between CT scan images of cancerous and non-cancerous tumours. We have used real patient's data from Kaggle Data Science Bowl 2017.

Experimental results demonstrate that bilinear CNN significantly outperforms linear CNN for the challenging task of classification of cancerous and non-cancerous tumours for lung cancer detection. One limitation of bilinear CNN is that they require more memory as compared to linear CNN. In our experiments, we had to resize images in order to run bilinear CNN on CT scan images. This resizing caused information loss which can be prevented by using GPU (Graphics Processing Unit).

REFERENCES

[1] National Cancer Institute, Data Science Bowl 2017

[2] Cortes Corinna and Vapnik Vladimir, "Support-vector networks," *Machine Learning*, vol.20, issue.3, pp.273–297, September 1995.

[3] Akira Motohiro, Hitoshi Ueda, Hikotaro Komatsu, Noboru Yanai, and Takashi Mori, "Prognosis of non-surgically treated clinical stage I lung cancer patients in Japan," *Lung Cancer Journal,* vol.36, issue.1, pp.65-69, April 2002.

[4] Marios Anthopoulos, Stergios Christodoulidis, Lukas Ebner, Andreas Christe and Stavroula Mougiakakou, "Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network," *IEEE Transaction on Medical Imaging,* vol.35, May 2016.

[5] J. Alamelu Mangai, Jagadish Nayak and V. Santhosh Kumar; "A Novel Approach for Classifying Medical Images Using Data Mining Techniques," *International Journal of Computer Science and Electronics Engineering*, vol.1, issue.2 2013.

[6] Petr Berka, Jan Rauch and Djamel Abdelkader Zighed, "Ontologies in the Health Field," in *Data Mining and Medical Knowledge management: Cases and Application*, Hershey, United States: IGI Global, March 2011, pp. 37-56.

[7] Petr Berka, Jan Rauch and Djamel Abdelkader Zighed, "Cost-Sensitive Learning in Medicine," in *Data Mining and Medical Knowledge management: Cases and Application*, Hershey, United States: IGI Global, March 2011, pp. 57–75.

[8] Petr Berka, Jan Rauch and Djamel Abdelkader Zighed, "Classification and Prediction with Neural Networks," in *Data Mining and Medical Knowledge management: Cases and Application*, Hershey, United States: IGI Global, March 2011, pp. 76–107.

[9] Thomas Erl, Wajid Khattak, Paul Buhler, "Understanding Big Data," in *Big Data Fundamentals*, Crawfordsville, Indiana; USA: RR Donnelley, December 2015, pp. 21–35.

[10] Kirill Eremenko, Hadelin de Ponteves(March 28, 2017), *Deep Learning A-Z™: Hands-On Artificial Neural Networks*[Online].Available: https://www.udemy.com/deeplearning/learn/v4/overview

[11] Kirill Eremenko, Hadelin de Ponteves, *Machine Learning A-Z™ : Hands-On Python & R In Data Science*[Online]. Available: https://www.udemy.com/machinelearning/learn/v4/overview

[12]Tsung-Yu Lin, Aruni RoyChowdhury, Subhransu Maji, "Bilinear CNN Models for Fine-Grained Visual Recognition," in IEEE International Conference on Computer Vision(ICCV), 2015, pp. 1449-1457.

[13] NCI, *Lung Image Database Consortium image collection* [ONLINE].Available: https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI

[14] *Lung Nodule Analysis 2016 dataset* [ONLINE]. Available: https://luna16.grand-challenge.org/data/