Deep Learning Techniques to Classify and Analyze Medical Imaging Data

Chucknorris Garikayi Madamombe
Department of Computer Science
Namibia University of Science and Technology, Windhoek, Namibia
chuckygari@gmail.com

***Abstract* - In recent years, deep learning techniques particularly Convolutional Neural Networks (CNNs) have been used in various disciplines. CNNs have shown an essential ability to automatically extract large volumes of information from big data. The use of CNNs have significantly proved to be useful especially in classifying natural images. Nonetheless, there have been a major barrier in implementing the CNNs in medical domain due to lack of proper training data. As a result, general imaging benchmarks such as ImageNet have been popularly used in the medical domain even though they are not so perfect as compared to the CNNs. In this paper, a comparative analysis of LeNet, AlexNet and GoogLeNet have been done. Thereafter, the paper has proposed an improved conceptual framework for classifying medicinal anatomy images using CNNs. Based on the proposed design of the framework, the CNNs architecture is expected to outperform the previous three architectures in classifying medical images.**

Keywords - Convolutional Neural Networks (CNNs), ImageNet, LeNet, AlexNet and GoogLeNet.

I. INTRODUCTION

In the United States of America (USA), millions of medical images are taken every day from both the private and state hospitals [1]. The same applies to all other countries throughout the world. There is a lot of images that are taken in hospitals on a daily basis [1]. As a result, there is a considerable increase in pressure on the healthcare providers for them to provide accurate and efficient diagnostic services. According to [2], the National Institute of Medicine has estimated that approximately more that 15 million patients in America are wrongly diagnosed on a yearly basis. This is quite a huge number of people who are being misdiagnosed and it is a major problem if this continues. This problem is emanating from the fact that there is large volumes of imaging data produced on a daily basis in the hospitals and there is no proper systems in place that can be used to accurately and efficiently handle such kind of data [1]. As a result, there is need for more precise and resourceful decision support systems for doctors so that they can significantly reduce the number of patients who are misdiagnosed [3].

According to [4], deep learning is a technology which is inspired from the working of the human brains. The networks of artificial neurons have the ability to examine enormous amounts of data in order to automatically discern essential patterns without the need for humans to be available [5]. In other words, artificial neural networks work in a similar way as human do. However, since they are systems, the artificial neural networks perform better as compared to human beings. In most cases, deep learning is mainly used where there is need of classifying patterns in amorphous data [6]. Thus deep learning is mostly used to identify unstructured patterns in various media such as text, sound, video and medical images [7].

When the doctors are examining the patients, they try always to solve sophisticated problems. In most cases, the goal of many medical doctors is to give the correct treatment to the patients based on the available previous medical reports like the lab test reports, signs and symptoms of the patients, medical images as well as the patient's medical history [8]. A study conducted by [9] indicated that the digital healthcare data is estimated to grow sharply from 500 petabytes to 25,000 petabytes by 2020 throughout the whole world. As noted earlier on, it is a big challenge for medical doctors to get precise understandings from billions of clinical data. As a result, a large number of researchers, medical professionals and data scientists are continuously finding solutions to advance patient care in the clinics and hospitals.

In this modern day of improved technology, there is need to implement deep learning in the medical industry. According to [7], machine learning algorithms are able to do information processing and pattern recognition and identification in a better way as compared to human beings. In addition, machine learning algorithms can be used to comprehend risk factors for diseases in a very large population. Furthermore, the machine learning algorithms can also be used to identify and predict dangerous diseases such as cancer, diabetes etc.

According to [8], the use of computer assisted diagnosis (CAD) to assess scans of women can detect approximately 52% of the cancer before the women were diagnosed officially.

Reference [10] noted that machine learning algorithms can be used in various disciplines of medicine including the discovery drugs, decision making in clinical. In addition, the use of machine learning algorithms can change by a huge margin the way in which medicine is practiced to date. The power of machine learning algorithms in recent years has come at a time when the medical records are being digitized. Unlike in the past when medical records were mainly paper based, these days, most medical records are being stored electronically. Machine learning algorithms cannot work with paper based medical records. They can only work if the medical records are digitized. This means that these machine learning algorithms has come at the right time when the medical records are now being digitized.

