Big Data Applications in Using Neural Networks for Medical Image Analysis

Tyler Peterson

Indiana University - School of Informatics, Computing, and Engineering

711 N. Park Avenue

Bloomington, Indiana 47408

typeter@iu.edu

ABSTRACT

Medical image analysis is proving to be a promising domain for disruption by machine learning. The analysis of medical imagery has long been within the purview of radiologists, a specialization in medicine that reviews medical imaging to form diagnoses and advise on treatment options. Historically, radiologists have relied on their training, senses and years of experience to evaluate images for medical issues, such as the presence of tumors, lung nodules, and hip osteoarthritis. The presence of technology, generally referred to as computer-aided diagnosis (CAD) tools, has been growing over the last several decades, but modern computing power and sizable datasets have accelerated the effectiveness of these tools. Machine learning algorithms, especially artifical neural networks (ANN), are being leveraged to help identify abnormalities present in medical images at a high level of accuracy. Several research studies conclude that ANN techniques can match, and occassionally outperform, the abilities of radiologists. Big data and the application of advanced algorithms show promise for evolving our ability to successfully evaluate medical images and save lives in the process.

KEYWORDS

i523, hid331, Big Data, Medical Image Analysis, Artificial Neural Networks, Medicine

1 INTRODUCTION

The analysis of medical imagery is primarily the responsibility of radiologists. These individuals are medical doctors who specialize in diagnosing diseases through review of images produced by various imaging modalities, such as x-ray, ultrasound, computerized tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET) [6]. Radiologists serve as an expert to other physicians by analyzing the medical images of patients suspected of having certain medical issues, and by making recommendations on subsequent care based on the observations [6]. The images reviewed by radiologists are generally stored digitially, and images are increasingly being stored in picture archiving and communications systems (PACS). These systems are required to keep up with the rapid accumulation of medical image data. Between 2005 and 2011, the medical image data in US hospitals increased from only 8,900 terabytes to 27,000 terabytes [7]. That number is expected to grow by 20 percent every year due to increasing image size and resolution, the adoption of 3D imaging, and an aging population who will likely bring an increasing demand for medical imaging studies [7].

It is estimated that one billion medical images are created worldwide each year, and most of these are assessed by radiologists [1]. Given that radiologists are human, their judgment is fallible. It is estimated that the lowest average error rate in analyzing medical imagery is 4 percent, which means collectively radiologists are estimated to make 40 million errors in judgement every year [1]. A particularly striking example of fallibility comes from a study that analyzed the first and second interpretations of radiologists from Massachusetts General Hospital. They reviewed abdominal CTs and re-reviewed studies that had either been interpreted by themselves or a colleague. The study found that the radiologists disagreed with their peers 30 percent of the time and even disagreed with themselves 25 percent of the time [1].

There are two major types of radiologic analysis error: perceptual error and interpretive errors [1]. Most errors, up to 80 percent, are perceptual errors, which occur when an abnormality is not perceived by the reviewer during the initial review, but is identified in a subsequent analysis [1]. Interpretive errors occur when the radiologist successfully identifies the abnormality, but incorrectly diagnosis the problem, which may lead to a less appropriate course of careaction [1]. There are several reasons why errors occur, including fatigue, excessive pace of analysis, distractions and insufficient knowledge of the practitioner. It is also asserted that the extreme complexity of a radiologist's job contributes to the errors. Errors occur in the practice of radiologists all across the world, at varying levels of training, in all imaging modalities and all clinical settings [1].

Over the last several decades, there has been an effort to develop computer-based tools to aid radiologists in the detection of abnormalities. Computer-aided diagnosis (CAD) tools are primarily intended to increase the rate at which problems are identified while also reducing the false negatives resulting from human error [9] . These systems are intended to supplement, not replace, the radiologist by reporting a second opinion to be considered alongside the radiologist's assessment. The earliest initiatives to develop these tools occurred in the 1960s, and concerted efforts began in the 1980s [2]. Despite the research being nearly 60 years old, widespread adoption is a relatively recent occurrence [4]. Clinical

studies reported early CAD implementations as being minimally effective. Specifically, CAD decisions included more false positives than human assessments, which created more work for radiologists and often led to additional, unnecessary medical tests and biopsies [5].

Several improvements in the field of computing have increased the accuracy of CAD tools and subsequently encouraged wider adoption of these tools into clinical workflows. The advancements includes increased access to digital imaging datasets, larger imaging data sets, increased used of imaging in healthcare and increased computing power [4][5]. These factors combine to create an ideal state for research related to advanced machine learning techniques, namely artifical neural networks, and the implementation of tools that can rival, and even outperform, the assessment of highly trained radiologists.

