UNIVERSITY OF CALICUT
CENTRE FOR COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
THALIKULAM-680569

DEEP LEARNING APPLICATIONS IN MEDICAL IMAGE ANALYSIS

SEMINAR PRESENTED BY

RESMI K.G

REG NO: TTARMCA022

SEMINAR REPORT

2017-2020

UNIVERSITY OF CALICUT
CENTRE FOR COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
THALIKULAM – 680569

CERTIFICATE

*This is to certify that the seminar entitled "**DEEP LEARNING APPLICATIONS IN MEDICAL IMAGE ANALYSIS**" is prepared and presented by **RESMI K.G (TTARMCA022)** of M.C.A Fifth Semester, a student of Calicut university,Center for Computer Science and Information Technology , Thalikulam , as a part of course curriculum.*

Lecture in charge:Associate Coordinator

1.

2.

3.

Date: …………………

ACKNOWLEDGEMENT

I feel great pleasure in submitting my seminar report on ”Deep Learning Applications in Medical Image Analysis” System Based on Residual Number System”. I take this occasion to thank **God Almighty** for blessing me with his grace and taking my endeavour to a successful culmination.

I wish to express true sense of gratitude towards **Mrs.Greeshma P V**, Administrative Officer of CCSIT Thalikulam, **Mrs.Deepthi K V**, AssociateCo-Ordinator and **staff in charge** of CCSIT Thalikulam , for their constant encouragement and valuable guidance and for providing all the necessary facilities, which were indispensable in the completion of this seminar. I would like to thank all staff members who extended their kind support during the accomplishment of the seminar report. Finally, I express my sincere thanks to all those who helped me directly indirectly in many ways in completion of this seminar.

ABSTRACT

The tremendous success of machine learning algorithms at image recognition tasks in recent years intersects with a time of dramatically increased use of electronic medical record sand diagnostic imaging. This review introduces the machine learning algorithms as applied to medical image analysis, focusing on convolutional neural networks, and emphasizing clinical aspects of the field. The advantage of machine learning in an era of medical big data is that significant hierarchal relationships within the data can be discovered algorithmically without laborious hand-crafting of features. We cover key research areas and applications of medical image classification, localization, detection, segmentation, and registration. We conclude by discussing research obstacles, emerging trends, and possible future directions

INDEX

DEEP LEARNING APPLICATIONS IN MEDICAL IMAGE ANALYSIS

Chapter 1

INTRODUCTION

Machine learning algorithms have the potential to be invested deeply in all fields of medicine, from drug discovery to clinical decision making, significantly altering the way medicine is practiced. The success of machine learning algorithms at computer vision tasks in recent years comes at an opportune time when medical records are increasingly digitalized. The use of Electronic Health Records (EHR) quadrupled from 11.8% to 39.6% amongst office based physicians in the US from 2007 to 2012 Medical images are an integral part of a patient's EHR and are currently analyzed by human radiologists, who are limited by speed, fatigue, and experience. It takes years and great financial cost to train a qualified radiologist, and some health-care systems outsource radiology reporting to lower-cost countries such as India via tele-radiology. A delayed or erroneous diagnosis causes harm to the patient. Therefore, it is ideal for medical image analysis to be carried out by an automated, accurate and efficient machine learning algorithm.

Medical image analysis is an active field of research for machine learning, partly because the data is relatively structured and labelled, and it is likely that this will be the area where patients first interact with functioning, practical artificial intelligence systems. This is significant for two reasons. Firstly, in terms of actual patient metrics, medical image analysis is a litmus test as to whether artificial intelligence systems will actually improve patient outcomes and survival. Secondly, it provides a testbed for human-AI interaction, of how respective patients will be towards health-altering choices being made, or assisted by a non-human actor.

1. Types of Medical Imaging

There is a myriad of imaging modalities, and the frequency of their use is increasing. Smith-Bindman et al.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 Ultra Sound(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. Fig.1 shows some example medical images. Some of these modalities examine multiple organs(such as CT,MRI) while others are organ specific(retinalphotography,dermoscopy). The amount of data generated from each study also varies. A histology slide is an image file 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.