According to [10], the use of electronic health records (EHR) in recent years has increased sharply from approximately 12% to 40% in the USA from 2007 to 2012. Despite the fact that medical images are an important component of any patient's EHR, they are currently being analyzed manually by human radiologists [11]. Human beings cannot be compared to machines because they are slower, they get tired and they might not have much experience. All these are the major limitations of using humans as compared to machine learning algorithms. According to [11], a diagnosis which is delayed and incorrect can be fatal to patients. As a result, it is crucial to automate medical image analysis through the use of precise and effective machine learning algorithms.

On a daily basis, there is an increase in the number of medical images for instance CT, MRI and X-ray. These type of medical images which are increasing on a daily basis are crucial because they provide vital information in order for doctors to provide accurate diagnosis, medical treatments, education as well as providing medical research [12]. In general, the usual methods used to retrieve medical images rely on the annotation of keywords. However, relying on images annotation is not efficient because the process takes a lot of time and also it is difficult to describe the contents of these images with words [13].

In recent years, the Content Based Image Retrieval (CBIR) has considerably gained popularity in the applications of medical image retrieval and classification due to the improvement of computing power and the abrupt developments in science and technology as well as broadband internet. In addition, the CBIR has also been applied in medical applications in recent years [15, 16].

Generally, the dawn of deep learning techniques have conveyed a complete change especially in the field of machine learning. The deep learning techniques and the machine learning algorithms have showed better performances as compared to the use of traditional approaches.

Other than its use in the classification, the other powerful aspect of deep learning is that they do not require any deigning features. Using deep learning, these features can be learned from the raw data which is not processed yet.

According to [16], deep learning is a rapidly emergent technology which is mainly focused on data analysis. Deep learning has been regarded as one of the top technologies in 2013 in analysing medical images. Generally, deep learning can be regarded as an enhancement of artificial neural networks. Deep learning differs in artificial neural networks in that it consists of increased layers that allow higher levels of abstraction and enhanced estimates from the data [16]. To date, it is emerging as the leading machine-learning tool in the general imaging and computer vision domains. In particular, convolutional neural networks (CNNs) have proven to be powerful tools for a broad range of computer vision tasks. Deep CNNs automatically learn mid-level and high-level abstractions obtained from raw data (e.g., images). Recent results indicate that the generic descriptors extracted from CNNs are extremely effective in object recognition and localization in natural images. Medical image analysis groups across the world are quickly entering the field and applying CNNs and other deep learning methodologies to a wide variety of applications. Promising results are emerging.

The whole paper has been outlined as follows; the review of related works have been outlined in Section II. Section III has outlined in detail the proposed methodology. The conclusion of the paper has been outlined in Section IV.

A. Types of Medical Imaging

There is a myriad of imaging modalities, and the frequency of their use is increasing. Reference [2] looked at imaging use from 1996 to 2010 across six large integrated healthcare systems in the United States, involving 30.9 million imaging examinations. The authors found that over the study period, CT, MRI and PET usage increased 7.8%, 10% and 57% respectively. Modalities of digital medical images include ultrasound (US), X-ray, computed tomography (CT) scans and magnetic- resonance imaging (MRI) scans, positron emission tomography (PET) scans, retinal photography, histology slides, and dermoscopy images. Some of these modalities examine multiple organs (such as CT, MRI) while others are organ specific (retinal photography, dermoscopy). The amount of data generated from each study also varies. A histology slide is an image of a few megabytes while a single MRI may be a few hundred megabytes. This has technical implications on the way the data is pre-processed, and on the design of an algorithm's architecture, in the context of processor and memory limitations.

II. RELATED WORKS

Over the past decades, a number of low level feature descriptors have been proposed as an image representation ranging from global features, such as shape and texture features as reported in [17] for classification of pulmonary nodules in lung ct images, edge features [18] to the recently used local feature representations, i.e SIFT with Bag of Visual Words [19]. On the other hand deep learning have shown promising results in image classification. Deep learning alludes to a category of machine learning techniques, where numerous layers of information processing stages in hierarchical architectures are exploited for pattern classification and feature learning.