2 MACHINE LEARNING AND ARTIFICIAL NEURAL NETWORKS

Broadly speaking, machine learning is a way of applying artificial intelligence to a problem through the analysis of data. Machine learning techniques assess the features, or attributes, of samples in a dataset to identify patterns in the data, and the resulting algorithm can be used to render conclusions about new inputs without human intervention [5]. Medical image analysis presents what is referred to as a classification problem. The typical example of a classifcation problem is handwritten numerical digit recognition. In this example, a handwritten digit between 0 and 9 is fed into the algorithm, and the algorithm decides which of the ten digits, or classes, that handwritten digit is most likely to belong. Specific to medical image analysis, the classifier detects abnormalities in images otherwise not present in images of the same area in healthy individuals, and renders a conclusion as to what that abnormality is.

There are several different machine learning techniques that have traditionally been used in classification problems, such as support vector machines (SVM). SVMs are an example of a supervised machine learning model, meaning this method uses labeled data. Every sample in the dataset includes the label, or correct answer, along with a value for each of the attributes. The SVM identifies patterns in the labeled samples to create an algorithm, and new, unlabeled samples can be processed by the algorithm and given a prediction. In the task of medical image analysis, attributes have historically been identified and designed by human experts [9]. For example, an expert would identify abnormalities by explicitly describing shape, texture, position and orientation of the abnormal biological structure [8]. In addition to this being labor intensive, the image features are specific to the immediate problem being explored and cannot be expected to work well for other image types [9].

Artifical neural networks (ANN) are another class of machine learning techniques that are increasingly being leveraged to tackle problems related to image analysis. ANNs, just like SVMs, are often utilized in a supervised manner, but do not require the painstaking process of expert-defined key attributes. Instead, ANNs learn the important features from the images themselves [8]. This is an obvious advantage when compared to traditional machine learning methods, but the complexity of these algorithms has prevented the achievment of mainstream use until recently. The theory of ANNs was introduced in the 1950s, and research in this field has begun to flourish recently due to the increase in computing power and availability of high quality datasets [5].

3 HOW ANNS WORK

The inspiration for the design of ANNs is drawn from the way a brain works.

deep neural network, a convolutional neural network and a auto encoder

AUC definition

Several statistics can be used for evaluating the accuracy of ANNs. Sensitivity is the true positive detection rate. This is the percentage of positive occurrences that are successfully identified [3]. A false positive may lead to unnecessary testing, unecessary expenses and unecessary stress on the patient who has been led to believe they have a certain condition. Specificity is the true negative detection. This is the percentage of negative occurrences that are successfully identified [3]. A false negative may lead a missed diagnosis, resulting in delayed treatment and a potentially avoidable death in some cases. Sensitivity and specificity can be assessed together by the receiver operating characteristic (ROC). The ROC curve plots the true positive rate against the true positive rate (100 - the true negative rate) for varying decision thresholds. This illustrates the trade-off between sensitivity and specificity and can provide guidance on which decision threshold is appropirate for the task [3]. ROC curves are often leveraged to evaluate the performance of ANNs by calculating the area under the ROC curve, also known as the AUC. The goal is the maximize the AUC value, and that value points to the optimal balance between sensitivity and specificity [3].

4 APPLICATIONS IN MEDICAL IMAGE ANALYSIS

There are several specialities in which medical image analysis has been studied and applied for the purpose of computer-aided detection and diagnosis, and the effectiveness of ANNs has been formally studied in different ways. This includes ANNs versus medical professional, ANNs combined with medical professional versus ANNs and medical professionals separately, and one type of ANN versus another type of ANN.

One study applied deep learning techniques to the evaluation of breast sentinel lymph nodes for the presence of metastasis. A positive test likely means the staging of the breast cancer will be higher and subsequent treatment will be more aggressive. A false negative means a patient's disease will be thought of as less advanced that it actually is, and subsequent treatment will not be as aggressive as necessary. The medical professionals in this study obtained a receiver operating curve (AUC) of 0.966, and the algorithm received

an AUC of 0.925. The AUC for the combination of the human and computerized assessment was 0.995, which equals an 85 percent reduction in error for the medical professional [11].