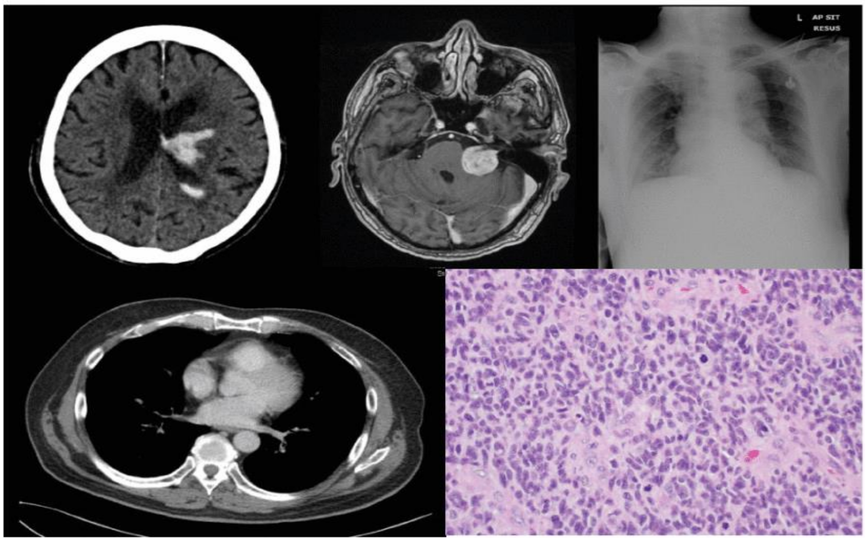


Figure 1.1: A collage of images depicting medical images

1. **History of Medical Image Analysis** The symbolic AI paradigm Of the 1970s led to the development of rule-based, expert systems.One early implementation in medicine was the MYCIN system by Shortliffe [3], which suggested different regimes of antibiotic therapies for patients. Parallel to these developments, AI algorithms moved from heuristics-based techniques to manual, handcrafted feature extraction techniques. and then to supervised learning techniques. Unsupervised machine learning methods are also being researched, but the majority of the algorithms from 2015–2017 in the published literature have employed supervised learning methods, namely Convolutional Neural Networks (CNN). Aside from the availability of large labelled data sets being available, hardware advancements in Graphical Processing Units (GPUs) have also led to improvements in CNN performance ,and their widespread use in medical image analysis.

2. **Convolutional Neural Networks** Both the 2-dimensional and 3-dimensional structures of an organ being studied are crucial in order to identify what is normal versus abnormal. By maintaining these localspatial relationships, CNNs are well-suited to perform image recognitiontasks. CNNs have been put to work in many ways, including image classification, localization, detection, segmentation and registration. CNNs are the most popular machine learning algorithm in image recognition and visual learning tasks, due to its unique characteristic of preserving local image relations, while performing dimensionality reduction. This captures important feature relationships in an image (such as how pixels on an edge join to form a line), and reduces the number of parameters the algorithm has to compute, increasing computational efficiency. CNNs are able to take as inputs and process both 2-

dimensional images , as well as 3-dimensional images with minor modifications. This is a useful advantage in designing a system for hospital use, as some modalities like X-rays are 2-dimensional while others like CT or MRI scans are 3dimensional volumes.

Chapter2

Machine Learning Architectures

Convolutional Neural Networks

Currently, CNNs are the most researched machine learning algorithms in medical image analysis. The reason for this is that CNNs preserve spatial relationships when filtering input images. As mentioned, spatial relationships are of crucial importance in radiology, for example, in how the edge of a bone joins with muscle, or where normal lung tissue interfaces with cancerous tissue. As shown in Fig. 2. a CNN takes an input image of raw pixels, and transforms it via Convolutional Layers, Rectified Linear Unit (RELU) Layers and Pooling Layers. This feeds into a final Fully Connected Layer which assigns class scores or probabilities, thus classifying the input into the class with the highest probability.

1. *Convolutional layer*

A convolution is defined as an operation on two functions. In image analysis, one function consists of input values (e.g. pixel values) at a position in the image, and the second function is a filter (or kernel); each can be represented as array of numbers. Computing the dot product between the two functions gives an output. The filter is then shifted to the next position in the image as defined by the stride length.