Reference [20] adopted the deep supervised back-propagation Convolution Neural Network (CNN) for digit recognition successfully. After that, the deep Convolutional Neural Networks (CNNs) proposed in [21] turned out to be a breakthrough, that was declared first in the image classification task of ILSVRC-2012. The model was trained on more than one million images, and has achieved a successful top-5 test error rate of 15.3% over 1000 classes. Since then, more work have been done by improving CNN models to improve the image classification results. Specifically, the CNN model consists of many convolutional layers and pooling layers that are stacked up with one on top of another. The convolutional layer shares several weights, and the pooling layer sub-samples the output of the convolutional layer and reduces the data rate from the layer below. The weight sharing in the convolutional layer, in conjunction with suitable chosen pooling schemes, subsidizes the CNN with some invariance properties e.g. invariance to translation.

On the other hand, CNNs have made a sound advancements in biomedical applications [18] too. Recent work has shown how the implementation of CNNs can significantly improve the performance of the state-of-the-art computer aided detection systems (CADe) [19]. However, in terms of research for classifying anatomies in medical images, there are only a few studies have been carried out using CNN [19]. One of the drawbacks of these studies is that they do not provide extensive evaluation of milestone deep nets [20] and are just focused on single modality, such as only CT images were used in [21]. In order to overcome these limitation, an architecture that can be generalized to various anatomies with different modalities is needed which leads to the main focus of this study.

III. PROPOSED METHODOLOGY

The anatomical classification problem is an important step in Computer Aided and Diagnosis Systems (CADs) [23]. Anatomical structures vary dramatically between individuals i.e normal lung structure as compared to deformed shaped due to pathological intervention, also small lumbar spine bone structure in one individual and same bone structure in other individuals appear to be elongated due to the advancement in the diseases. As a result, a robust Convolutional Neural Network (CNN) architecture is required to achieve better accuracy and that should generalize to all medical image types regardless of normal or abnormal.

Our proposed model of the CNN architecture is a modification of the basic architecture of AlexNet [15]. This architecture contains four convolutional layers (conv) followed by two fully connected layers (fc). The first convolutional layer i.e conv1 subjected to local response normalization, with kernel size 11, which depicts that each unit in each feature map is connected to 11 X11 neighborhood in the input and stride of 4, which means after every four pixels perform the convolution on the input images. The output of the first convolution layer are 96 feature maps. The first layer i.e. conv1 layer is followed by pooling. The kernel size for the pooling is set to 3 with stride 2. Pooling is followed by convolution conv2 with kernel size 5 and stride 2. The pooled feature maps are again convolved in layer conv3, with parameter setting of kernel size equal to 3, stride of 2.

These convolved features are again convolved in layer conv4 with parameter setting same as in layer conv3. Which is followed by fully connected layers (fc), i.e. fc5, fc6. In the layer fc6 in Alexnet two operations are applied, i.e. relu6 and drop6. While as in our proposed architecture, fully connected layer 5 (fc5) is only subjected to rectified linear unit operation. The output of our con4 layer are 256 where as in AlexNet 384 feature maps are generated. The layer fc5 is followed by fully connected layer while fc6 which results in 4096 dimensional vector for each image. The architecture of the proposed CNN used for medical image anatomy classification is as shown in Figure 1 below while the hyper parameter specifications of the proposed CNN framework are given in Table I below.

Table 1: Hyper parameter Specifications of the proposed CNN framework in units.

