Spring 2020

The Effectiveness of Transfer Learning Systems on Medical Images

James Boit
Dakota State University

THE EFFECTIVENESS OF TRANSFER LEARNING SYSTEMS ON MEDICAL IMAGES

A dissertation submitted to Dakota State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Information Systems

Spring 2020

By

James Boit

Dissertation Committee:

Dr. David Zeng- Chair

Dr. Ronghua Shan

Ms. Linda Parks

DISSERTATION APPROVAL FORM

This dissertation is approved as a credible and independent investigation by a candidate for the Doctor of Philosophy degree and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this dissertation does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department or university.

Student Name: James Boit

Dissertation Title: The Effectiveness of Transfer Learning Systems on Medical Images

Dissertation Chair/Co-Chair: David Zeng Date: June 4, 2020

Name: David Zeng

Dissertation Chair/Co-Chair: ________ Date: ________

Name:

Committee member: Ronghua Shan Date: June 4, 2020

Name: Ronghua Shan

Committee member: Linda Parks Date: June 4, 2020

Name: Linda Parks

ABBREVIATIONS

AI	Artificial Intelligence
CIFAR	Canadian Institute for Advanced Research
CNN	Convolutional Neural Network
COVID-19	Coronavirus 2019
CP	Conference Paper
CT	Computed Tomography
CXR	Chest X-ray
DCNN	Deep Convolutional Neural Network
DIM	Dense Inception Module
DL	Deep Learning
DNN	Deep Neural Network
GAN	Generative Adversarial Networks
GPU	Graphical Processing Units
JA	Journal Article
JSTR	Japanese Society of Radiological Technology
LSTM	Long-Short Term
MNIST	Modified National Institute of Standards and Technology
MPDBM	Multi-Prediction Deep Boltzmann Machine
MRI	Magnetic Resonance Imaging
MUNIT	Multi-modal Unsupervised Image-to-image Translation
OCL	Optimal Cut-off Layer
OCT	Optical Coherence Tomography
PA	Publication Article
PRISMA	Preferred Reporting Items for Systematic Literature review
RCNN	Recurrent Convolutional Neural Network
RNN	Recurrent Neural Network
RQ	Research Question
SGD	Stochastic Gradient Descent

ABSTRACT

Deep neural networks have revolutionized the performances of many machine learning tasks such as medical image classification and segmentation. Current deep learning (DL) algorithms, specifically convolutional neural networks are increasingly becoming the methodological choice for most medical image analysis. However, training these deep neural networks requires high computational resources and very large amounts of labeled data which is often expensive and laborious. Meanwhile, recent studies have shown the transfer learning (TL) paradigm as an attractive choice in providing promising solutions to challenges of shortage in the availability of labeled medical images. Accordingly, TL enables us to leverage the knowledge learned from related data to solve a new problem.

The objective of this dissertation is to examine the effectiveness of TL systems on medical images. First, a comprehensive systematic literature review was performed to provide an up-to-date status of TL systems on medical images. Specifically, we proposed a novel conceptual framework to organize the review. Second, a novel DL network was pretrained on natural images and utilized to evaluate the effectiveness of TL on a very large medical image dataset, specifically Chest X-rays images. Lastly, domain adaptation using an autoencoder was evaluated on the medical image dataset and the results confirmed the effectiveness of TL through fine-tuning strategies.

We make several contributions to TL systems on medical image analysis: Firstly, we present a novel survey of TL on medical images and propose a new conceptual framework to organize the findings. Secondly, we propose a novel DL architecture to improve learned representations of medical images while mitigating the problem of vanishing gradients. Additionally, we identified the optimal cut-off layer (OCL) that provided the best model performance. We found that the higher layers in the proposed deep model give a better feature representation of our medical image task. Finally, we analyzed the effect of domain adaptation by fine-tuning an autoencoder on our medical images and provide theoretical contributions on the application of the transductive TL approach. The contributions herein

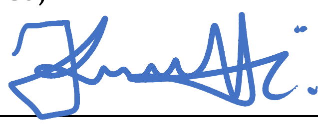
reveal several research gaps to motivate future research and contribute to the body of literature in this active research area of TL systems on medical image analysis.

DECLARATION

I hereby certify that this dissertation constitutes my product, that where the language of others is set forth, quotation marks so indicate, and that appropriate credit is given where I have used the language, ideas, expressions, or writings of another.

I declare that the dissertation describes original work that has not previously been presented for the award of any other degree of any institution.

Signed,



James Boit

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GENERAL INTRODUCTION

1.1 Background and motivation

That an image is worth a thousand words, is an English adage that is very relevant today in the medical imaging domain. Medical imaging techniques such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), produce pictures that contain hidden information about the medical diagnosis. Recent advances in deep learning (DL) technologies have achieved tremendous breakthroughs in many computer vision problems, for example, object detection, localization, segmentation, and classification outperforming existing traditional approaches that rely on the extraction of handcrafted features. Moreover, these deep learning networks such as Convolutional Neural Network (CNNs) have found great success in computer vision recognition tasks (Farabet et al., 2013; Girshick et al., 2014; Hariharan et al., 2014; Krizhevsky et al., 2012) that automatically reveal hidden patterns from image features with a high degree of accuracy that has surpassed human judgment on many tasks. Further, deep learning methods present a phenomenal opportunity when employed on very large-scale datasets. However, the increasingly high costs of medical image acquisition and processing is a big hindrance for researchers who want to produce robust models from analysis of very large datasets. Deep learning techniques, for example, supervised transfer learning systems and data augmentation techniques can provide researchers with the opportunities to mitigate the challenges of limited datasets and resource-intensive processes of medical image acquisition or annotations.

Recent literature reveals remarkable progress has been made in the field of computer vision and recognition with the availability of very large-scale datasets such as ImageNet (Deng et al., 2009). The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) (Russakovsky et al., 2015), a widely popular online competition in object recognition, have contributed tremendously in pushing the boundary of possibilities for providing solutions to image recognition problems sustained by improvements of CNNs since 2012. These developments can be attributed to improvements in deep learning algorithms and model architectures sustained by the availability of increased computational power and big data. Ultimately, these benefits can translate into better accuracy, timely diagnosis of medical images, and help alleviate diagnostic decisions in healthcare environments with resource constraints.

Over the years, there has been a paradigm shift from traditional machine learning to deep learning, more specifically on computer vision recognition tasks. Deep learning pipelines facilitate the autonomous learning of complicated features to discover patterns that are useful in solving many application problems in the computer vision domain. Figure 1 shows the shift from traditional machine learning to deep learning systems (Agarwal, 2019).

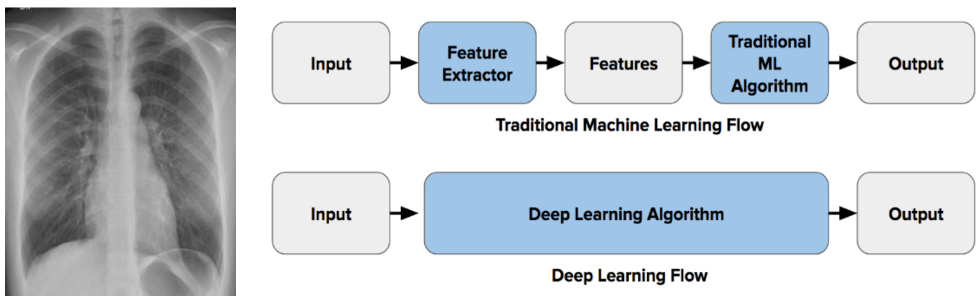


Figure 1. Traditional versus Deep Learning flow

This dissertation is inspired by the recent advances in CNN architecture developments, limited availability of large public data sets for domain-specific tasks, and by the goal to improve existing deep CNN (DCNN) to leverage transfer learning strategies. Moreover, the ability of pretrained CNN architectures to generalize across different domains and tasks is an open and ongoing research problem. The objective of this dissertation is, therefore, is to examine the effectiveness of transfer learning systems on medical images. Drawing on an examination of literature review, and development of novel CNN methods, this work is organized into three distinct but related parts to theoretically and empirically evaluate the effectiveness of transfer learning systems on medical images.

1.2 Scope

This dissertation is focused on the effectiveness of transfer learning systems on medical images. More specifically, we aim to conduct a comprehensive systematic literature on transfer learning systems on medical images. Next, we propose a novel deep model to experiment on the effectiveness of supervised transfer learning on the classification of medical images. Lastly, we also demonstrate the use of domain adaptation approach in transductive transfer learning using an autoencoder on medical images. This work will not cover other CNN architectures outside those we used for development nor attempt to explain the technical aspects and mathematical computations involved. However, succinct explanations are given throughout this work to answer the research questions and objectives outlined in sections 1.8 and 1.4 respectively.

1.3 Statement of the problem

The objective in image analysis, specifically in image classification tasks, is to map images to class labels. The images are treated as input labels where the image pixels are extracted and represented as feature vectors or feature maps. On the other hand, the output labels are represented as a probability distribution containing a probability value of either a binary or multiclass problem. One of the main problems in developing a robust image classification model is to identify a feature space that can clearly and efficiently distinguish feature maps from a probability distribution of feature vectors. This can be achieved through feature selection or extraction techniques and more recently, the use of CNNs to learn the feature maps of different image classes. The other problem is to describe how data or new samples can be generated from a probabilistic model; this is called the *generative model*. The Generative model predicts the joint distribution $p(x,y)$ by using the rules of the Bayes Theorem. Similarly, the use of decision boundaries to distinguish classes by learning the conditional distribution $p(y|x)$, this technique is referred to as the *discriminant model*. Examples of generative and discriminant models are Hidden Markov models and logistic regression classifiers respectively.

The two problems identified earlier are classical issues facing the computer vision and recognition communities. However, with limited availability of large training samples of medical images, transfer learning offers an attractive choice to address the second problem. Although, transfer learning presents many research issues such as: What to transfer; how to transfer; and when to transfer (Pan & Yang, 2010), the benefits of using transfer learning in many applications can ensure knowledge transfer across domains or tasks with similar or different feature spaces over single or multiple source domains. Moreover, advances in classes of deep neural networks (DNN) have significantly made it easier to solve the first problem such that with a few to an intermediate number of CNN layers, output labels of different classes can be classified, for example, using SoftMax activation functions or classifiers. These problems provide an ongoing research problem in the visual recognition of tasks such as classification and segmentation. Therefore, transfer learning provides a large playing ground to explore the use of deep learning models to further investigate these two problems within the context of medical images.

1.4 Objectives of the research

The primary research objective of this work is to evaluate the effectiveness of transfer learning systems on medical images. Throughout this work, supervised inductive and transductive transfer learning are evaluated on the medical image datasets. Other objectives of this work include:

1. Perform an up-to-date systematic literature review of transfer learning systems on medical images using the PRISMA guidelines.
2. Implement and evaluate a novel deep model on a very large medical dataset.
3. Evaluate the effectiveness of transfer learning strategies such as the use of pretrained models and fine-tuning techniques on medical images.
4. Identify the Optimal Cut-off Layer (OCL) that produces a robust model performance for generalizability purposes.
5. Demonstrate the effectiveness of applying domain adaptation with an autoencoder on medical images.

The knowledge transfer from *natural images* to solve computer vision tasks on *medical images* via transfer learning is a current trend in deep learning with the potential to improve learning performance on use cases where there is limited availability of medical images. Therefore, we hypothesize that transfer learning with fine-tuning of a pretrained model and jointly trained on medical images can significantly increase learned feature interactions inside our novel deep model thus improving the goal of finding an optimal model that is robust and can potentially perform across multiple but similar domains.

1.5 Contributions

Corresponding to the research objectives, the primary expected contributions of this work can be summarized in the following:

From a *theoretical perspective*, we aim to inform the body of knowledge of the application of transfer learning on medical images. Most notably, a novel systematic literature review was conducted to summarize the findings. Next, we propose a new conceptual framework to classify the results to inform future research directions. Similarly, we provide theoretical explanations on the behavioral performances of strategically positioning of the *Dense-Inception Network* (DINET) modules at different parts of the deep neural network that gives the optimal feature learning of medical images on vision recognition tasks while mitigating the problem of vanishing gradients (i.e. the gradual minimization of gradient error term through backpropagation) phenomenon. One key observation was the increased learning activity inside the model among the feature vectors due to the enforcement of feature re-use. We believe that the findings of this work will contribute to the growing body of literature in transfer learning systems on medical images.

From a *methodological perspective,* we analyze the effectiveness of an autoencoder, specifically the U-Net model architecture using transfer learning approaches and explore potential opportunities for further research. Furthermore, we experiment with public datasets for chest radiographs, a challenging problem in medical image segmentation. We also exploit data augmentation techniques to examine performance improvements in applying transfer learning techniques.

From a *practical perspective*, we propose a novel architecture that leverages multi-scale representations of learned features from shallow layers (generic layers) and forms high-level features transformed from deeper layers which can classify and distinguish medical images much more accurately into different classes. Also, through rigorous experimentations of our novel deep model, we show that the problem of vanishing gradients can be alleviated due to increased activity of learned feature maps inside the network. Moreover, we show that by using transfer learning systems, knowledge learned and transferred from both natural and domain-specific images can be effective in medical image diagnosis.

1.6 Publication

The work in this of this dissertation is connected to the publication efforts from the following:

Zeng, D, Boit, J, and Winston Z. (2019) Effectiveness of transfer learning on medical image classification using chest X-ray 14 dataset.