The computation is repeated until the entire image is covered, producing a feature (or activation) map. This is a map of where the filter is strongly activated and 'sees' a feature such as a straight line, a dot, or a curved edge. If a photograph of a face was fed into a CNN, initially low-level features such as lines and edges are discovered by the filters. These build up to progressively higher features in subsequent layers, such as a nose, eye or ear, as the feature maps become inputs for the next layer in the CNN architecture.

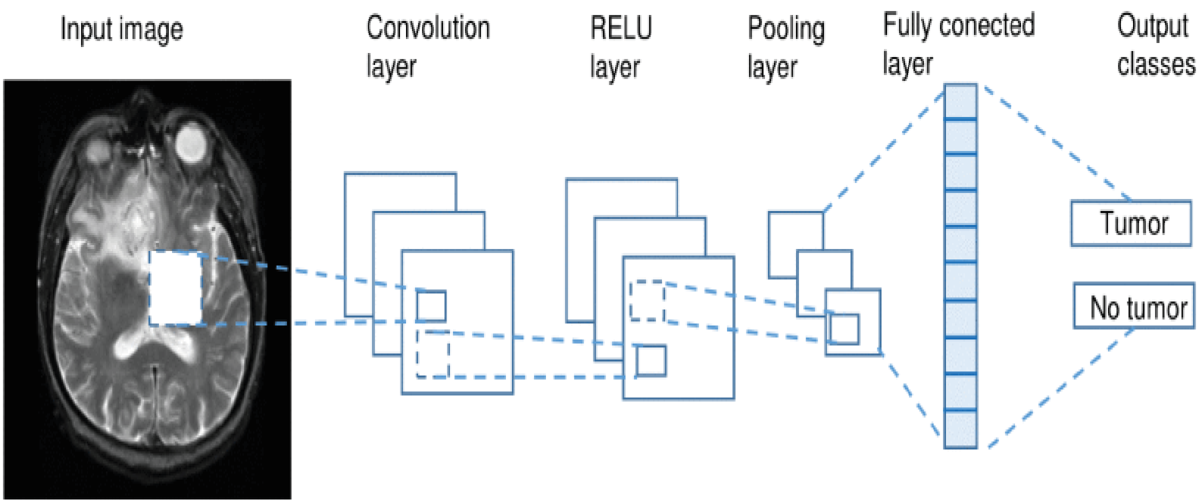


Figure2.1: disease classification task, an input image of an abnormal axial slice of a T2-weighted MRI brain is run through a schematic depiction of a CNN. Feature extraction of the input image is performed via the Convolution, RELU and pooling layers, before classification by the fully connected layer.

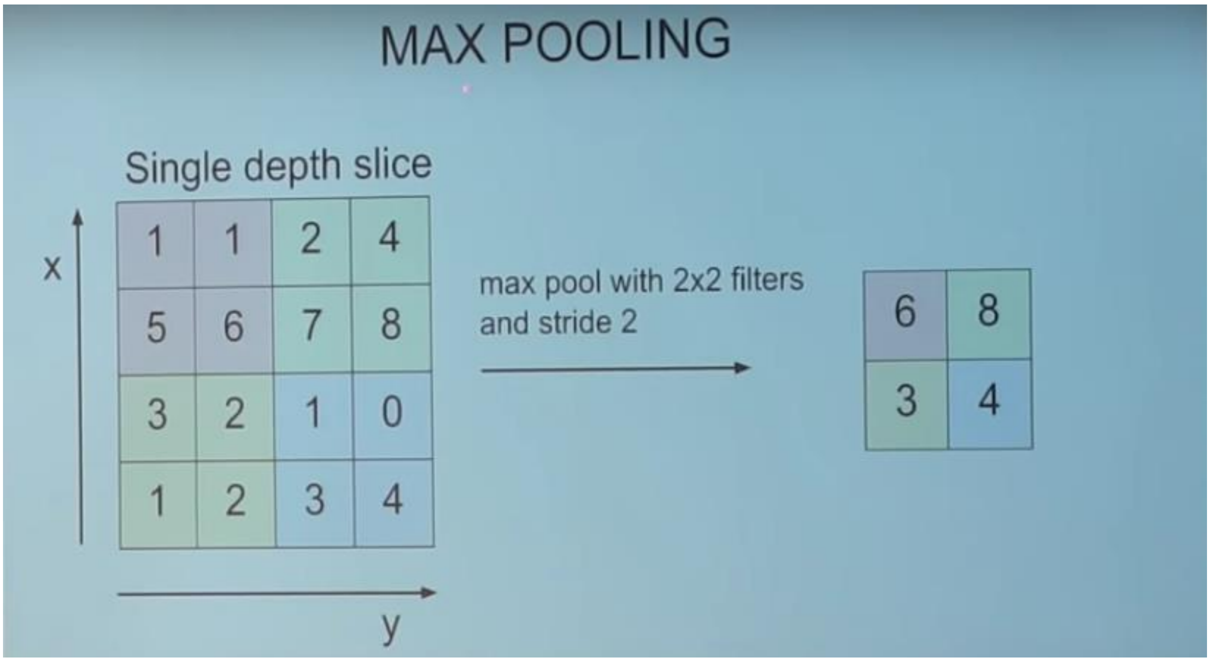
2. ***Rectified Linear Unit(RELU)Layer***