HyperParameter	Layer1	Layer 2	Layer 3	Layer 4
Number of filters	96	256	384	256
kernel size	11×11	5×5	3×3	3×3
stride	4	2	2	2
Learning rate	0.01			
Momentum	0.9			
Weight Decay	0.0005			
Training epochs	30-60			
Number of units in fully connected layer				4096

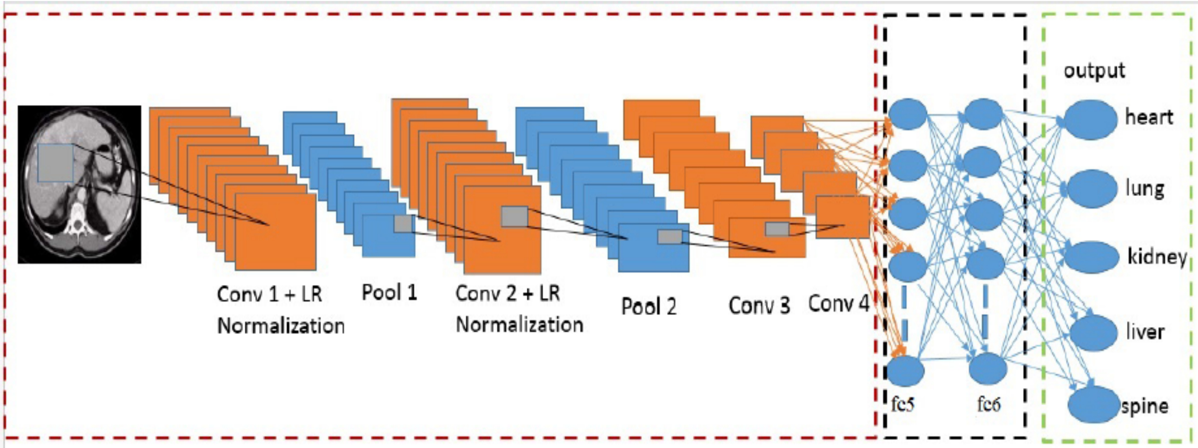

Figure 1: Proposed CNN architecture

In AlexNet [15], five convolutional and three fully connected layers were used, whereas our architecture contains only four layers with two fully connected layers (fc): fc5 and fc6. We did not use the dropout layers that have been used with fc6 and fc7 layers in AlexNet, because looking at the visualization of the feature maps most of the activations are dumped out in higher layers. The result of which it does not control any overfitting but rather adds complexity to the network. Outputs from the convolution layer 4 are calculated as:

$$Y_{i,j}^{l} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} W_{ab} * X_{(i+a)(j+b)}^{l-1} \qquad (1)$$

The features maps resulted from convolution are subjected to rectified linear unit operation as follows:

$$y_{ij} = max\left\{0, Y_{ij}\right\} \qquad (2)$$

In AlexNet, layers fc6 and fc7 are subjected to dropout for regularization. Dropout prevents co-adaptation of hidden units by randomly dropping out i.e., setting to zero a proportion p of the hidden units during forward back-propagation. The dropout is formulated as :

$$y = w(l \odot r) + b, \qquad (3)$$

Gradients are backpropagated only through the unmasked units. So if the Drop out masks the maximum unit it will cause the weights to update in such a way that the neuron will never activate on any data point again. If this happens, then the gradient flowing through the unit will forever be zero from that point on. So these activated units will ultimately vanished during training process. In our modified architecture, this does not subject to fully connected layer, fc5 to drop out operation, rather feed it with the output of the conv4 layer, as shown in a simplified expression,

$$y = w.l + b \tag{4}$$

IV. CONCLUSION

In this paper, we proposed a modified CNN architecture that combines multiple convolution and pooling layers for higher level feature learning. The experiments for medical image anatomy classification has been carried out and it shows that the proposed CNN feature representation outperforms the three baseline architectures for classifying medical image anatomies. The modification of CNN has been done on the basis of experimentation that is carried out with the three milestone architectures. These models over fit due to the number of layers and the hyper-parameters used in these architectures have been used for large set of natural images. However, medical image datasets are different in terms of their acquisition medium and less availability because of privacy and security policies as compared to natural images. In this paper, we also provide an insight into the deep features that have been learned through training that will help in analyzing various abstraction of features ranging from low level to high level and their role in final classification. Our future work will extend to recognition and classification of pathological structures from these classified anatomies, leading to a fully automated medical image classification system.