A second study compare the performance of three different types of ANNs on the detection of cancerous lung nodules in CT images. Lung cancer is a disease that greatly benefits from early detection. Over 220,000 new cases were identified in 2015, and an early detection of the disease improves the 5 year survival rate by roughly 50 percent [10]. CT images provide three-dimensional (3D) views of the chest and are a key component of the clinical workflow of this specialty. These image are analyzed to understand if the structures in the image are part of the expected anatomy or if a nodule is a tumor. If a nodule is present, the goal is to understand if the nodule is benign or malignant. This determination closely depends on the size, shape and texture of the nodule, all of which can be analyzed by an ANN. This study compared the performance of a deep neural network (DNN), a convolutional neural network (CNN) and a auto encoder (SAE). The CNN performed best with an AUC of 0.916, while the DNN and SAE recorded AUCs of 0.877 and 0.884 respectively [10].

A third study compares the performance of an ANN to two experienced physicians in the evaluation of x-rays for the presence of hip osteoarthritis. Hip osteoarthritis causes pain and stiffness which can diminish quality of life through an inability to perform daily tasks or go to work [13]. Hip osteoarthritis is diagnosed through x-ray imaging studies, which are traditionally evaluated by radiologists. The time consuming and error-prone nature of this work is well documented, and with an increasingly aging population, efficient and timely diagnosis of hip osteoarthritis is growing in importance. The study used a CNN to achieve a sensitivity rate of 95.0 percent and a specificity rate of 90.7 percent, with an AUC of 0.94. Both physicians achieved a sensitivity of 100 percent and specificity rates of 86.0 percent and 93.0 percent. While the CNN recorded slightly lower sensitivty compared to the two physicians, it did record a higher specificity than one of the physicians [13]. This shows promise for the performance of ANNs in evaluating hip osteoarthritis.

Diabetic retinopathy is the third leading cause of blindness worldwide, contributing to

5   INFRASTRUCTURE
6   CHALLENGES

There are several challenges that may inhibit the success of widespread use of CAD tools built upon a neural network algorithm. First, overfitting occurs when an algorithm is trained based on a dataset that does not generalize well to examples outside of the data used to train the algorithm. Many studies demonstrating the value of ANNs were trained using relatively small datasets, and because the significant features present in a small dataset may not be the exact same features present in a large dataset, the algorithm derived from the small dataset may not perform well when analyzing images from the large dataset [12] [5]. This issue can be addressed by training algorithms on larger datasets, but of course requires access to larger datasets, longer training periods and more computing power.

Second, algorithms derived from ANNs are frequently considered to be black boxes, meaning that the results are uninterpretable. This contrasts with several other types of machine learning techniques that produce equations that much more easily lend themselves to understanding which features are significant and why [5]. Instill belief and trust in a sytem that is difficult, if not impossible, to explain is a barrier, even if that system produces accurate responses the vast majority of time.

Third, ethical and legal considerations must be made. Adopters of this technology must grapple with questions pertaining to the liability if and when the system makes a prediction that harms a patient [5]. If a radiologist is led to a conclusion by an algorithm, and the algorithm presents a false positive, a false negative or presents one conclusion while missing another, who is responsible for the error?

Fourth, the ANNs are dependent on the quality and nature of the imaging data used to train algorithms. There is variability around the world in regards to the type and quality of imaging machines and the imaging protocols that dictate why and how images are taken [5]. It's conceivable that two different imaging machines or two different techniques capturing two images of similar abnormalities could appear different to an algorith such that the conclusion is not the same. This issue could at least partially be addressed by sufficiently large datasets containing labeled images of abnormalities that were taken using machines of varying quality and executed using differing methods.

Fifth, computing power

7   CONCLUSION

math used - sigmoid function - why what does this do

back propagation

what is deep learning, how does it compare with human brains

ACKNOWLEDGMENTS

REFERENCES

[1] Michael Bruno, Eric Walker, and Hani Abujudeh. 2015. Understanding and Confronting Our Mistakes: The Epidemiology of Error in Radiology and Strategies for Error Reduction. *Radio-Graphics* 35, 6 (2015), 1668–1676. https://doi.org/10.1148/rg.2015150023

[2] Kunio Doi. 2007. Computer-Aided Diagnosis in Medical Imaging: Historical Review, Current Status and Future Potential. *Comput Med Imaging Graph* 31, 4-5 (2007), 198–211.

[3] Christopher M Florkowski. 2008. Sensitivity, Specificity, Receiver-Operating Characteristic (ROC) Curves and Likelihood Ratios: Communicating the Performance of Diagnostic Tests. *Clinical Biochemistry Review* 29 (August 2008), S83–S87.