The RELU layer is an activation function that sets negative input values to zero. This simplifies and accelerates calculations and training, and helps to avoid the vanishing gradient problem. Mathematically it is defined as: $f(x)=\max(0,x)$ where x is the input to the neuron. Other activation functions include the sigmoid, tanh, leaky RELUs, Randomized RELUs and parametric RELUs.

3. ***PoolingLayer***

The Pooling layer is inserted between the Convolution and RELU layers to reduce the number of parameters to be calculated, as well as the size of the image (width and height, but not depth). Max-pooling is most commonly used; other pooling layers include Average pooling and L2-normalization pooling.Max-pooling simply takes the largest input value within a filter and discards theother values; effectively it summarizes the strongest activations over a neighborhood. The rationale is that the relative location of a strongly activated feature to another is more important than its exact location.

For Example,



4.*Fully Connected Layer*

The final layer in a CNN is the Fully Connected Layer, meaning that every neuron in the preceding layer is connected to every neuron in the Fully Connected Layer. Like the convolution, RELU and pooling layers, there can be 1or more fully connected layers depending on the level of feature abstraction desired. This layer takes the output from the preceding layer (Convolutional, RELU orPooling)as its input, and computes a probability score for classification into the different available classes. In essence, this layer looks at the combination of the most strongly activated features that would indicate the image belongs to a particular class. For example, on histology glass slides, cancer cells have a high DNA to cytoplasm ratio compared to normal cells. If features of DNA were strongly detected from the preceding layer, the CNN would be more likely to predict the presence of cancer cells. Standard neural network training methods with back propagation [10] and stochastic gradient descent help the CNN learn important associations from training images.

Recurrent Neural Networks(RNNs)

RNNs have traditionally been used in analyzing sequential data, such as the words in a sentence. Due to their ability to generate text, RNNs have been employed in text analysis tasks, like machine translation, speech recognition, language modelling, text prediction and image caption generation. In a plain RNN, the output of a layer is added to the next input, and this is fed back into the layer, resulting in a capacity for contextual 'memory'. To avoid vanishing gradient problems with back propagation through time, plain RNNs have evolved into Long Short Term Memory (LSTM) networks and Gated Recurrent Units (GRUs). These are modifications of RNNs to hold long term dependencies, and to discard or forget some of the accumulated information.

In the medical image analysis space, RNNs have been used mainly in segmentation. Chen et al.combined CNN and RNN to segment neuronal and fungal structures from three-dimensional electron microscope images. Using a multi dimensional LSTM, Stollenga et al.segmented both three-dimensional electron microscope images of neurons as well as MRI brain scans. Shin et al.describe annotating X-ray images with captions trained on radiology reports.

Chapter3

Applications in Medical Image Analysis

To the researcher, CNNs have been put to task for classification, localization, detection, segmentation and registration in image analysis. Machine learning research draws a distinction between localization (draw a bounding box around a single object in the image), and detection (draw bounding boxes around multiple objects, which may be from different classes). Segmentation draws outlines around the edges of target objects, and labels them (semantic segmentation). Registration refers to fitting one image (which may be 2 or 3 dimensional) onto another. This separation of tasks is based on different machine learning techniques and is maintained below.