Transfer learning is significantly gaining rapid adoptions as an important tool for diagnosis and interpretation of medical images by decreasing the time spent in predictions, improving the accuracy in identifying abnormalities, and, therefore, enhancing the clinical outcomes of patients. We test the effectiveness of transfer learning (TL) techniques, namely, transferring knowledge from deep learning models pretrained with general-purpose images to medical image classification using the Chest X-ray 14 dataset, comprising of 112,120 frontal-view chest X-ray images from 30,805 unique patients. We use the DenseNet-121 architecture, pretrained on Image Net, as our baseline model, and perform binary classification on our dataset. The results show that fine-tuning with data augmentation gives a more robust model performance and we propose that identifying the optimal cut-off layer during fine-tuning provides a novel approach for higher-order representation of medical features. For future research, we will combine fine-tuning approaches with hyperparameter optimization, adding non-image patient data, finding optimal data augmentation and model architecture, and generating high-resolution medical images using generative adversarial networks to improve model performance.

1.8 Dissertation outline

This dissertation is structured into the following three distinct but related parts.

Part 1: This section provides an up-to-date systematic literature review of transfer learning systems on medical image analysis. To evaluate the extent to which transfer learning strategies and techniques have been used on medical images across anatomical areas of the human body, peer-reviewed articles are methodologically analyzed using the Preferred Reporting Items for Systematic Literature review (PRISMA). Also, a taxonomy and conceptual classification framework are developed to classify the findings, to show the trends over time of how transfer learning systems have been applied to medical images within the context of the human anatomy. Therefore, this work aims to address the following main research questions **(RQ)**:

1. How have transfer learning systems been applied to medical images?
2. What are the learning settings used for transfer learning on medical images?

Part 2: This section examines the effectiveness of transfer learning systems on medical images using a novel deep model as our primary objective. In particular, the motivation for this work is to implement and evaluate a novel architecture to stimulate learning of relevant feature maps of the medical images, to help diagnose thoracic pathologies. Besides, pretrained models and fine-tuning strategies, key components of transfer learning systems are used to investigates the proposed model's effectiveness in transferring features maps from the source domain (*natural images*) to the target domain (*medical images*). Moreover, rigorous experiments are conducted with fine-tuning operations to determine the optimal cut-off layer (OCL) of the proposed model that gives the optimal feature representation on a medical image recognition task. To accomplish this objective, the following research questions are proposed:

1. Does the proposed deep model alleviate the problem of vanishing gradients efficiently?
2. Does the proposed deep model improve the effectiveness of transfer learning for medical image classification?
3. What is the Optimal Cut-off Layer that produces the best model for deep feature representation for medical image classification?

To answer the above questions, we fully trained a pretrained model and extracted feature maps that were relevant for retraining for our medical image dataset.

Part 3: This is the last section of the dissertation that employs the transductive learning approach to transfer learning where domain adaptation is evaluated on the medical image dataset. More specifically, an autoencoder, a type of neural network is used to learn source data distributions from a good performing model to solve a task on a related target distribution (*medical images*). In this section, supervised domain adaptation applying the re-training method is used to reconstruct the errors from the source data which are beneficial for classification or segmentation tasks. The research objectives of this work include:

1. What is the effectiveness of transfer learning systems in using autoencoders for medical images?

In summary, this dissertation seeks to answer fundamental research questions highlighted in this three-part series format regarding the effectiveness of transfer learning systems on medical images with an emphasis on theoretical and practical perspectives. The contributions are of significant interest to the multi-disciplinary audience in information systems, health practitioners, policy-makers on matters health, computer vision researchers, and the medical imaging community.

PART I: A SURVEY OF TRANSFER LEARNING IN MEDICAL IMAGE ANALYSIS

2.1 Overview

This section provides an overview of the extant literature reviews on medical image analysis and more specifically on aspects of deep learning technologies on medical images. We will summarize the related works in Table 1 where the reader will find an up-to-date overview of related studies that focus on deep learning technologies and architectures on medical images. The main idea in this section is to provide a current status of reviews that cover the use of deep learning technologies on medical image analysis.

2.2 Literature review

In recent years, the adoption of deep learning technologies has led to remarkable growth in visual recognition tasks such as object detection, classification, localization and segmentation, key components of computer vision problems. By definition, Deep learning is a computation model comprised of several processing layers with the ability to learn representations of data in many abstraction levels (LeCun et al., 2015). Applications of Deep learning in the computer vision field relies on CNNs, one of the most popular deep learning architectures, utilized for image recognition tasks. Over the years, CNNs have been widely used in visual recognition problems applying both supervised and unsupervised methods with notable achievements in translation invariance in computer vision tasks (Krizhevsky et al., 2012; LeCun et al., 1989). With the resurgence of increased interests in CNN architectures on visual recognition tasks, starting from AlexNet (Krizhevsky et al., 2012) in 2012, the growth of deep learning application on medical image analysis and to a greater extent, the use of pretrained models have dramatically increased over the last seven years. For example, seminal contributions have been made by Litjens et al., (2017) in the field of deep learning applications, where they comprehensively reviewed a plethora of published studies estimated at over 300 papers focused on medical image analysis. Another study by Altaf et al., (2019) presented a survey of recent publications in 2018 focusing on the application of deep learning methods on medical imaging analysis. In related works by Anwar et al., (2018) a review of deep learning techniques leveraging on CNN architectures and their application on medical image analysis was presented. In a similar attempt, Sengupta et al., (2019) presented a review of the algorithmic foundation of deep learning architectures, their application, and current trends in medical image analysis. In another work, Kumar and Bindu, (2019) provided a systematic literature review of the deep learning application on medical imaging following the PRISMA guidelines. Although the findings were categorized according to the visual recognition tasks, the use of transfer learning strategies and techniques on medical image analysis was not discussed.

Pehrson et al., (2019) surveyed the use of deep learning algorithms to detect pulmonary nodules in thoracic CT scans. In this study, the authors used the PRISMA methodology to investigate deep learning algorithms that were used on the Lung Image Database Consortium Image Collection (LIDC-IDRI), a very large database comprising 7371 lesions. Additionally, Mazurowski et al., (2018) conducted a review of deep learning in radiology covering visual recognition tasks such as classification, segmentation, and detection. Moreover, transfer learning approaches were highlighted in the classification task. This approach remains briefly discussed in the literature. In a different work by Yi et al., (2019) the authors presented a survey of Generative Adversarial Networks (GANs) in medical imaging. Although there are several studies on different deep learning dimensions on medical images, the review focusing on transfer learning remains limited. To fill this literature gap, our study identifies and focuses on the application of transfer learning systems on medical images following the scheme by Litjens et al., (2017). The following Table 1. provides a summary of surveys from the literature that contributed towards the application of different dimensions of deep learning on medical image analysis.

Table 1: Summary of review of papers related to DL on medical image analysis

Author	**Focus**	**PRISMA**	**Remarks**
Altaf et al., (2019)	To provide a review of recent deep learning techniques and methods on medical imaging published in 2018.	NONE	Best practices and concepts of transfer learning are discussed. Also, transfer learning strategies are highlighted in studies that use detection or localization tasks.
Anwar et al., (2018)	This study focused on a survey of deep learning techniques and their application on medical image analysis.	NONE	An estimate of 10 papers was summarized using the CNN method and its performance accuracy reported.
Ker et al., (2018)	This work focused on applications of medical image classification, localization, detection, segmentation, and registration.	NONE	Reviewed several highly cited articles to articulate the application of different supervised and unsupervised deep learning models on medical image analysis tasks
Litjens et al., (2017)	This study examines over 300 studies on the use of deep learning for image classification, segmentation, registration, etc.	NONE	Summary of deep learning techniques, and challenges to medical imaging tasks
Mazurowski et al., (2018)	This study reviews the use of deep learning in radiology with a focus on classification, segmentation, and detection. Other tasks reviewed in the study included: image registration, image generation, image enhancement, content-based image retrieval, and objective image quality assessment.	NONE	Transfer learning strategies and deep features are covered only for image classification tasks.
Pehrson et al., (2019)	This study surveyed the literature on the use of deep learning algorithms to detect pulmonary nodules derived from the Lung Image Database Consortium Image Collection (LIDC-IDRI)	YES	No mention of transfer learning strategies in the study.
Sengupta et al., (2019)	The focus of this study is to review deep learning architectures, applications, and trends with an estimate of 12 papers on medical image processing.	NONE	The theoretical and mathematical explanation of different types of CNN is provided alongside application in other industries such as financial services and power systems
Shorten and Khoshgoftaar, (2019)	Types of Data augmentation methods	NONE	Developed taxonomy for image data augmentations.
Kumar and Bindu, (2019)	The objective of this research was to provide a systematic literature review on deep learning methods used in medical image analysis tasks such as classification, localization, detection and segmentation	YES	The review was conducted between 2012 and 2018. Also, Transfer learning techniques or strategies were not explicitly mentioned or covered.
Yi et al., (2019)	To present a comprehensive overview of the literature on Generative Adversarial Networks (GAN) on medical imaging	NONE	None of the transfer learning strategies was reviewed together with GAN application on medical imaging.
Zhuang et al., (2019)	To review and summarize current transfer learning approaches with a focus on data and models	NONE	Comprehensively summarizes the mechanisms and strategies of transfer learning with model examples on Natural language processing (NLP) problems.

2.3 Classification framework

We propose a graphical representation of the conceptual framework for categorizing the literature on the transfer learning systems on medical images. The framework shown in Figure 2 is based on ideas from prior literature of transfer learning (Pan & Yang, 2010), deep learning on medical images (Litjens et al., 2017; Mazurowski et al., 2018), and medical imaging modalities (Elangovan & Jeyaseelan, 2016). Figure 2 consists of several layers: (1) Transfer learning settings; (2) Methods; (3) Tasks; (4) Imaging modalities; (5) Anatomical application areas; and (6) Systems of the human body. Moreover, Figure 5 organizes the review of the literature to reflect the categories of the several identified layers. A brief overview of each of the layers is given to motivate the review and analysis process.

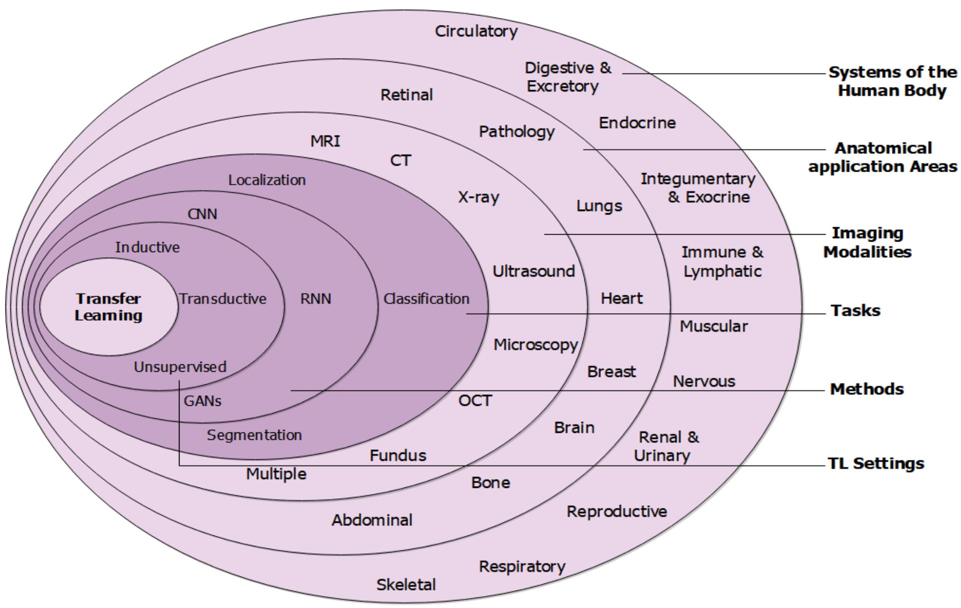


Figure 2. A conceptual framework for the application of TL on medical images

2.4 Deep learning architectures

Convolutional Neural Networks (CNNs or ConvNets) are a group of feed-forward neural networks comprising of components such as convolutional, pooling, and fully connected layers and mostly used to solve computer vision challenges. With the latest developments of faster Central Processing Units (CPUs) and efficient Graphical Processing

Units (GPUs), deep learning is progressively permeating every aspect of our daily lives such as diagnosis of Alzheimer's (Awate, 2019), fraud detection, facial recognition, self-driving vehicles, and virtual assistant. Recently, deep convolutional neural networks (H. Shin et al., 2016; Krizhevsky et al., 2012) leveraged on CNN architectures have proven significant success on image classification tasks (Rawat & Wang, 2017), pushing the performance boundary that surpasses human judgment. The primary idea of DL is the training of features from raw data. Conventionally, the DL algorithm will create a hierarchical representation for feature detection such as eyes, nose, and ears.

DCNNs comprise of multiple hidden layers that enable representations of complex features learned from the relationships between the inputs and outputs. In the computer vision field, modern CNN architectures that have found tremendous success in many image recognition tasks include AlexNet (Krizhevsky et al., 2012), DenseNet (Huang et al., 2017), Inception (Szegedy et al., 2015), VGGNet (Simonyan & Zisserman, 2015), and ResNet (He et al., 2016). It is worth mentioning that several versions of the prevenient CNN architecture have already been developed in terms of going deeper and wider with intentions of improving model performances and parameter efficiency. Other approaches in the design and development of CNN architectures have adopted an architectural combination of hybrid networks, for example, Inception-ResNet-v2 (Szegedy, Ioffe, et al., 2016). All these different variants of CNN architectures have played a great role in contributing innovations towards the diagnosis of diseases in the medical domain.