REFERENCES

[1] Litjens, G., T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. V. Ginneken & S´anchez. A survey on deep learning in medical image analysis, Medical Image Analysis, vol. 42, 60–88, 2017.
[2] Zhao, F., & X. Xie. An overview of interactive medical image segmentation. Annals of the BMVA, 7, 1–22, 2012.
[3] Rajchl, M., M. Lee, O. Oktay, K. Kamnitsas, J. Passerat-Palmbach, W. Bai, M. Rutherford, J. Hajnal, B. Kainz, & D. Rueckert. DeepCut: Object segmentation from bounding box annotations using convolutional neural networks, TMI, vol. 36, no. 2, 674–683, 2017.
[4] Li, W., G. Wang, L. Fidon, S. Ourselin, M. J. Cardoso, & T. Vercauteren. On the compactness, efficiency, and representation of 3D convolutional networks: Brain percolations as a pretext task, in IPMI, 348–360, 2017.
[5] Gal, Y. & Z. Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning, in ICML, 1050–1059, 2016.
[6] Kamnitsas, K., C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, & B. Glocker. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Medical Image Analysis, 36, 61–78, 2017.
[7] Wang, G., Li, W., Maria, A., Zuluaga, Pratt, R, Premal, A., Patel, Aertsen, M., Doel, T., Anna. L. Jan., D., Ourselin, S. & Vercauteren, T. Interactive medical image segmentation using deep learning with image-specific fine-tuning. IEEE Transactions on Medical Imaging, 1-12, 2018.
[8] Singh, S. & Singh, N. Object classification to analyze medical imaging data using deep learning. International Conference on Innovations in information Embedded and Communication Systems (ICIIECS), 1 – 4, 2018.
[9] Yigzaw, K. Y. & Bellika, J. G. Evaluation of secure multi-party computation for reuse of distributed electronic health data, IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), 219-222, 2014.
[10] Ker, J., Wang, L., Rao, J., & Lim, T. Deep learning applications in medical image analysis. IEEE Access, 6, 9375 – 9389, 2017.
[11] Hsiao, C.J., Hing, E., & Ashman, J. Trends in electronic health record system use among office-based physicians: United States, 2007 – 2012. Nat. Health Stat. Rep, 75, 1-18, 2014.
[12] Müller, H., Michoux, N., Bandon, D., & Geissbuhler, A. A review of content based image retrieval systems in medical applications–clinical benefits and future directions. International Journal of Medical Informatics, 73, 1-23, 2004.
[13] Qiu, C., Cai, Y., Gao, X., & Cui, Y. Medical image retrieval based on the deep convolution network and hash coding. International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, 1-6, 2017.
[14] Muller, H., Rosset, A., Vallee, J.P., & Geisbuhler, A. Comparing feature sets for content-based image retrieval in a medical case database. SPIE Med. Image, PACS Image. 99–109, 2004.
[15] Felipe, J. C., Traina, A. J. M., & Traina, C. Retrieval by content of medical images using texture for tissue identification. IEEE Symp. Computer-Based Med. 175–180, 2003.
[16] LeCun, Y., Bengio, Y., & Hinton, G. Deep learning. Nature, vol. 521, no. 7553, pp. 436–444, 2017.
[17] Ashis Kumar Dhara, Sudipta Mukhopadhyay, Anirvan Dutta, Mandeep Garg, and Niranjan Khandelwal. A combination of shape and texture features for classification of pulmonary nodules in lung ct images. Journal of digital imaging, 29(4):466–475, 2016.
[18] Mohammad Reza Zare, Ahmed Mueen, and Woo Chaw Seng. Automatic medical x-ray image classification using annotation. Journal of digital imaging, 27(1):77–89, 2014.
[19] Wei Yang, Zhentai Lu, Mei Yu, Meiyan Huang, Qianjin Feng, and Wufan Chen. Content-based retrieval of focal liver lesions using bag-of-visual-words representations of single-and multiphase contrast-enhanced ct images. Journal of digital imaging, 25(6):708–719, 2012.
[20] Yann LeCun, L´eon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998.
[21] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.