To the clinician this separation of tasks is not that crucial, and it is the author’s opinion that a pragmatic machine learning system will incorporate some or all of the tasks into a unified system. It would be ideal to, in a single workflow, detect a lung tumor on a CT chest scan, and then localize and segment it away from normal tissue, and to prognosticate various treatment options, such as chemo therapy or surgery. Indeed, some of these tasks blur into one another in the papers discussed here. From the clinician’s perspective, classification ascertains if a disease state is present or not, ,i.e., is blood present on this MRI brain scan signifying a hemorrhagic stroke? Localization implies the identification of normal anatomy, for example, where is the kidney in this ultra sound image? This is in contrast to detection, which implies an abnormal, pathological state, for example, where are all the lung tumors in this CT scan of the lung? Segmenting the outline of a lung tumor helps the clinician determine its distance from major anatomical structures, and helps to answer a question such as, should this patient be operated on, and if so, what should be the extent of resection?

1. Classification

Classification is sometimes also known as Computer-Aided Diagnosis (CADx). Lo et al. described a CNN to detect lung nodules on chest X-rays as far back as 1995. They used 55 chest x-rays and a CNN with 2 hidden layers to output whether or not a region had a lung nodule. The relative availability of chest x-ray images has likely accelerated deep learning progress in this modality. Rajkomar et al. augmented 1850 chest x-ray images into 150,000 training samples. Using a modified pre-trained GoogLeNet CNN , they classified the orientation of the images into frontal or lateral views with near 100% accuracy. Although this task of identifying the orientation of the chest x-ray is of limited clinical use, it does demonstrate the effectiveness of pre-training, and data augmentation in learning the relevant image metadata, as part of an eventually fully-automated diagnostic work-flow. Diabetic Retinopathy (DR) can also be diagnosed using CNNs. Using digital photographs of the fundus of the eye, Pratt et al. [56] trained a CNN with 10 convolutional layers and 3 fully connected layers on approximately 90,000 fundus images. They classified DR into 5 clinically used classifications of DR severity, with 70% accuracy. Abramoff et al. evaluated a commercial device, the IDx-DR version X2.1 (IDx LLC, Iowa City, Iowa, USA) to detect DR. The author does
Not disclose the CNN architectures but states they are inspired by Alexnet and VGGNet. The device, trained on up to 1.2 million DR images, obtained an AUC score of 0.98.

pneumonia classification in particular achieved an Area Under Curve (AUC) score of 0.7632 with Receiver Operating Characteristics (ROC) analysis.

Moreover, on a test set of 420 images, CheXNet matched or bettered the performance of 4 individual radiologists, and also the performance of a panel comprising of 3 radiologists. Shen et al. [50] used CNNs combined with Support Vector Machine (SVM) and Random Forest (RF) classifiers to classify lung nodules into benign or malignant, based on 1010 labelled CT lung scans from the Lung Image Database Consortium (LIDC-IDRI) dataset. They used 3 parallel CNNs with 2 convolution layers each, with each CNN taking image patches at different scales to extract features. The learned features were used to construct an output feature vector, which was then classified using either a SVM with radial basis function (RBF) filter or RF classifier into benign or malignant. Their method classified nodules with 86% accuracy and they also found that it was robust against different levels of noise inputs. Li et al. [51] used 3-dimensional CNNs to interpolate missing imaging data between MRI and PET images. 830 patients with MRI and PET scans from the Alzheimer Disease Neuro imaging Initiative (ADNI) database were studied. 3-D CNNs were trained with MRI and PET images as input and output respectively, and used to reconstruct PET images from patients who did not have them. Their reconstructed PET images almost matched ground truth results of disease classification, but one caveat is that issues of overfitting were not addressed, limiting the potential generalizability of their technique. Unsupervised learning methods are also an active area of research. Plis et al.used Deep Belief Networks to extract features from functional fMRI (fMRI) images, and MRI scans of patients with Huntington Disease and Schizophrenia. Suk et al. classified fMRI images into diagnoses of Healthy or Mild Cognitive Impairment, using a stacked architecture of RBMs to learn hierarchal functional relationships between different brain regions. Looking outside the usual CNN models, Kumar et al. compared the performance of the well-known CNNs Alexnet and VGGNet to other techniques, namely Bag of Visual Words (BOVV) and Local Binary Patterns (LBP). Interestingly, the BOVV technique performed the best at classifying histopathological images into 20 different tissue types.