Beyond the CNNs, variants of the feedforward neural networks with the addition of feedback connections, include Recurrent Neural Networks (RNNs) (Bengio et al., 1994; Jordan, 1997), Long Short-term memory (LSTM) (Hochreiter & Schmidhuber, 1997) and GANs (Goodfellow et al., 2014) have become increasingly popular in segmentation tasks, NLP and unsupervised learning. In biomedical image analysis, U-Net (Ronneberger et al., 2015), has proven to be very successful for medical image segmentation tasks. Over time, a more comprehensive literature review on historical and state-of-the-art deep learning algorithms can be found in Pouyanfar et al.,(2018).

2.5 Overview of transfer learning approaches

Transfer learning (Caruana, 1997; Thrun, 1996) is a popular concept currently used in deep learning to refer to the reuse of features from a pretrained network (usually from natural images) to be applied to a new problem. Conventionally, Deep neural networks require vast amounts of large labeled datasets, for example, ImageNet (Deng et al., 2009), one of the largest datasets for image classification (with over 14 million images belonging to over 1000 object classes) and very powerful computing resources to solve many challenging computer vision problems. However, now TL presents an attractive proposition in solving real-world problems such as medical image recognition tasks, where there is a shortage of labeled datasets and training DCNNs from scratch could be computationally expensive. In this context, two transfer learning techniques have been widely applied for image recognition tasks: (1) Pretrained networks as a feature extractor and (2) fine-tuning a pretrained network (Litjens et al., 2017). In principle, transfer learning translates knowledge that has already been learned in one domain (*source*) and applied to solve a new task in a different but related problem (*target*). For example, a person who knows how to ride a motorbike can utilize those acquired skill sets to learn (transferred knowledge) how to drive a vehicle. In a seminal paper, Pan and Yang (2010), categorized TL under three broad settings:

Inductive transfer learning: In this TL setting, the source and target domain can be the same or different but the *source* (task) and *target* (task) is considered different regardless of the domain. Additionally, the two variants of inductive transfer learning include multi-task and self-taught learning. On one hand, in multi-task learning, the goal is to learn the features from both the target and the source tasks simultaneously. On the other hand, self-taught learning is analogous to inductive learning with the exception where the labeled data in the source domain are absent.

Transductive transfer learning: In this approach, the source and target domains differ while the source and target tasks are the same. In other words, the source domain may contain large amounts of labeled data while labeled data in the target domain are unavailable. Other special variants of this approach have been identified in the literature to include domain adaption, co-variant shift, and sample selection bias (Patel et al., 2015a).

Unsupervised transfer learning: In unsupervised TL, the source and target domains have no labels for training. However, it is much like inductive TL, where the target task can be dissimilar from but comparable with the source task. Table 2 summarizes the different TL settings available for various deep learning scenarios (Pan & Yang, 2010).

Table 2. Different settings of transfer learning

Transfer Learning Settings	Related Areas	Source Domain Labels	Target Domain Labels	Tasks
Inductive Transfer Learning	Multi-task learning	Available	Available	Regression, Classification
	Self-taught learning	Unavailable	Available	Regression, Classification
Transductive Transfer Learning	Domain Adaptation, Sample Selection Bias, Co-variate Shift	Available	Unavailable	Regression, Classification
Unsupervised Transfer Learning		Unavailable	Unavailable	Clustering, Dimensionality Reduction

2.6 Medical imaging modalities

The goal of medical imaging is to enable correct diagnosis, therapy, and treatment of patients centered around the managing of the disease protocols while mitigating any harmful side effects from the diagnostic procedure. In the healthcare domain, hospitals rely on different medical imaging modalities such as X-rays, to institute preventative care, improve diagnostics, and patient outcomes. With medical imaging, medical practitioners can use several scanning methods to visualize and monitor the activities in the human body without the need for invasive procedures, for example, surgeries for diagnostic and treatment purposes. Each of the medical imaging modalities is functionally specialized to produce digital images with varying degrees of information. In the medical domain, some of the common types of medical imaging include X-rays, ultrasound, Magnetic Resonance Imaging (MRI), microscopy, fundus, Computed Tomography (CT), Optical Computed Tomography (OCT) (Elangovan & Jeyaseelan, 2016; Litjens et al., 2017). Other imaging modalities are organ-specific, for example, retinal photographs and dermoscopy (Ker et al., 2018). Figure 3 shows a graphical representation and includes brief notes on the types of medical imaging modalities (Speicher, 2019).

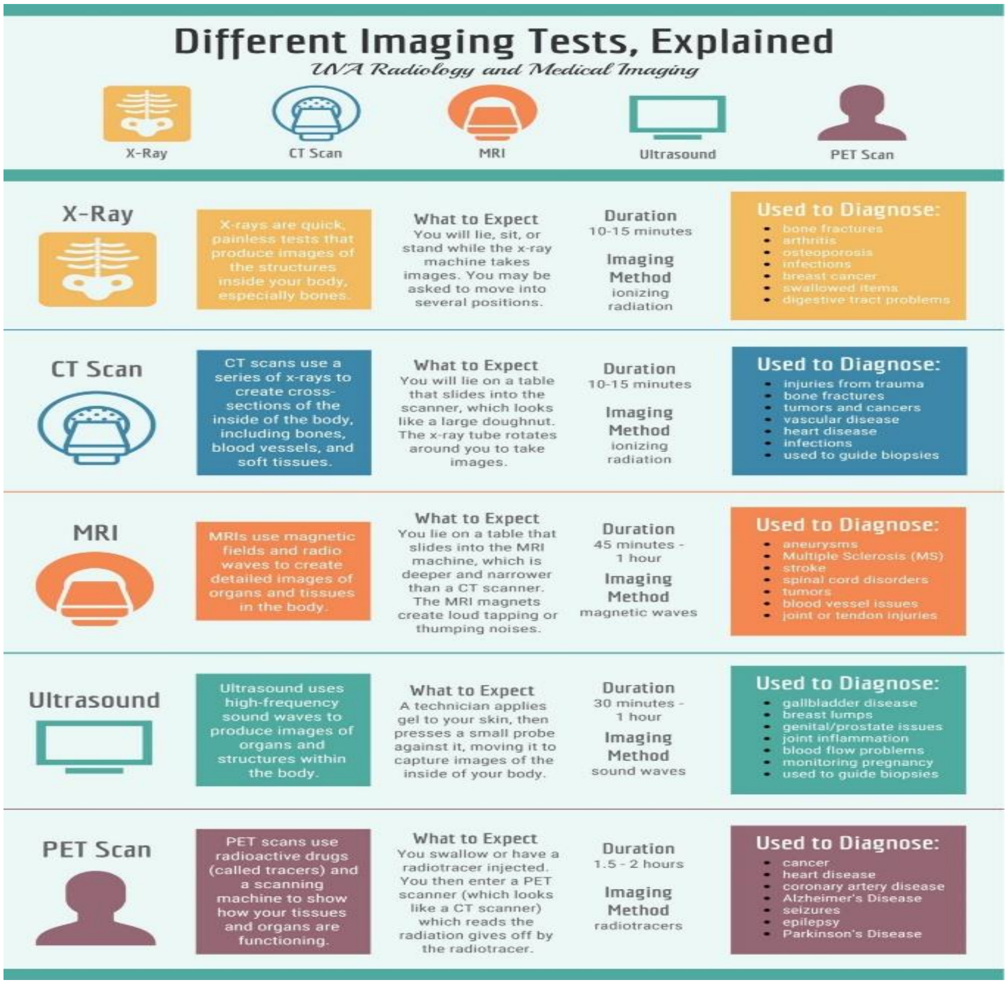


Figure 3. Types of common medical imaging modalities

2.7 Systems of the human body

The human body is made up of several organs working together towards a common objective known as the human body system. The human body can be classified into several broader categories (Kerrigan, 2020):

1. Endocrine: deals with regulation of the body.
2. Nervous: responsible for all sensory activities e.g. brain.
3. Respiratory: the organs here ensure gaseous exchange in the in and out of the body e.g. lungs.
4. Cardiovascular/Circulatory: ensures the circulation of blood in the body e.g. heart.
5. Digestive and Excretory: provides a system for food absorption and excretion of wastes from the body e.g. gastrointestinal tract.

6. Lymphatic and immune system: protects the body from harmful substances e.g. lymph.
7. Reproductive: the organ that allows the production of offspring.
8. Urinary and renal system: includes the kidneys responsible for filtering blood and removal of wastes from the body.
9. Integumentary: parts of the boy that ensure regulation from temperature e.g. skin, hair.
10. Skeletal: organs that maintain the body structure e.g. bones.
11. Muscular: allows the body to move using muscles.

A diagrammatic representation of the human body systems is presented in Figure 4.

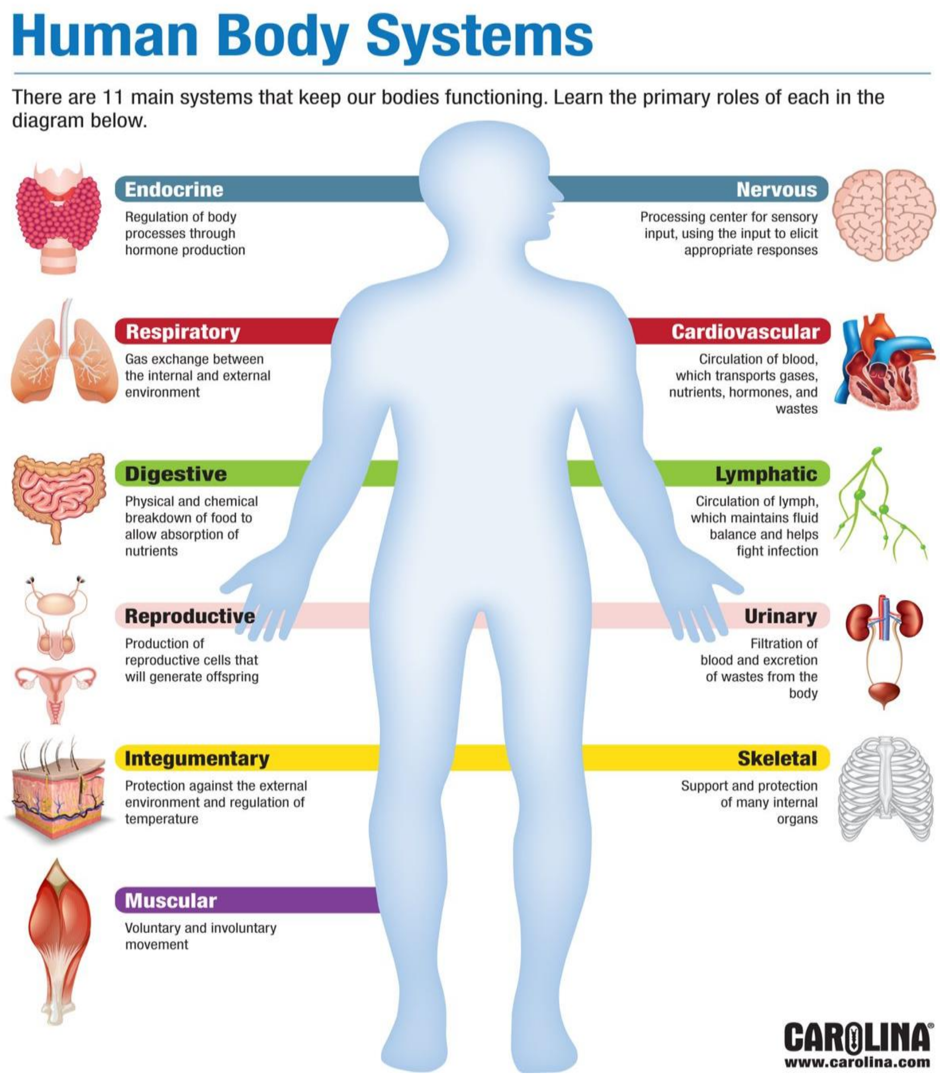


Figure 4. The Human body system

After reviewing the relevant literature, we propose a framework to conceptualize and organize the analysis and results from the systematic literature review as shown in Figure 5. This framework connects several concepts identified and developed in Figure 2.