[4] Tae-Yun Kim, Jaebum Son, and Kwang-Gi Kim. 2011. The Recent Progress in Quantitative Medical Image Analysis for Computer Aided Diagnosis Systems. *Healthcare Informatics Research* 17, 3 (September 2011), 143–149. https://doi.org/10.4258/hir.2011.17.3.143

[5] June-Goo Lee, Sanghoon Jun, Young-Won Cho, Hyunna Lee, Guk Bae Kim, Joon Beom Seo, and Namkug Kim. 2017. Deep Learning in Medical Imaging: General Overview. *Korean Journal of Radiology* 18, 4 (2017), 570–584. https://doi.org/10.3348/kjr.2017.18.4.570

[6] Radiological Society of North America (RSNA). 2017. What Does a Radiologist Do? Online. (April 2017). https://www.radiologyinfo.org/en/info.cfm?pg=article-your-radiologist

[7] Morris Panner. 2015. Whats Next for the Health-Care Data Center? Online. (April 2015). http://www.datacenterjournal.com/whats-healthcare-data-center/

[8] Faizan Shaikh. 2017. Deep Learning vs. Machine Learning the essential differences you need to know! Online. (April 2017). https://www.analyticsvidhya.com/blog/2017/04/comparison-between-deep-learning-machine-learning/

[9] Dinggang Shen, Guorong Wu, and Heung-Il Suk. 2017. Deep Learning in Medical Image Analysis. *Annual Review of Biomedical Engineering* 19 (June 2017), 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442

[10] QingZeng Song, Lei Zhao, and XingKe Luo andXueChen Dou. 2017. Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images. *Journal of Healthcare Engineering* 2017 (August 2017), 1–7.

[11] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H Beck. 2016. Deep Learning for Identifying Metastatic Breast Cancer. Online. (June 2016).

[12] Shijun Wang and Ronald M. Summers. 2012. Machine Learning and Radiology. *Medical Image Analysis* 16, 5 (July 2012), 933–951. https://doi.org/10.1016/j.media.2012.02.005

[13] Yanping Xue, Rongguo Zhang, Yufeng Deng2, Kuan Chen, and Tao Jiang. 2017. A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis. *PLoS ONE* 12, 6 (June 2017), 1–10. https://doi.org/10.1371/journal.pone.0178992

A ISSUES

DONE:

> Example of done item: Once you fix an item, change TODO to DONE

A.1 Assignment Submission Issues

> Do not make changes to your paper during grading, when your repository should be frozen.

A.2 Uncaught Bibliography Errors

> Missing bibliography file generated by JabRef

> Bibtex labels cannot have any spaces, _ or & in it

> Citations in text showing as [?]: this means either your report.bib is not up-to-date or there is a spelling error in the label of the item you want to cite, either in report.bib or in report.tex

A.3 Formatting

> Incorrect number of keywords or HID and i523 not included in the keywords

> Other formatting issues

A.4 Writing Errors

> Errors in title, e.g. capitalization

> Spelling errors

> Are you using *a* and *the* properly?

> Do not use phrases such as *shown in the Figure below.* Instead, use *as shown in Figure 3*, when referring to the 3rd figure

> Do not use the word *I* instead use *we* even if you are the sole author

> Do not use the phrase *In this paper/report we show* instead use *We show*. It is not important if this is a paper or a report and does not need to be mentioned

> If you want to say *and* do not use *&* but use the word *and*

> Use a space after . , :

> When using a section command, the section title is not written in all-caps as format does this for you

`\section{Introduction}` and NOT `\section{INTRODUCTION}`

A.5 Citation Issues and Plagiarism

> It is your responsibility to make sure no plagiarism occurs. The instructions and resources were given in the class

> Claims made without citations provided

> Need to paraphrase long quotations (whole sentences or longer)

> Need to quote directly cited material

A.6 Character Errors

> Erroneous use of quotation marks, i.e. use "quotes" , instead of " "

> To emphasize a word, use *emphasize* and not "quote"

> When using the characters & # % _ put a backslash before them so that they show up correctly

> Pasting and copying from the Web often results in non-ASCII characters to be used in your text, please remove them and replace accordingly. This is the case for quotes, dashes and all the other special characters.