2. Localization

Localization of normal anatomy is less likely to interest the practicing clinician although applications may arise in anatomy education. Alternatively, localization may find use in fully automated end-to-end applications, whereby the radiological image is autonomously analyzed and reported without any human intervention. Yan et al. looked at transverse CT image slices and constructed a two stage CNN where the first stage identified local patches, and the second stage discriminated the local patches by various body organs, achieving better results than a standard CNN. Roth et al. trained a CNN with 5 convolution layers to discriminate approximately 4000 transverse axial CT images into one of 5 categories: neck, lung, liver, pelvis, legs. He was able to achieve a 5.9% classification error rate and an AUC score of 0.998, after data augmentation techniques. Shin et al. used stacked autoencoders on 78 contrast-enhanced MRI scans of the abdominal region containing liver or kidney metastatic tumors, to detect the locations of the liver, heart, kidney and spleen. Hierarchal features were learned over the spatial and temporal domains, giving detection accuracies of between 62% and 79%, depending on the organ.

3. Detection

Detection, sometimes known as Computer-Aided Detection (CADe) is a keen area of study as missing a lesion on a scan can have drastic consequences for both the patient and the clinician. The task for the Kaggle Data Science Bowl of2 017 [64] involved the detection of cancerous lung nodules on CT lung scans. Approximately 2000 CT scans were released for the competition and the winner Fangzhou achieved a logarithmic loss score of 0.399. Their solution used a 3-D CNN inspired by U-Net architecture to isolate local patches first for nodule detection. Then this output was fed into a second stage consisting of 2 fully connected layers for classification of cancer probability. Shin et al. evaluated five well-known CNN architectures in detecting thoraco-abdominal lymph nodes and Interstitial lung disease on CT scans. Detecting lymph nodes is important as they can be a marker of infection or cancer. They achieved a mediastinal lymph node detection AUC score of 0.95 with a sensitivity of 85% using GoogLeNet, which was state of the art. They also documented the benefits of transfer learning, and the use of deep learning architectures of up to 22 layers, as opposed to fewer layers which was the norm in medical image analysis. Overfeat was a CNN pretrained on natural images that won the ILSVRC 2013 localization task . Ciompi et al. applied Overfeat to 2-dimensional slices of CT lung scans oriented in the coronal, axial and sagittal planes, to predict the presence of nodules within and around lung fissures. They combined this approach with simple SVM and RF binary classifiers, as well as a Bag of Frequencies, a novel 3-dimensional descriptor of their own invention. Other than lung lesions, there are also a myriad of other applications, including detecting malignant skin cells. Esteva et al. used 130,000 dermatological photographs and dermoscopic images to train a GoogLe Net Inception V3 CNN, with no hand-crafting of features. The CNN out performed human dermatologists in classifying the images as benign, malignant or non-neoplastic lesions, reaching an accuracy of 72% compared to the 65% and 66% accuracies obtained by 2 human dermatologists. The CNN again bettered 21 human dermatologists at deciding treatment plans for two types of skins cancers: carcinoma and melanoma. This task involved 376 biopsy-proven images, and the CNN achieved AUC scores of between 0.91 to 0.96.

4. Segmentation

CT and MRI mage segmentation research covers a variety of organs such as liver, prostate and knee cartilage, but a large amount of work has focused on brain segmentation, including tumor segmentation. The latter is especially important in surgical planning to determine the exact boundaries of the tumor in order to direct surgical resection. Sacrificing too much of eloquent brain areas during surgery would cause neurological deficits such as limb weakness, numbness and cognitive impairment. Traditionally, medical anatomical segmentation was done by hand, with a clinician drawing outlines slice by slice through an entire MRI or CT volume stack, therefore it is ideal to implement a solution that automates this laborious task. An excellent review of brain MRI segmentation was written by Akkus et al. , who reviewed various CNN architectures and metrics used in segmentation. Additionally, he also detailed the numerous competitions and their datasets, such as Brain Tumor Segmentation (BRATS), Mildtraumatic brain injury outcome prediction (MTOP) and Ischemic Stroke Lesion Segmentation (ISLES). CNNs, each with a different 2-dimensional input patch size, running in parallel to classify and segment MRI brain images of 22 pre-term infants and 35 adults into different tissue classes such as white matter, grey matter and cerebrospinal fluid. The advantage of using 3 different input patch sizes is that each focuses on capturing different aspects of the image, with the smallest patch focused on local textures while the larger patch sizes assimilated spatial features. Overall, the algorithm achieved good accuracy, with Dice coefficients between 0.82 and 0.87. Most segmentation research has been on 2-dimensional image slices, but Milleterai et al. applied 3-dimensional CNN to segment MRI prostate images from the PROMISE 2012 challenge data set. Their proposed V-net was inspired by