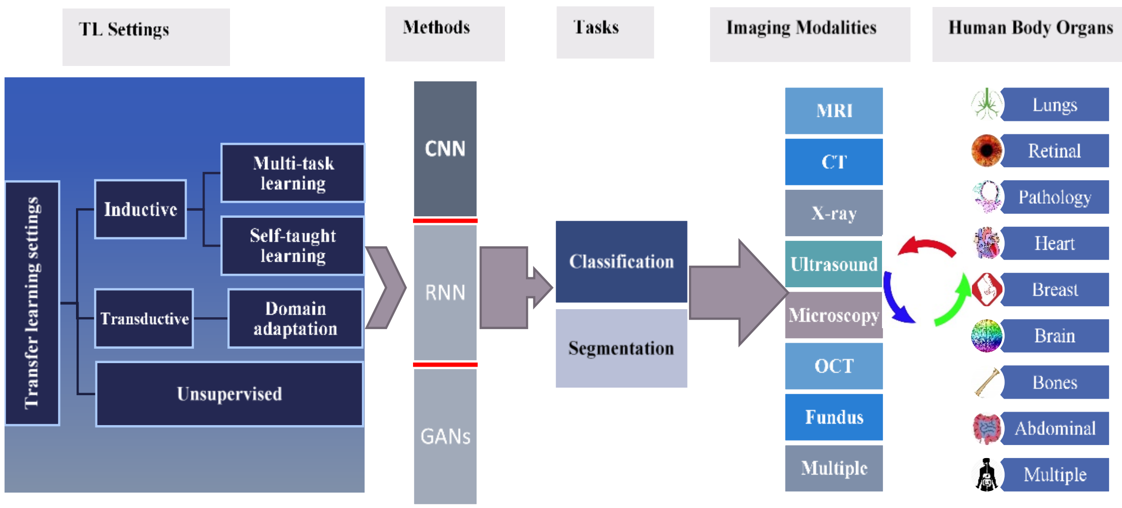


Figure 5. Framework for organizing the review articles

To our knowledge, no prior studies have examined transfer learning settings, strategies, and techniques specifically focused on medical images using the PRISMA methodology and to evaluate their impact on current trends to identify future research opportunities. A new survey is, therefore, needed to address the research gap revealed so far in the literature. This promising opportunity, therefore, inspires our research efforts to present an up-to-date comprehensive systematic literature review, following the PRISMA methodology, specifically on the application of transfer learning settings and techniques on medical images covering the visual recognition tasks such as classification and segmentation.

3.0 RESEARCH METHODOLOGY

3.1 Overview

For the methodology, we used the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) principles and guidelines to complete the literature review process (Moher et al., 2009). The research topic reviewed focused on the transfer learning strategies and settings applied to medical image recognition problems. In addressing the research questions highlighted earlier in the general introduction, we conducted a systematic search of articles from targeted scientific and journal online databases that included PubMed, Web of Science, IEEE Xplore Digital Library (IEEE), and Arxiv databases based on their relevance to the research questions and domain topic. The survey topic is relatively nascent, and therefore, the scope of the search period was limited to the time frame between April 1, 2010, to April 2020, a 10-year timeline that was deemed representative of the research topic under investigation.

3.2 Search Strategy

The literature search was based on the keywords "*transfer learning*", "*medic**" and "*image**". The key phrases were concatenated using Boolean expressions and applied to search through the selected online databases yielding a total of 223 articles. The online databases were selected based on the relevancy of content with the research title, research questions, and the domain application. The literature review search, process, and classification were carefully conducted guided by two defined measures; 1.) Inclusion criteria; and 2.) Exclusion criteria. These measures define the criteria for searching and extracting relevant articles relating to the research title and domain of interest. They include:

Inclusion criteria:

- Article presenting the application of transfer learning techniques or strategies and discussing their application on medical images.
- Articles that use neural networks such as CNN to implement transfer learning approaches.

Exclusion Criteria:

- Articles discussing the use of transfer learning in other application domains other than medical images or other aspects of transfer learning regardless of the application domain.
- Articles that focus on image captioning, annotations, concepts, and topic extraction.
- Articles that focus on literature review or survey.
- Articles that mentioned the detection of surgical devices or tools.
- Masters, doctoral dissertations, and unpublished works are excluded.
- All articles other than the English language are excluded.

4.0 RESULTS

4.1 Overview

The literature review process involved several important steps to extract the relevant articles as shown in Figure 6., resulting in a total of 223 articles. First, duplicate articles were carefully removed through manual and automatic methods. In some cases, where multiple publications were observed to have the same authors, for example, in both a conference and a journal publication, the journal entry was included for further analysis. Second, the articles were screened for the title and abstract relevancy guided by the inclusion and exclusion criteria. Third, the articles were further screened for full-text availability using both the inclusion and exclusion criteria. Lastly, the results were included for further analysis and organized to provide answers to address the research questions formulated earlier in the introduction section. When in doubt about information on an article's title or abstract, the articles were carefully read in full text to extract the exact parameters for the conceptual classification framework. The organization of the extracted information was based on the proposed classification framework shown in Figure 5. Moreover, Figures 7-10 shows the numeric distribution of articles according to the publication period, imaging modalities, imaging analysis task, and application areas in the human body.

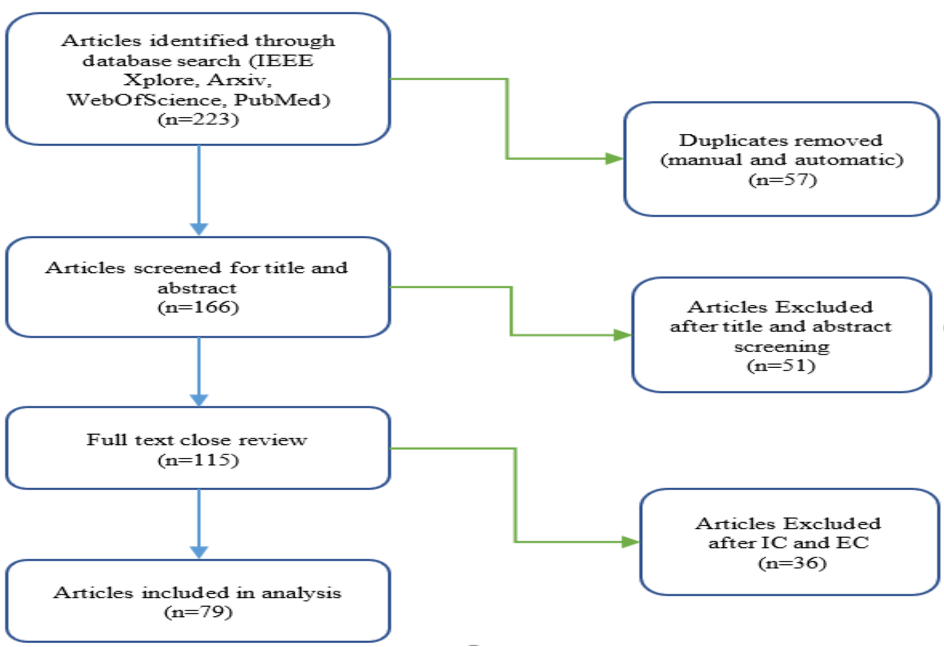


Figure 6. Flow diagram of the review process using PRISMA

4.2 Analysis of results

The vast majority of the reviewed articles were extracted from the journal articles (75%) followed by conference papers (25%) spanning the period 2015-2020 as shown in Figure 9. In Figure 7, the reviewed articles are classified according to the area of application regarding the human body organs. Lungs and abdominal areas of the body have the highest attention in terms of research publications. Additionally, Tables 3-13 presents a detailed breakdown of the classification of peer-reviewed articles according to the proposed conceptual framework (see Figure 5). The organization into tables helps to reveal resultant patterns and insights for further discussion. Unless stated otherwise, it is worth noting that most of the articles summarized in Tables 3-13 had their methods as classified as CNNs and TL approach as inductive transfer learning.

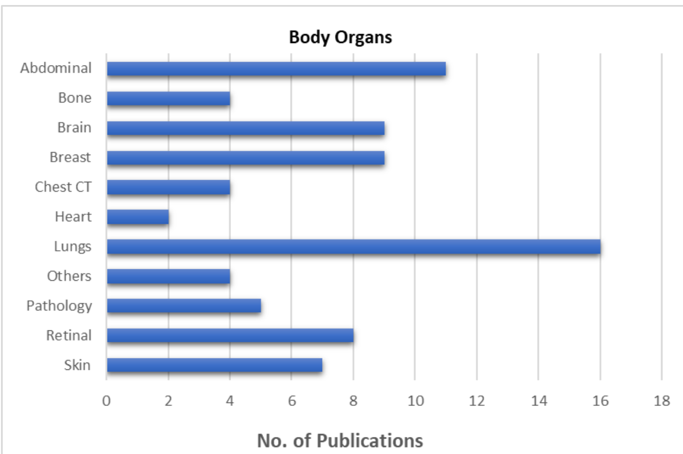


Figure 7. Distribution by body organs

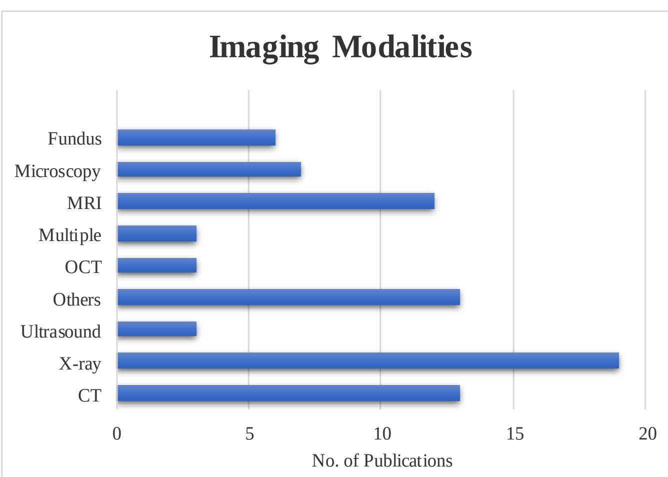


Figure 8. Distribution by imaging modalities

Looking at the distribution of publication by year, it is noteworthy that the growing appeal in transfer learning on medical images can be evidenced from the increased frequency in scientific publications of conference and journal papers since 2015 as shown in Figure 9. This observation can largely be attributed to the huge amounts of labeled datasets and computational power required to train deep neural networks. However, we argue that the limited number of relevant articles gathered between 2015 to 2018 reflects the relative newness of the transfer learning concept and gradual evolution in the application on the medical imaging domain.

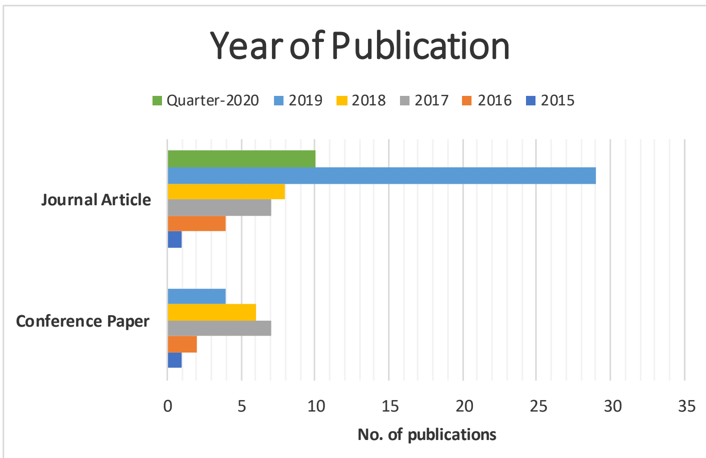


Figure 9. Distribution by year of publication

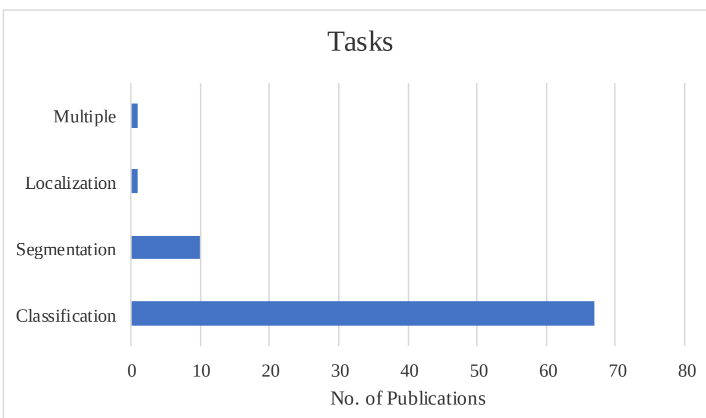


Figure 10. Distribution by image analysis task

For the image analysis tasks, the results in Figure 10 reveal that most of the medical imaging tasks that use transfer learning approaches are classification followed by segmentation. It is interesting to note that classification is one of the most popular and heavily used tasks in image recognition, that is, identifying what object or anomaly appears in an image. Similarly, a word cloud was created to give insights on how transfer learning systems are applied to which specific systems of human anatomy (see the outer layer in Figure 2.). All the relevant articles were classified into the systems of human anatomy to reveal the resultant insights (Rettner, 2016). In Figure 11, the word cloud reveals that the top three systems of the human anatomy that receive high research attention are Nervous, Respiratory, and Integumentary systems. However, immune, Circulatory, Renal, and Endocrine systems remain under-examined, perhaps due to the unavailability or limited size of labeled datasets.