> If you see a gure and not a figure in text you copied from a text that has the fi combined as a single character

A.7 Structural Issues

> Acknowledgement section missing

> Incorrect README file

> In case of a class and if you do a multi-author paper, you need to add an appendix describing who did what in the paper

> The paper has less than 2 pages of text, i.e. excluding images, tables and figures

The paper has more than 6 pages of text, i.e. excluding images, tables and figures

Do not artificially inate your paper if you are below the page limit

A.8   Details about the Figures and Tables

Capitalization errors in referring to captions, e.g. Figure 1, Table 2

Do use *label* and *ref* to automatically create figure numbers

Wrong placement of figure caption. They should be on the bottom of the figure

Wrong placement of table caption. They should be on the top of the table

Images submitted incorrectly. They should be in native format, e.g. .graffle, .pptx, .png, .jpg

Do not submit eps images. Instead, convert them to PDF

The image files must be in a single directory named "images"

In case there is a powerpoint in the submission, the image must be exported as PDF

Make the figures large enough so we can read the details. If needed make the figure over two columns

Do not worry about the figure placement if they are at a different location than you think. Figures are allowed to oat. For this class, you should place all figures at the end of the report.

In case you copied a figure from another paper you need to ask for copyright permission. In case of a class paper, you must include a reference to the original in the caption

Remove any figure that is not referred to explicitly in the text (As shown in Figure ..)

Do not use textwidth as a parameter for includegraphics

Figures should be reasonably sized and often you just need to add columnwidth

e.g.

```
/includegraphics[width=\columnwidth]{images/myimage.pdf}
```

re

B   BIBTEX ISSUES

C   ISSUES

DONE:

Example of done item: Once you fix an item, change TODO to DONE

C.1   Assignment Submission Issues

Do not make changes to your paper during grading, when your repository should be frozen.

C.2   Uncaught Bibliography Errors

Missing bibliography file generated by JabRef

Bibtex labels cannot have any spaces, _ or & in it

Citations in text showing as [?]: this means either your report.bib is not up-to-date or there is a spelling error in the label of the item you want to cite, either in report.bib or in report.tex

C.3   Formatting

Incorrect number of keywords or HID and i523 not included in the keywords

Other formatting issues

C.4   Writing Errors

Errors in title, e.g. capitalization

Spelling errors

Are you using *a* and *the* properly?

Do not use phrases such as *shown in the Figure below*. Instead, use *as shown in Figure 3*, when referring to the 3rd figure

Do not use the word *I* instead use *we* even if you are the sole author

Do not use the phrase *In this paper/report we show* instead use *We show*. It is not important if this is a paper or a report and does not need to be mentioned

If you want to say *and* do not use *&* but use the word *and*

Use a space after . , :

When using a section command, the section title is not written in all-caps as format does this for you

\section{Introduction} and NOT \section{INTRODUCTION}

C.5   Citation Issues and Plagiarism

It is your responsibility to make sure no plagiarism occurs. The instructions and resources were given in the class

Claims made without citations provided

Need to paraphrase long quotations (whole sentences or longer)

Need to quote directly cited material

C.6 Character Errors

Erroneous use of quotation marks, i.e. use "quotes" , instead of " "

To emphasize a word, use *emphasize* and not "quote"

When using the characters & # % _ put a backslash before them so that they show up correctly

Pasting and copying from the Web often results in non-ASCII characters to be used in your text, please remove them and replace accordingly. This is the case for quotes, dashes and all the other special characters.

If you see a gure and not a figure in text you copied from a text that has the fi combined as a single character

C.7 Structural Issues

Acknowledgement section missing

Incorrect README file

In case of a class and if you do a multi-author paper, you need to add an appendix describing who did what in the paper

The paper has less than 2 pages of text, i.e. excluding images, tables and figures

The paper has more than 6 pages of text, i.e. excluding images, tables and figures

Do not artificially inate your paper if you are below the page limit

C.8 Details about the Figures and Tables

Capitalization errors in referring to captions, e.g. Figure 1, Table 2

Do use *label* and *ref* to automatically create figure numbers

Wrong placement of figure caption. They should be on the bottom of the figure

Wrong placement of table caption. They should be on the top of the table

Images submitted incorrectly. They should be in native format, e.g. .graffle, .pptx, .png, .jpg

Do not submit eps images. Instead, convert them to PDF

The image files must be in a single directory named "images"

In case there is a powerpoint in the submission, the image must be exported as PDF

Make the figures large enough so we can read the details. If needed make the figure over two columns

Do not worry about the figure placement if they are at a different location than you think. Figures are allowed to oat. For this class, you should place all figures at the end of the report.

In case you copied a figure from another paper you need to ask for copyright permission. In case of a class paper, you must include a reference to the original in the caption

Remove any figure that is not referred to explicitly in the text (As shown in Figure ..)

Do not use textwidth as a parameter for includegraphics

Figures should be reasonably sized and often you just need to add columnwidth

e.g.

```
/includegraphics[width=\columnwidth]{images/myimage.pdf}
```

re