Ronner berger's U-Net architecture, and was trained on 50 MRI prostate scans and tested on 30 similar scans. V-net achieved a dice similarity coefficient score of 0.869, which was similar to that of the top placed teams in the challenge. Pereira et al. applied deliberately small filters of 33 size, to allow the design of a deeper 11 convolution layer CNN, and to reduce overfitting. Their CNN was trained on 274 MRI brain tumor scans of gliomas, a type of brain tumor with significant malignant potential, obtaining first place in the BRATS 2013 and second place in the BRATS 2015 challenge. Havaei et al.also looked at gliomas, and explored various 2-dimensional CNN architectures on the BRATS 2013 data set. Their algorithm performed better than the BRATS 2013 winner, and took 3 minutes to run, compared to 100 minutes. Their Input Cascade CNN had a cascaded architecture, with the output of a first CNN being fed into a second CNN. Chen et al. [79] proposed using up-sampled filters, atrous spatial pyramid pooling, and fully connected Conditional Random Fields (CRFs). These aid in enlarging the field of each filter's view at multiple scales and improve localization accuracy. With this architecture which they called Deep Lab, Chen et al. achieved state-of-the-art performance in the PASCAL VOC-2012 Image segmentation task, reaching 79.7% mean Intersection over Union (mIOU). There is some overlap with Moeskops'use of input patches at different scales, and it would be interesting to see how this work in image segmentation can be advanced further. A more recent study by Casamitjana et al. compared various 3dimensional CNN architectures.

5. Registration

Although the registration of medical images has many potential applications, which were reviewed by El-Gamal et al., their actual clinical use is encountered in niche areas. Image registration is employed in neurosurgery or spinal surgery, to localize a tumor or spinal bony landmark, in order to facilitate surgical tumor removal or spinal screw implant placement. A reference image is aligned to a second image, called a sense image and various similarity measures and reference points are calculated to align the images, which can be 2 or 3-dimensional. The reference image may be a pre-operative MRI brain scan and the sense image may be an intra operative MRI brain scan done after a first-pass resection, to determine if there is remnant tumor and if further resection is required. Using MRI brain scans from the OASIS dataset, Yang et al. stacked convolution layers in an encoder-decoder fashion, to predict how an input pixel would morph into its final configuration. They invoked the use of a Large deformation diffeomorphic metric mapping (LDDMM) registration model and achieved dramatic improvements in computational time. Miao et al. trained a 5 layer CNN on synthetic X-ray images in order to register 3-dimensional models of a knee implant, a hand implant, and a trans-esophageal probe onto 2-dimensional X-ray images, in order to estimate their pose. Their method obtained successful registrations 79–99% of the time, and took 0.1 seconds, a significant improvement over traditional intensity-based registration methods.

CHALLENGES

A recurring theme in machine learning is the limit imposed by the lack of labelled datasets, which hampers training and task performance. Conversely, it is acknowledged that more data improves performance, as Sun et al. shows using an internal Google dataset of 300 million images. In general computer vision tasks, attempts have been made to circumvent limited data by using smaller filters on deeper layers, with novel CNN architecture combinations, or hyper parameter optimization.