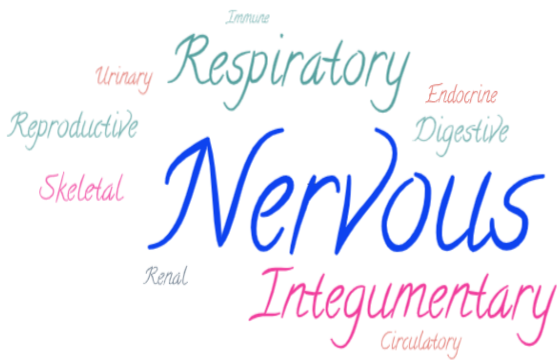


Figure 11. Word cloud representing systems of the human body

4.3 Distribution of articles by Abdomen image analysis

Out of the total reviewed articles, the abdomen contributed about 12.7% of the total relevant articles. The majority of articles reviewed aimed to classify and segment different organs found in the abdomen region such as kidney, liver, and lymph nodes as shown in Table 3. The imaging modalities were fairly a mixed bag, for example, microscopy, CT, ultrasound, and staining with none dominating the other. Given the popularity and success of U-Net (Ronneberger et al., 2015) on biomedical image segmentation, it was interesting to note that Motamed et al. (2019), trained U-Net architecture to segment prostate glands. Other articles that focused on segmentation tasks, included, Sital et al., (2020) where they used autoencoders with both labeled and unlabeled datasets in identifying 3D liver images from CT scans. The remaining articles were on classification tasks of different body organs.

Some abbreviations used in the tables include PA: Publication Article; JA: Journal Article; H&E: hematoxylin and eosin; Optical coherence tomography (OCT).

Table 3. Overview of papers using TL for abdominal image analysis.

Author	PA	TL setting	Method	Architecture	Task	Imaging Modality	Comments
Agrawal et al.,(2019)	JA	Inductive	CNN	VGGNet;ResNet50;Inception-V3;XceptionNet;MobileNet	Classification	Others	Evaluating CNN performance on low resource medical images(e.g. gastrointestinal images)
Ayyar et al.,(2018)	CP	Inductive	CNN	ResNet50;InceptionV3; InceptionResNet; VGG19	Classification	Others (staining)	Classification of abnormal glomeruli in the renal tissues
Chen et al.,(2019)	JA	Inductive	CNN	InceptionV3; VGG16	Classification	MRI	Detection of prostate lesions from multiparametric MRI images
Kang and Gwak,(2019)	JA	Inductive	CNN	Ensemble of Mask R-CNN; ResNet50 and ResNet101 as backbone	Segmentation	Others	The first study to use Mask-RNN for the task of Polyp instance segmentation
Lan et al.,(2019)	JA	Inductive	CNN	Cascade proposal with ZF network as a baseline and Fast R-CNN with VGG-16	Classification	Others	Wireless capsule endoscopy abnormality detection
Ma and Peng,(2019)	JA	Inductive	CNN	Custom CNN	Classification	CT	Use of multi-source transfer learning for lymph node detection
Motamed et al.,(2019)	JA	Inductive	CNN	U-net	Segmentation	MRI	Segmentation of prostate glands and transition zones
Nadimi et al.,(2020)	JA	Inductive	CNN	ZF-Net; Faster R-RNN	Classification	Others	Detection and localization of colorectal polyps
Ravishankar et al.,(2017)	JA	Inductive	CNN	Custom CNN	Classification	Ultrasound	Investigates the effectiveness of transfer learning to kidney detection
Sital et al.,(2020)	JA	Inductive	CNN	Auto-Encoder(AE)	Segmentation	CT	Segmentation task on 3D liver CT images
Sun et al.	JA	Inductive	CNN	ResNet50	Classification	Microscopy	Detection of liver cancer histopathological images

4.4 Distribution of articles by brain image analysis

The brain constitutes an important part of the nervous system of the human anatomy as illustrated in the last layer of the conceptual framework in Figure 2. As shown in Table 4, we identified a large number of articles addressing classification tasks with many of the imaging modality being MRI. Multimodal and fusion representation of the brain in 2D and 3D targeting diagnosis of Alzheimer's and stroke lesions. Most papers applied classification tasks except for Amin et al.,(2019) and Malla et. al,(2019) where they diagnosed the presence of ischemic stroke lesions from MRI images.

Table 4. Overview of papers using TL for brain image analysis

Author	PA	TL setting	Method	Architecture	Task	Imaging Modality	Comments
Amin et. al.,(2019)	JA	Inductive	CNN	AlexNet and GoogLeNet	Segmentation	MRI	Detection of multimodal Brain tumor and ischemic stroke lesion via segmentation tasks
Awate, (2019)	JA	Inductive	CNN	Inception V3	Classification	MRI	Detection of Alzheimer's disease from MRI images
Dawud et al.,(2019)	JA	Inductive	CNN	AlexNet-SVM	Classification	CT	Identification of brain hemorrhage from CT images
Hermessi et al.,(2018)	JA	Inductive	CNN	Custom CNN	Classification	CT and MRI	Developed a multi-modal fusion method for fusion and classification in the shearlet domain
Lao et al.,(2017)	JA	Inductive	CNN	CNN-S	Classification	MRI	Developed a radiomic model for prediction of patients with Glioblastoma multiforme (GBM)
Liang et al.,(2018)	JA	Inductive	CNN	M3D-DenseNet	Classification	MRI	Prediction of isocitrate dehydrogenase (IDH) genotype using a novel M3D DenseNet on multimodal MRI images
Malla et. al.,(2019)	JA	Inductive	CNN	Custom CNN	Segmentation	MRI	Developed DeepMedic method to identify and segment ischaemic stroke lesions from MR images
Wingate et al.,(2019)	JA	Transductive (Domain adaptation)	CNN-RNN	ResNet-50	Classification	MRI	Use of DNN's to predict Parkinson's disease from MRI images
Zhang et al.,(2017)	JA	Inductive (multi-task)	CNN	AlexNet	Classification	MRI	Prediction of Alzheimer's disease using a multi-task learning strategy

4.5 Distribution of articles by bone image analysis

Bones can be categorized under the skeletal system of the human anatomy. The DCNNs have been applied to the bone images for different tasks such as segmentation, classification, and localization. The findings are summarized in Table 5. An interesting study by Feng et al., (2020) combined CNN and RNN (LSTM) to detect the presence of rheumatoid arthritis from CT images. As expected, the preferred imaging modality for diagnosis of bone-related diseases was X-ray and CT and illustrated in Figure 8 as well.

Table 5. Overview of papers using TL for bone image analysis

Author	PA	TL setting	Method	Architecture	Task	Imaging Modality	Comments
Feng e. al.,(2020)	JA	Inductive	CNN; RNN (LSTM)	ResNet; Highway Networks	Classification	CT	Detection of rheumatoid arthritis based on diffuse optical tomography
Jodeiri et. al.,(2019)	CP	Inductive (multi-task)	CNN	Mask R-NN; U-Net	Segmentation	X-ray	Segmentation of pelvic radiographs using multi-task learning
Tiulpin et al.,(2019)	JA	Inductive	CNN	Hour-glass CNN	Localization	X-ray	Localization of knee landmarks using hourglass networks
Zhou et al.,(2017)	CA	Inductive	CNN	VGGNET	Classification	X-ray	Automatic bone age assessment based on the region of interests (ROI)

4.5 Distribution of articles by breast image analysis

Breast cancer is a type of cancer mostly afflicting women and manifests itself in breast tissues. Mammograms are the most common type of modality for detecting breast cancer via X-rays. Other imaging modalities like ultrasound, MRI, and microscopy are equally important for examining signs of cancer from breast tissues. In Table 6, we summarize the findings of several studies from the literature on breast image analysis. All the articles focused on image classification with one paper by Samala et al., (2017) using multi-task transfer learning to classify breast cancer. Although ultrasound is considered the most safest and cost-effective imaging modality, only one study employed ultrasound images (Byra et al., 2018). Regarding the architecture used, VGGNet variants were the most popular for the diagnosis of breast cancer with a specific case where VGG19 (Hu et al., 2019) was used to extract 4D information from MRI breast images.

Table 6. Overview of papers using TL for breast image analysis

Author	PA	TL setting	Method	Architecture	Task	Imaging Modality	Comments
Byra et al.,(2018)	JA	Inductive	CNN	Inception V3 and VGG19	Classification	Ultrasound	Impact of image reconstruction on breast lesion classification using transfer learning
Chang et al., (2017)	CP	Inductive	CNN	Inception V3	Classification	Microscopy	Detection of breast cancer using histopathology images
Guan and Loew, (2017)	CP	Inductive	CNN	VGG16	Classification	X-ray	Detection of breast cancer
Hadad et al., (2017)	CP	Inductive	CNN	VGG128	Classification	MRI	Detection of breast lesion
Hu et al.,(2019)	JA	Inductive	CNN	VGG19	Classification	MRI	Breast cancer diagnosis using 4D information
Huynh et al., (2016)	JA	Inductive	CNN	AlexNet	Classification	X-ray	Detection of breast lesions using 5-fold cross-validation
Samala et al.,(2017)	JA	Inductive (multi-task)	CNN	ImageNet DCNN	Classification	Others	Breast cancer diagnosis using multi-task transfer
Suzuki et al., (2016)	CP	Inductive	CNN	AlexNet	Classification	X-ray	Mass detection on mammographic images
Valerio et al., (2019)	CP	Inductive	CNN	Inception V3;Inception-ResNet-v2;NASNet-Large	Classification	X-ray	Detection of Breast lesion

4.6 Distribution of articles by chest image analysis

Chest radiography, also known as a Chest X-ray, is considered as one of the most common forms of radiological examination used to diagnose thoracic pathologies. For many years, X-rays and CT scans have been the standard imaging modality for detecting different chest conditions by radiologists. In recent years, two large scale datasets have been made publicly available: (1) ChestX-Ray-14 (Wang et al., 2017) and (2) CheXpert (Irvin et al., 2019) for medical image classification tasks. Many types of different deep learning architectures have been applied to these large Chest X-Ray datasets for both binary and multiclass classification with varying success. Recently, Choudhary and Hazra (2019) used a pretrained ImageNet model of VGG-16 (Krizhevsky et al., 2012) on the ChestX-ray14 dataset. They achieved state-of-the-art training accuracy of 98% and test accuracy of 97% on a binary classification task using transfer learning approaches by fine-tuning the higher layers of the pretrained model. An interesting observation was from a study by Oliveira and dose Santos (2018) who used unsupervised learning by training a variant of U-Net to segment and translate Chest X-ray images. We expect to see more research exploiting transfer learning approaches for both classification and segmentation tasks on Chest X-ray images.

Table 7. Overview of papers using TL for Chest CT image analysis

Author	PA	TL Setting	Method	Architecture	Task	Imaging Modality	Comments/Detection
Alsabahi et al.,(2018)	CP	Inductive	CNN	Inception V3	Classification	X-ray	Detection of abnormal or normal digital radiographic chest images
Choudhary and Hazra (2019)	JA	Inductive	CNN	VGG16	Classification	X-ray	Multi-classification of Chest X-Ray 14 dataset using transfer learning strategies
Feng et al., (2020)	JA	Inductive	CNN	3D U-Net	Segmentation	CT	Classification of Organs at Risk (OAR) from thoracic images
Oliveira and dos Santos (2018)	JA	Unsupervised	CNN	FCN;U-nets; SegNets;MUNIT	Segmentation	X-ray	Use of DNNs for Semantic segmentation and image translation of Chest X-ray images

4.7 Distribution of articles by heart image analysis

Transfer learning approaches have been applied to the heart or cardiac image analysis as shown in Table 8. Despite the high rate of global deaths resulting from heart-related diseases, for example, ischaemic heart disease (see Figure 12.), few articles mentioned the use of transfer learning techniques on cardiac image analysis. Imaging modalities were varied from CT to Ultrasound, but the imaging task was reported as classification. Gupta et al., (2019) used Inception v3 architecture to investigate the impact of transforming 3D to 2D coronary CT images applying transfer learning with data augmentation techniques. We predict to see more research targeting cardiac image analysis as more large datasets of heart-related diseases become publicly available to researchers in the computer vision and medical domain.

Table 8. Overview of papers using TL for heart image analysis

Author	PA	TL setting	Method	Architecture	Task	Imaging Modality	Comments
Gupta et al.,(2019)	JA	Inductive	CNN	Inception V3	Classification	CT	Impact of 3D-to-2D transformation using TL and data augmentation on small datasets (coronary CT images)
Miyagawa et al.,(2019)	JA	Inductive	CNN	Custom CNN	Classification	Ultrasound	Detecting vascular bifurcation on coronary CT images

4.8 Distribution of articles by lung image analysis

This is the category with the largest amount of reported studies related to lungs where transfer learning techniques were used for classification and segmentation tasks. The results from this category had the most balanced and varied applications regarding the TL settings and architectures utilized for tuberculosis diagnosis. The findings are summarized in Table 9. Although the unsupervised TL approach had the lowest reported application from all the results, it is interesting to note that three papers employed unsupervised learning for classification and segmentation tasks. Chen et al., (2018) proposed a semantic-aware generative adversarial network for unsupervised domain adaptation (*SeUDA*) utilizing a modified ResNe-t101 and CycleGAN which achieved a dice coefficient of 93.42% on lung segmentation task. Similarly, Sawada and kozuka, (2015) combined unsupervised transfer learning with a multi-prediction deep Boltzmann machine (MPDBM) on two datasets: (1) MNIST dataset (*source domain*) and (2) X-ray dataset (*target domain*) to classify lung tissue with an improved classification performance of 99.6%.

Moreover, Hussein et al., (2019) proposed a novel unsupervised learning algorithm combining multi-task learning with 3D CNNs achieving a lung nodule classification accuracy of 78.06%. Currently, the Coronavirus disease (COVID-19) declared as a pandemic by World Health Organization,(2020) targeting the respiratory systems of human beings, has ravaged almost every continent with hundreds of thousands of fatalities recorded to date. There is no known cure for COVID-19, however, research efforts are on-going to diagnose and find treatment for the disease. In the medical imaging field, these research efforts include the use of ResNet-18 and VGG-19 together with other DCNNs to diagnose COVID-19 from Chest X-ray datasets (Abbas et al., 2020; Apostolopoulos & Mpesiana, 2020). Remarkably, a large number of papers are devoted to addressing respiratory diseases, such as lung-related conditions. This is evidenced by the global statistics depicted in Figure 12.

Table 9. Overview of papers using TL for lung image analysis

Author	PA	TL Setting	Method	Architecture	Task	Imaging modality	Comments
Abbas et al.,(2020)	JA	Inductive	CNN	ResNet18	Classification	X-ray	Detection of COVID-19 from chest X-ray images
Ahsan et. al, (2019)	JA	Inductive	CNN	VGG16	Classification	X-ray	Detection of Tuberculosis from Chest x-ray dataset
Apostolopoulos et. al.,(2020)	JA	Inductive	CNN	VGG19; MobileNet v2;Inception; Xception; Inception ResNet-v2	Classification	X-ray	Performance of several standard CNN architectures to diagnose COVID-19
Chen et al.,(2018)	JA	Transductive (Unsupervised Domain Adaptation)	GAN	CycleGAN; ResNet101	Segmentation	X-ray	Unsupervised domain adaptation using CycleGAN to segment Chest X-ray images (Montgomery dataset as the source domain and JSRT dataset as target domain)
Christodoulidis et al., (2017)	JA	Inductive (multi-task)	CNN; RNN	Custom CNN	Classification	CT	Developed a method that improved accuracy and stability of lung tissue pattern characterization
Han et al.,(2020)	JA	Inductive	CNN	Xception; VGG16; Mask R-CNN	Classification and segmentation	CT	Classification and segmentation of lung and Hemorrhagic stroke from CT images
Hussein et al., (2019)	JA	Unsupervised (multi-task learning)	CNN	3D CNN	Classification	CT	Detection of Lung nodules and Pancreatic cysts using a novel unsupervised algorithm
Hwang and Kim,(2016)	JA	Inductive (multi-task)	CNN	Custom CNN	Classification	X-ray	Tuberculosis detection and localization via heatmaps
Nóbrega et al., (2018)	CP	Inductive	CNN	ResNet50	Classification	CT	Classification of lung nodules from CT images
O'quinn et al., (2019)	CP	Inductive	CNN	AlexNet	Classification	X-ray	Diagnosis of Pneumonia from Chest radiographs
Paul et al., (2016)	JA	Inductive	CNN	Vgg-f;vgg-m; vgg-s	Classification	CT	Prediction of patient's survival with lung Adenocarcinoma using deep transfer learning
Sawada and Kozuka,(2015)	CP	Unsupervised	GAN	Deep Boltzmann Machines (DMB)	Classification	X-ray and CT	Classification of lung tissues using multi-prediction deep Boltzmann machines.
Suzuki et al., (2018)	CP	Inductive	CNN	AlexNet	Classification	X-ray	Used two-stage transfer learning to classify diffuse lung diseases (DLDs)
Wang et al., (2018)	JA	Inductive	CNN; Auto-Encoder	VGGNet16	Classification	X-ray	Detection of lung lesions from multiple chest X-ray datasets using a novel multiple metric indexes.
Xiong et al., (2019)	JA	Inductive	CNN	ResNet101	Classification	CT	Identification of EGFR mutation status in patients with lung adenocarcinoma
Zhang et al., (2019)	JA	Inductive	CNN	LeNet-5; AlexNet	Classification	CT	Using transfer learning to classify pulmonary nodules from CT images

4.9 Distribution of articles by pathology image analysis

Deep learning techniques specifically transfer learning strategies have been applied for classification and segmentation of pathology image analysis. The most common imaging modality in this category is microscopy and hematoxylin and eosin (H&E) staining. In Liang et al., (2018) a CNN combined with RNN was presented with the best model performance from Xception-LSTM resulting in a classification accuracy of 90.79% when used for blood cell image classification. Alternatively, Gaur et al., (2016) performed membrane segmentation using transductive TL demonstrating its effectiveness in handling domains with scarce labeled datasets. More recent studies have continued to employ CNN architectures for the classification of blood cells. Table 10 presents a summary of each of the relevant articles in this category.

Table 10. Overview of papers using TL for pathology image analysis

Author	PA	TL Setting	Method	Architecture	Task	Imaging Modality	Data	Comments
Gaur et al.,(2016)	CP	Transductive	CNN	ConvNet	Segmentation	Microscopy	Stained data	Membrane segmentation using active learning-based feature transfer
Liang et al.,(2018)	JA	Inductive	CNN-RNN	Xception-LSTM; ResNet50-LSTM; InceptionV3; InceptionV3-LSTM Xception-ResNet50-LSTM;	Classification	Microscopy	Blood cells	Classification of blood cell images
Ponzio et al.,(2019)	JA	Inductive	CNN	VGG-16	Classification	Microscopy (H&E staining)	Stain data	Classification of histopathological images
Sun and Binder, (2017)	CP	Inductive	CNN	Caffenet; GoogLeNet; ResNet50	Classification	Microscopy (H&E staining)	Stain data	Transfer learning on H&E stained histopathology images
Talo, (2019)	JA	Inductive	CNN	ResNet50; DenseNet161	Classification	Microscopy (H&E staining)	Stain data	Classification of histopathology images into 24 classes via transfer learning

4.10 Distribution of articles by retinal image analysis

Retinal disease detection is an emergent domain where DNNs have recently found success in diagnosing retinal diseases from fundus images. Apart from Sekou et al., (2019) who developed a framework for retinal image segmentation and performed experiments using U-Net on retinal images, all the other papers focused on classification tasks. Most recently, a new type of imaging modality for evaluating retinal disorders known as Optical coherence tomography (OCT), has emerged allowing the application of DCNNs to classify the high-resolution images for macular degeneration disorders (Karri et al., 2017; Motozawa et al., 2019). As summarized in Table 11, all the papers used CNNs for the diagnosis of retinal diseases and conditions.

Table 11. Overview of papers using TL for retinal image analysis

Author	PA	TL setting	Method	Architecture used	Task	Imaging Modality	Comments
Choi et al.,(2017)	JA	Inductive	CNN	VGG19;MatConvnet;AlexNet	Classification	fundus	Multi-categorical classification of fundus images
Hagos and Kant,(2019)	JA	Inductive	CNN	Inception V3	Classification	fundus	Detection of diabetic retinopathy from fundus images
Karri et al., (2017)	JA	Inductive	CNN	GoogLeNet	Classification	OCT	Identification of pathologies on diabetic macular edema and macular degeneration images
Li et al.,(2017)	CP	Inductive	CNN	VGG-m (128,1024,2048); GoogLeNet; AlexNet	Classification	fundus	Impact of transfer learning using small datasets for fundus classification tasks
Motozawa et al.,(2019)	JA	Inductive	CNN	Custom CNN	Classification	OCT	Detection of age-related macular degeneration from healthy OCT images
Raghu et al., (2019)	JA	Inductive	CNN	ResNet50 and Inception-v3	Classification	fundus	Effectiveness of standard and pretrained architectures on medical imaging tasks
Sekou et al., (2019)	JA	Inductive	CNN	FCNN; U-Net	Segmentation	fundus	Classification performance on retinal image segmentation
Xu et al.,(2018)	CP	Inductive	CNN	DenseNet	Classification	fundus	Performance classification of fundus images

4.12 Distribution of articles by skin image analysis

Diagnosis of skin cancer is another application area that is attracting attention based on recently published articles where CNNs are increasingly used to detect different forms of skin cancer from dermoscopic images. Melanoma is the most common form of skin cancer and can be fatal if left untreated (Burdick et al., 2018). A recent work by Serte and Demirel, (2020), demonstrated the performance of ResNet-18 and ResNet-50 on dermoscopic images to classify malignant melanoma and seborrheic keratosis achieving the state-of-the-art sensitivity value of 78.66%. Another study by Burdick, (2018) focused on the segmentation of skin cancer cells to detecting Melanoma showing improved metric results when combined with the transfer learning paradigm. The remaining papers emphasized classification tasks of detecting the presence of Melanoma and the imaging modalities were all reported as digital photographs except for a study by Lopez et al., (2017) who used X-ray images to diagnose skin lesions. Table 12 summarizes the findings from the literature.

Table 12. Overview of papers using TL for skin image analysis

Author	PA	TL setting	Methods	Architecture	Task	Imaging Modality	Comments
Burdick et al.,(2018)	JA	Inductive	CNN	VGG16; Inception V3	Segmentation	Others (Total body photography	Segmentation of skin cancer cells to diagnose Melanoma
Lopez et al.,(2017)	CP	Inductive	CNN	VGGNet16	Classification	X-ray	Diagnosis of skin lesions through classification
Mahbod et al., (2020)	JA	Inductive	CNN	EfficientNetB0; EfficientNetB1; SeReNeXt-50	Classification	Others (Clinical Photographs)	Effect of image size for skin lesion classification
Menegola et al.,(2016)	JA	Inductive	CNN	VGG-M + SVM	Classification	Others (digital photographs)	Detection of Melanoma from skin lesion images
Sachdev et al.,(2018)	CP	Inductive	CNN	VGG19:RESNet50; Inception V3	Classification	Others(Clinical Photographs)	Classification of skin lesions using an Android application
Serte and Demirel ,(2020)	JA	Inductive	CNN	ResNet-18;ResNet-50	Classification	Others (Clinical Photographs)	Classification of malignant melanoma and seborrheic keratosis from skin images
Wu et al.,(2019)	JA	Inductive	CNN	ResNet-50; Inception-v3; DenseNet121; Xception and Inception-ResNet-v2	Classification	Others (Digital Photographs)	Diagnosis of six common facial skin diseases

4.13 Distribution of articles by others

This is the last section where the articles addressed human anatomical regions ranging from reproductive to hearing organs. Table 13 gives a summary of the studies covering the different anatomical districts of the human body system. All the articles addressed image classification tasks with AlexNet being the CNN of choice to experiment with the effectiveness of using the inductive transfer learning approach.

Table 13. Overview of papers using TL for other medical image tasks

Author	PA	TL Setting	Method	Architecture	Task	Imaging Modality	Anatomical District	Comments/Detection
Hermessi et. al.,(2019)	JA	Inductive	CNN	AlexNet	Classification	MRI	Uterus	Classification of liposarcoma and leiomyosarcoma through the transfer learning approach.
Kudva et al.,(2019)	JA	Inductive	CNN	AlexNet and VGGNet -16	Classification	CT/MRI	Cervix	Classification of Uterine Cervix Images for cervical cancer using hybrid transfer learning
Patrini et al, (2020)	JA	Inductive	CNN	VGG16;Inception v4; ResNet V1 101; ResNet V1 152; ResNet V2 152; Inception - ResNet V2	Classification	Others [Narrow Band Imaging (NBI)]	Throat	Multi-class classification using transfer learning with SVM in selecting of laryngoscopic features (laryngeal tract)
Shie et al., (2015)	JA	Inductive	CNN	Variant of AlexNet	Classification	OCT	Ear	Employed transfer learning with SVM classifier to classify Otitis Media images (ear) using an unsupervised codebook construction from ImageNet images.

In summary, we analyzed all the deep learning architectures used in all of the articles we extracted for review and found that among the top five most used algorithms for transfer learning systems on medical images includes: (1) Inception-v3, (2) ResNet50, (3) AlexNet, (4) VGG-16 and (5) VGG-19. A complete detailed representation of all the deep learning architectures can be found in Appendix 1.

5.0 DISCUSSION

In this research, we provided a systematic literature review of transfer learning systems on medical images. Out of the 79 articles analyzed and reviewed in this survey, it is evident that the application of transfer learning systems on medical images is gaining attention in the medical domain even with the existing challenges of limited domain knowledge and scarcity of annotated images. The growth has been gradual over the last few years with the last two years contributing most of the articles which were published between the period 2018-2020. This work reveals that a majority of the applications target the nervous, respiratory, and integumentary systems of human anatomy.

Notably, these results are consistent with statistics from the World Health Organization (WHO) survey on top global causes of deaths that consists of lung cancer, tuberculosis, respiratory infections, and stroke among others (WHO, 2020a). However, an interesting observation is a paradoxical correlation between the findings in Table 8 and the statistical summary in Figure 12 about the heart disease. Whereas Ischaemic heart disease is the top global killer, limited research still exists in using transfer learning systems on cardiac image analysis. This discovery can be attributed twofold: First, the multi-dimensional nature of the brain which requires the preservation of much more information in 2D or 3D while training a deep neural network; Second, the limitation of large labeled datasets could conceivably explain the fewer publication. This phenomenon is also supported in literature where Litjens et. al, (2017) observed that most research papers focused on 2D, 3D, and 4D CNNs with exceptional cases where authors combined a CNN with an RNN to preserve more information during segmentation tasks. Moreover, CNNs remain the primary workhorse for most of the medical image analysis tasks with the classification task ahead on applications involving medical images. In other words, the application of RNNs and GANs is still limited further opening research questions for practice in employing transfer learning systems on medical images.