In medical image analysis, the lack of data is two - fold and more acute: there is general lack of publicly available data, and high quality labelled data is even more scarce. Most of the datasets presented in this review involve fewer than 100 patients. Yet the situation may not be as dire as it seems, as despite the small training datasets, the papers in this review report relatively satisfactory performance in the various tasks. The question of how many images are necessary for training in medical image analysis was partially answered by Cho et al. He ascertained the accuracy of a CNN with GoogLeNet architecture in classifying individual axial CT images into one of 6 body regions: brain, neck, shoulder, chest, abdomen, pelvis. With 200 training images, accuracies of 88–98% were achieved on a test set of 6000 images. While categorization into various body regions is not a realistic medical image analysis task, his report does suggest that the problem may be surmountable. Being able to accomplish classification with a small dataset is possibly due to the general intrinsic image homogeneity across different patients, as opposed to the near-infinite variety of natural images, such as a dog in various breeds, colors and poses.

An important, non-technical challenge is the public reception towards their health results being studied by a nonhuman actor. This situation is not helped by the apocalyptic artificial intelligence scenarios painted by some. Machine learning algorithms have surpassed human performance in image recognition tasks, and it is likely that they will perform better than humans in medical image analysis as well. Indeed, some of the papers in this review report that dermatologists and radiologists have already been bested by machine learning. Yet the question regarding legal and moral culpability arises when a patient is misdiagnosed, or suffers morbidity as a result of AI or AI-assisted medical management. This is accentuated by our inability to fully explain how the black-box of machine algorithms work. However, it is likely that our relationship will continue evolve and recalibrate as AI-based technologies mature and inexorably permeate different facets of our lives.

FUTURE APPLICATIONS

The traditional applications for medical image analysis were discussed in Section 3. New areas of research include prognostication, content-based image retrieval, image report or caption generation [100], [101], and manipulation of physical objects with LSTMs and reinforcement learning, involving surgical robots. A few innovative applications that span across traditional medical image analysis categories are described below.

An interesting application was reported by Nie et al. in which GANs were used to generate CT brain images from MRI images. This is remarkable, as it means that patients can potentially avoid the ionizing radiation from a CT scanner altogether, lowering cost and improving patient safety. Nie also exploited the ability of GANs to generate improved, higher resolution images from native images and reduced the blurriness in the CT images. A useful extension of resolution improvement techniques would be applying them to generate MRI images of higher quality. High quality MRI images require high tesla (and correspondingly costlier) MRI scanners. Algorithmically-generated high quality MRI images on a lower field-strength scanner would thus lower healthcare costs.

Chang demonstrated a novel application in the nascent area of radio-genomics, which uses radiological images to predict the underlying molecular origin of a tissue. He first used an auto encoder to learn latent features from MRI images of glioblastoma multiforme (GBM), a malignant brain tumor, from The Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM) data collection. The learned features were then fed into a fully connected classifier layer to classify a MRI scan into one of 4 known molecular sub-types of GBM.

Although still early, Chang’s work could potentially diagnose a GBM sub-type and obviate the need for surgical biopsy and molecular assays. The generalizability of this technique to tumors elsewhere in the body is also promising.

Coudray et al. accomplished an analogous task, but used histopathological images to classify lung cancer subtypes, and to predict common genetic mutations. Knowing the genetic mutations is helpful in prognosticating length of survival and guiding the choice of chemotherapy. Their method outperforms a human pathologist, and the prediction of genetic mutations had AUC scores of between 0.73 to 0.86.

Tsochatzidis et al. [112] described an original work combining content-based image retrieval (CBIR) and computer aided diagnosis (CADx). In essence, their model segmented a lesion on a query image, and compared this to the segmented lesions in their database, consisting of 400 Regions of interest derived from the Digital Database for Screening Mammography (DDSM). The basis of comparison were the Euclidean distances between the representation vectors of the query lesion and database lesions. The model then outputs both reference images and a likelihood of a lesion being benign or malignant. They reported that their combined CBIR and CADx method resulted in state of the art prediction accuracy of 81%. These examples highlight how the field of machine learning in medical image analysis is changing rapidly, and that there may still be numerous applications which have not been conceived of yet.

Chapter4

CONCLUSION

Medical image analysis is an active field of research for machine learning, partly because the data is relatively structured and labelled, and it is likely that this will be the area where patients first interact with functioning, practical artificial intelligence systems. This is significant for two reasons. Firstly, in terms of actual patient metrics, medical image analysis is a litmus test as to whether artificial intelligence systems will actually improve patient outcomes and survival. Secondly, it provides a testbed for human-AI interaction, of how receptive patients will be towards health-altering choices being made, or assisted by a non-human actor.