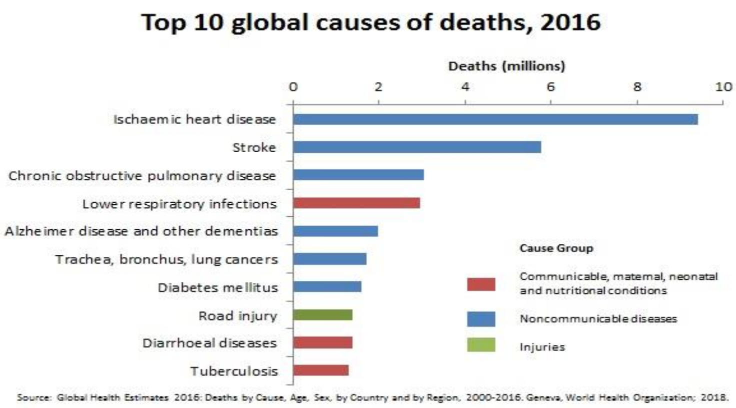


Figure 12 Top global causes of deaths

Unlike medical imaging scenarios that require training of models from scratch, results from the reviewed articles reveal that transfer learning approaches have significantly contributed to improved model performance in many medical imaging tasks across a variety of imaging modalities. Overall, we can state that currently, CNNs are the gold standard for many medical recognition problems and there exists plenty of playing ground for utilization of other algorithms such as recurrent convolutional neural networks (RCNNs) and GANs which can be a game-changer in providing alternative methods to data augmentations.

In transfer learning settings, the most common type of approach is inductive transfer learning, popular with the majority of the reviewed articles. Although widely interchanged with transfer learning in research texts, Inductive transfer learning as described earlier by (Pan & Yang, 2010), is a TL approach where the source and target tasks are different, domains can be similar or different but with the condition of having labeled data for both source and target domains. The popularity of the inductive transfer learning approach can be attributed to the compatibility and flexibility of most CNNs to allow for pretrained models and fine-tuning of different higher layers to improve model performance which offers an inexpensive strategy. However, this approach can be severely limited by the availability of labeled datasets and access to the powerful configuration of computational resources (GPUs) to train deep neural networks that can vary in width and depth of hidden layers.

Although transfer learning approaches on medical imaging have shown promising results across many medical imaging tasks and modalities, several untapped research opportunities remain for exploration. Looking at the findings summarized in the tables and observations of current trends in transfer learning systems on medical images, we can uncover insights regarding unsupervised transfer learning approaches which until recently remains an attractive research area that still has the potential to solve the problem of unlabeled datasets. Furthermore, the resurgence of RCNNs (Ming Liang & Xiaolin Hu, 2015), unsupervised techniques such as Variational Autoencoders (Doersch, 2016; Kingma & Welling, 2014), GANs (Goodfellow et al., 2014) coupled with domain adaptation could potentially extend the successes of CNNs to integrate and enhance contextual information which is relevant for most object recognition tasks. Most importantly, these unsupervised methods can be trained to learn internal and discriminative features from unlabeled data in both source and target domains, mitigating the scarcity of labeled datasets in the medical domain by leveraging the already existing huge amounts of unlabeled medical images. We also predict a gradual growth in the use of multi-instance learning (Kotzias et al., 2014), multi-task learning (Samala et al., 2017), multi-source learning (Christodoulidis et al., 2017) in combination with unsupervised domain adaptation.

Finally, transfer learning systems not only will have a great deal of research impact in medical image analysis but gain traction in the growth of sustained research publications in the medical imaging domain into the foreseeable future. Participatory involvement of Physicians and medical practitioners in the deployment of these deep learning models for medical imaging tasks will significantly benefit the complementary efforts of accurate, timely, and efficient diagnosis of diseases from mobile devices to cloud computing scenarios. The net effect is to advance the goals of public health organizations of enhancing diagnosis and offering effective treatment strategies to different population groups by leveraging Artificial Intelligence (AI) initiatives. Cutting edge deep learning systems will play a greater role in the healthcare industry by providing newer and efficient models that can learn multi-faceted and complex data types from heterogenous sources, for example, early diagnoses of Alzheimer’s disease.

6.0 CONCLUSION, LIMITATIONS AND FUTURE WORK

6.1 Conclusion

Deep learning, specifically TL approaches has shown great promise in the fast-changing landscape of medical imaging for efficient diagnosis and improved treatment strategies with demonstrated evidence of pervading almost every facet of medical image analysis. Throughout this research, we examined the extent to which transfer learning systems have been applied to medical images through a comprehensive systematic literature review. Using PRISMA methodology, the insights from the analysis reveal that transfer learning systems have proved its beneficial application in the medical domain by offering various approaches for handling scarcity of labeled datasets and alleviating the shortage of domain expertise in the medical field. However, the present findings confirm that there continue to be opportunities for exploitation of other innovations such as unsupervised learning, and GANs, needing further research in the medical imaging domain.

In this research, firstly we summarized recent surveys of deep learning in medical imaging analysis, and definitions of transfer learning concepts were also provided. Secondly, a comprehensive literature review was conducted following the PRISMA methodology and results categorized in tables. Afterward, a novel conceptual framework was developed grounded in literature and guided by concepts from both medical and deep learning domains. Next, a detailed analysis of the survey findings was summarized in Tables 3-13 Subsequently, we discussed in detail the implications of this work in section 5. The insights from the survey identify several research gaps, for example, very few studies addressed transductive and unsupervised transfer learning approaches. Also, applications in some anatomical areas were identified that can further inform future research developments.

Our major contributions in this research, therefore, include a novel systematic literature review. In prior research, several surveys focused on general aspects of deep learning technologies ranging from the impact of data quality to review of CNNs used for medical image analysis. To our knowledge, this is the first systematic literature review that specifically focuses on transfer learning approaches in medical image analysis emphasizing on the application areas of the human anatomy. The hallmark of a novel survey is the creation of an appropriate conceptual framework to organize the findings from the literature. By doing so, we developed a new conceptual framework to map the results to inform both practitioners and academic researchers in the medical imaging fields. Accordingly, this work can motivate researchers in the computer vision field to explore the underrepresented TL approaches, methods, tasks, and application areas of the human body concerning the use of transfer learning systems as a potential research opportunity area in the future.

6.2 Limitations and future work

Since this is a fledgling research area, this work had two major limitations. First, our survey employed only four online databases. A future literature review could expand the scope to include additional databases and further refine the search terms. Second, despite using articles written only in English for the review process, future research could provide for the inclusion of peer-reviewed articles published in different languages. Accordingly, our future work may benefit from an expanded scope and dimensions to include the general medical domain, for example, application of transfer learning systems in drug discovery.

For many years, deep learning systems have been criticized as a “black-box” due to its complex, opaque, and unpredictable nature. In other words, the operations that begin from the input and end with the output are not very transparent. Additionally, it remains inherently unclear to users how these DL systems arrive at a decision (less transparent). Current research efforts could continue focusing on developing and streamlining deep learning libraries that improve interpretability, transparency, and visualization of the internal workings of the deep learning models to explain their predictions for a much easier decision-making process. Likewise, the impact of data augmentation on transfer learning systems may be investigated further as a research dimension and its effect on medical image analysis especially in cases where limited labeled datasets exist. Correspondingly, the future of using unsupervised transfer learning techniques presents a fresh opportunity for further exploration within the medical imaging domain, specifically with the recent development and application of GANs (Goodfellow et al., 2014).

Currently, most TL systems leverage deep neural networks to accomplish image recognition tasks. Future research could focus on streamlined architectures, specifically lightweight deep neural networks like MobileNets (Howard et al., 2017), MobileNetV2

(Sandler et al., 2019) that are optimized for mobile devices. These lightweight architectures can be memory-efficient hence maximizing the trade-off on size, latency, and accuracy in environments with restricted computational resources, for example, in mobile devices. This area is still nascent and remains an open research area for the exploration of transfer learning systems on medical images with the potential of improving predictive performance even further.

PART II: THE EFFECTIVENESS OF TRANSFER LEARNING ON MEDICAL IMAGES

1.0 INTRODUCTION

The fundamental building blocks, models, and underlying algorithms of Deep Learning (DL) have existed for many years. In the era of big data, data has become much more pervasive in everyday life. These DL algorithms require huge amounts of data for training to succeed in many machine learning tasks such as medical image recognition. Current DL architectures are massively parallelized allowing them to tremendously take advantage of modern GPU's parallel processing platform and specialized hardware to accelerate the speed training of very large scale datasets. Equally, the availability of improved techniques and open-source software libraries for deep neural networks, for example, TensorFlow (Abadi et al., 2016) and Keras (Chollet & others, 2015), have streamlined the efficiency of deploying DL algorithms.

Despite these impressive advancements, training of deep neural networks with large amounts of data is computationally expensive and time-consuming, a challenge that remains in the medical domain. Furthermore, the scarcity of large annotated datasets, specifically medical images persists. To solve these challenges, recent studies have demonstrated the successes of using the transfer learning (TL) paradigm on medical image recognition problems (Litjens et al., 2017; Shie et al., 2015; Shin et al., 2016a; Wang et al., 2017).TL, a popular approach in deep learning, allows us to use a pretrained model and transfer knowledge learned from one domain to solve a new problem in a related domain. Motivated by the success of TL in resolving the shortage of very large labeled datasets, we address these challenges by developing a novel DL network to evaluate the effectiveness of transfer learning on medical images.

The remainder of this section is organized as follows. In section 2, the literature review is discussed. We describe the methodology in Section 3. The experiments and results are discussed in Section 4. Next, we present a discussion of the findings in Section 5. Finally, section 6 presents the conclusion, limitations, and future research.

2.0 LITERATURE REVIEW

2.1 Overview

This section provides a general review of the literature using transfer learning on medical images. We also provide a background of deep learning and transfer learning for medical image classification followed by a detailed explanation of the proposed artifact.

2.2 Deep learning

Deep learning (DL) is an ongoing and popular research area in the mature field of Machine Intelligence providing researchers tools and techniques to solve complex tasks involving very large-scale datasets. Furthermore, DL is a subclass of machine learning applying mathematical algorithms to mimic the cognitive abilities of the human brain to perform human-level tasks, for example, speech and image recognition. The idea of deep learning is to train a computer to mimic the structure and function of the human brain, specifically the biological neural networks through the fundamental concepts of Artificial Neural Networks (ANN) to solve computer vision and recognition problems. For example, Waymo, formerly Google self-driving car project, is an autonomous driving company that uses deep learning techniques to develop self-driving vehicles trained on millions of miles of public roads and over a billion simulated miles to handle complex scenarios such as traffic signs, pedestrians, car surroundings, etc. within the context of real-world driving conditions (Waymo, 2019). According to LeCun et al., (2015) deep learning “allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction”. Deep learning models, for example, DCNNs have demonstrated tremendous achievements in the classification of medical images (Ker et al., 2018; Litjens et al., 2017). Within the DCNNs, the convolutional filters provide the workhorse power by convolving across the underlying images to extract features necessary for the learning and prediction of outputs.

Deep learning models can comprise hundreds of hidden layers or convolutions generating millions of parameters that require powerful and large-scale computing resources to train the deep neural network. Often in practice, training convolutional neural networks

(CNN) on large datasets from scratch requires an enormous amount of resources such as time, money, and computing infrastructure. Although a primary challenge with DCNN's is the enormous data and huge computational resources required for training purposes, recent advancements in Next-generation computing architectures; modern innovations in DL algorithms, such as Generative Adversarial Networks (GAN); and the dawn of powerful and efficient Graphical Processing Units (GPUs) known for their high performance in computing large matrix operations and convolutions simultaneously, have accelerated the use of efficient hyperparameters necessary for the optimization of deep learning models. Additionally, deep learning frameworks, for example, TensorFlow (Abadi et al., 2016), PyTorch (Paszke et al., 2017) and Keras (Chollet & others, 2015) have been optimized to scaleup and speed up the use of federated and distributed learning algorithms for deep neural networks (O. Gupta & Raskar, 2018; McMahan et al., 2017). Therefore, with transfer learning, we do not need to retrain the entire CNN network from scratch. As has been previously reported in the literature, using different transfer learning approaches can efficiently improve performance (Yosinski et al., 2014) and accuracy of training deep learning models. Table 14 below summarizes the different types of transfer learning settings (Pan & Yang, 2010).

Table 14. Types of transfer learning and their different settings

Learning Strategy	Related areas	Source and Target domains	Source domain labels	Target domain labels	Source and target tasks	Tasks
Inductive Transfer Learning	Multi-task Learning	Similar	Available	Available	Different but Related	Regression Classification
	Self-taught Learning	Similar	Unavailable	Available	Different but Related	Regression Classification
Transductive Transfer Learning	Domain Adaptation, Sample Selection bias, Co-variate shift	Different but Related	Available	Unavailable	Similar	Regression Classification
Unsupervised Transfer Learning		Different but Related	Unavailable	Unavailable	Different but Related	Clustering, Dimensionality Reduction

The following scenarios also provide classifications or further categorization on different settings of when to apply transfer learning (Pan & Yang, 2010).

1. **Instance transfer:** This approach aims to reuse knowledge gained from parts of the *source domain* data to learn the tasks of a *target domain*. In other words, the importance is given to re-weighting and sampling.
2. **Feature-representation learning:** The idea behind this approach is to apply optimal features learned from the source domain that can better represent features in the target domain. The objective, in this case, is to minimize error rates and significantly improve the performance of learned features of the target tasks.
3. **Parameter transfer:** The assumption in this scenario is that models used for similar tasks between the *source* and *target* domains share common parameters or prior distribution of priors. Essentially, this means that learned weights can be transferred across different domain tasks.
4. **Relational-knowledge transfer:** The intuition behind this case, is that unlike the other three previous approaches, the contextual relationship is between the *source* and the *target* tasks based on similar data. The implication is that the relational knowledge learned from similar data can be inherently transferred across *source* and *target* tasks.

   The relationship between the different settings and approaches to transfer learning is summarized in Table 15.

Table 15. Methodological approaches to transfer learning

	Inductive Transfer Learning	Transductive Transfer Learning	Unsupervised Transfer Learning
Instance transfer	x	x	
Feature-Representation Learning	x	x	x
Parameter-Transfer	x		
Relational-Knowledge Transfer	x		

2.3 Transfer learning

Transfer learning using pretrained DCNN's are increasingly finding wide adoption in solving a considerable number of challenging problems in the medical domain. We introduce this section with a formal definition of transfer learning following the works of Pan and Yang, (2010) to describe the problem domain in transfer learning. According to Pan and Yang (2010), they proposed two unified definitions and notations for transfer learning as follows:

Definition 1: Given a *source domain* DS and *learning task* TS, a *target domain* DT and *learning task* TT, transfer learning aims to help improve the learning of the target predictive function fT (·) in DT using the knowledge in DS and TS, where DS /= DT, or TS /= TT.

In the above definition, a domain is a pair D = {X, P(X)}, thus the condition DS /= DT implies that either XS /= XT or PS(X) /= PT (X). Similarly, a task is defined as a pair T = {Y, P (Y |X)}. Thus, the condition TS /= TT implies that either YS /= YT or P(YS|XS) /= P (YT |XT). When the target and source domains are the same, i.e. DS = DT, and their learning tasks are the same, i.e., TS = TT, the learning problem becomes a traditional machine learning problem.

Definition 2: Given a *source domain* DS and a *learning task* TS, a *target domain* DT and a *learning task* TT, inductive transfer learning aims to help improve the learning of the target predictive function fT (·) in DT using the knowledge in DS and TS, where TS /= TT. This formal definition applies to inductive transfer learning as summarized in Table 14.

From the above definition, the inductive transfer learning scenario refers to where the *source* and *target domains* are similar while the specific *target tasks* to be performed are different but related. Perhaps, Multi-Task Learning (MTL), an inductive transfer approach is one of the most successful and widely adopted approaches because of its ability to improve generalization by utilizing specific domain information obtained from training signals of related tasks (Caruana, 1997; Ruder, 2017). Some of the many successful applications of MTL range from drug discovery (Ramsundar et al., 2015) to NLP (McCann et al., 2018). Regardless of the selected TL approach, the goal is to find an accurate marginal or conditional distribution difference in the source domain or a combination of both (Weiss et al., 2016). In this work, we focus on the feature-representation learning of the transfer learning approach where the source and target class labels are different, but we pick the optimal features from the *source domain* to classify our target labels.

2.4 Transfer learning for medical image classification

Humans have the inherent capacity to conceptualize complex concepts learned in one domain and use that knowledge learned to solve another related task in either a similar or different domain. For example, a person can learn how to play the guitar and use this knowledge to learn how to play the violin. The underlying concepts are such that it is much easier to cross-reference related tasks and using the knowledge learned from these tasks to solve other related tasks. Researchers and data scientists believe that this concept of knowledge transfer to solve related tasks from one domain to another domain is very paramount towards achieving the goal of strong AI. Traditionally, in the context of deep learning, accessing very large datasets with labeled data for supervised learning is not only tedious but time-consuming and expensive. Therefore, the concept of transfer learning is gaining wide adoption and finding success in computer recognition tasks.

Given the formal definition and notations of transfer learning (TL) described earlier, TL can also be defined as the ability to identify deep connections (Cook et al., 2013); the ability to extend what has been learned from one domain to a new domain (Z. Li et al., 2018). The insight behind TL is to leverage pretrained models to transfer knowledge gained through solving a specific task and then reuse that knowledge learned to decode different problems unrelated to the same domain. One popular computer vision problem is medical image classification, localization and segmentation tasks where one can retrieve knowledge learned from a non-medical image domain (source) and make predictions in a medical image domain (target). Recent studies in computer and vision literature have provided empirical evidence on the successes of using TL with CNNs to represent learned features trained on very large-scale datasets. For instance, the following studies demonstrated the use of CNNs architectures pretrained on ImageNet as either feature extractors (Donahue et al., 2014; Razavian et al., 2014; Bolei Zhou et al., 2018) or fine-tuning (Girshick et al., 2014; Oquab et al., 2014) networks. With the success of CNNs, the transferability of deep representations across tasks has been comprehensibly investigated, especially using the transfer learning paradigm (Azizpour et al., 2016; Huh et al., 2016; Yosinski et al., 2014). Also, the use of pretrained networks in TL enables deep convolutional neural network (DCNN) models to improve its generalization performance to new classification tasks previously unseen by the model. In a pretrained model, the trained weights enable the bottom hidden layers of the ConvNets to learn low-level universal features such as curves, edges, and lines useful for most image analysis tasks. The top convolutional layers tend to specialize in learning more abstract features (e.g., eyes, nose, ears) and fit those features to the specific classification task of interest (e.g., face or jawlines). Transfer learning improves the learning of interaction of relationships or patterns in the target domain by leveraging knowledge from related domains.

In many image recognition tasks where limited dataset exists, application of TL techniques has achieved considerable success in transferring knowledge from one domain to another to solve different image recognition problems. Medical image classification is considered a sub-domain of image classification problems and inherent neural networks such as CNNs for classification challenges can also be applied to it. Furthermore, in the computer vision domain, specifically the medical domain, prior research suggests that using TL with ImageNet pretrained models can have a significant impact on the success of medical image classification tasks (Bar et al., 2015; Shin et al., 2016b). The literature review also shows that the use of a transfer learning system on the classification of Optical Coherence Tomography (OCT) images yielded an accurate model that rivaled the judgment of six human experts (Kermany et al., 2018). Regarding medical image classification tasks, Wang et al., (2017), demonstrated a multi-label classification task using DCNN architecture to evaluate the performance on the Chest X-ray 14 dataset.

Over the past years, several studies have contributed to the development of deep learning networks with capabilities to learn representations of feature maps with multiple levels of abstraction (Goodfellow et al., 2016), for example, AlexNet (Krizhevsky et al., 2012), GoogleNet (Szegedy et al., 2015), VGGNet (Simonyan & Zisserman, 2015), ResNet (He et al., 2016), and DenseNet (Huang et al., 2017). Representation learning refers to a collection of methods that allow machines to automatically discover insights from raw data (Hwang & Kim, 2016). Prior studies have also focused on the use of pretrained ResNet-50 and DenseNet-121 architectures on medical images to develop state-of-the-art models to address classification and detection problems. For example, Baltruschat et al.,(2018) leveraged the ResNet-50 architecture pretrained on ImageNet to build a deep model using TL on the Chest X-ray 14 dataset. Moreover, weakly supervised learning has been used to examine pathology localization through the classification of thoracic diseases (Hwang & Kim, 2016; Z. Li et al., 2018; Sedai et al., 2018; Yan et al., 2018; Yao et al., 2018; Bo Zhou et al., 2018). In binary classification tasks, researchers used the ChestX-ray 14 dataset for pneumonia detection using the CheXNet model (Q. Guan et al., 2018; X. Guan et al., 2018; Rajpurkar et al., 2017). CheXNet is a recent DCNN effort on the classification of chest X-ray images using a fine-tuned DenseNet-121 with a modified fully connected layer. Researchers have found that higher resolution images can improve model performance especially with the use of spatial location information, which greatly improves classification accuracy (Guendel et al., 2018). Our approach extends the depth of the proposed network and introduces active interactions among the learned feature maps into the network. We use pretrained models for our novel method and apply transfer learning strategies, which benefit from less time spent in learning new tasks.

2.5 Proposed artifact

In recent years, the popularity of deep CNNs with the ability to learn multi-scale features in different visual recognition tasks has given rise to the design of other multi-scale CNNs to improve some of the inherent computer vision challenges. Inception network (Szegedy et al., 2015) is one such heavily engineered class of CNN, consisting of 22 layers of neural networks to solve classification and detection tasks. The Inception model based on the prior work by Line et al., (2014) incorporated a dimensionality reduction layer (1 x 1 convolutional filter) to improve the expensive computation and training process of the network. Additionally, the other notable novel contributions of Inception networks were the introduction of inception modules, and kernel tricks to improve performance. The introduction of auxiliary classifiers in the Inception network helped to mitigate the problem of vanishing gradients, in other words, preventing parts of the network from ‘dying out’. Similarly, later versions of Inception networks introduced further improvements such as the use of residual connections which in effect dramatically improved the speed and efficiency of training the network (Szegedy, Ioffe, et al., 2016). An illustration of the inception module is shown in Figure 13. Our proposed novel method uses the architectural configuration of the Inception network but excludes the auxiliary classifiers as shown in Appendix 2 (Szegedy et al., 2015).

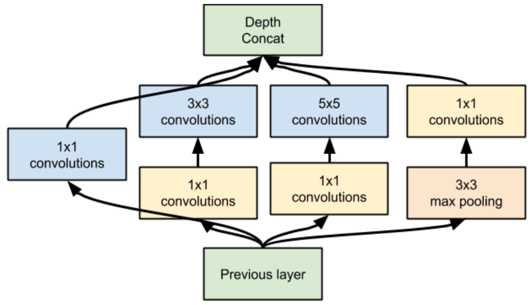


Figure 13. An inception module with dimensionality reduction

Another class of deep CNN is the DenseNet architecture, which proved that CNNs can substantially go deeper and give much more accuracy without sacrificing performance, especially when training deep networks. DenseNet comprises 121 layers concatenated together in a feed-forward version as shown in Figure 14 (Huang et al., 2017). Like Inception, DenseNet has the advantage of alleviating the problem of vanishing gradients. Other compelling benefits that DenseNets offers include feature reuse, strong feature propagation, and a significant reduction of the number of parameters thus improving the efficiency of training deep networks (Huang et al., 2017). Our proposed novel method incorporates the beneficial aspects of DenseNet such as parameter efficiency, and concatenation of feature maps which promotes feature reuse and replaces the inception modules in favor of DenseNet modules.

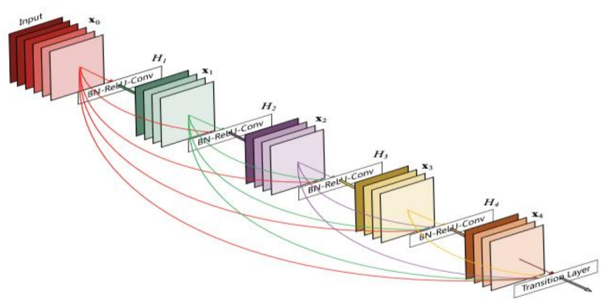


Figure 14. A 5-layer dense block with a growth rate of k=4

Thus, TL techniques and results from medical image classification have been reported in the literature but it is not clear to what extent the findings are effective towards generalizing the models in the medical images domain. In this study, we seek to answer the central objective of determining the effectiveness of transfer learning using our novel deep model on medical images and finding an optimal cut-off point through the fine-tuning technique.

In this work, we develop a deep model based on a deep transfer learning method with increased depth for medical image classification. Then we further investigate the effectiveness of TL on medical image classification and potentially advance the learned knowledge (model) to generalize for other unrelated problems in the medical image domain. Finally, in medical image diagnosis, it is worth noting that a deep learning system that minimizes the occurrences of false positives is much more beneficial in mitigating the risks associated with misdiagnosis.

3.0 METHODOLOGY

This section presents the steps and procedures taken for the medical image classification task. Section 3.1 describes the datasets and data pre-processing steps taken for the classification task. Sections 3.2 and 3.3 describes the experimental setup and implementation environment for achieving the task. Moreover, in this work, we apply a machine learning methodology, specifically using non-linearity functions on medical imaging recognition tasks. A pretrained DIM network is used for transfer learning purposes to leverage feature learning from source data (non-image data) to our target data (medical images). Finally, the DIM network is trained using versions of the novel deep architecture to investigate the effectiveness of fine-tuning on the medical images.

3.1 Datasets

To ensure the robustness of our proposed method, we trained and evaluated it on two publicly available datasets: CIFAR-10, and ChestX-ray 8. The datasets were randomly split into 90:10 (CIFAR-10) and 80:10:10 (ChestX-ray 14) proportions for training and validation respectively. The two datasets are used for the training and validation phase of our novel method. The following is a description of each of the datasets:

3.1.1 CIFAR-10

The CIFAR-10 dataset comprises of 60,000 color images from diverse objects with an image size of 32×32 pixels for each image and categorized into 10 classes (*airplane, bird, dog, frog, deer, dog, horse, ship, truck, automobile)* for a total of 6000 images per class (Krizhevsky, 2009). During the training of the proposed method, the datasets were automatically split into 50000 training images and 10000 test images.

3.1.2 ChestX-ray 14

Chest X-ray is one of the popular imaging modalities due to its cost-effectiveness in performing medical examinations. Although the diagnosis of Chest X-rays can be challenging enough, this dataset is one of the largest publicly available medical images focusing on clinical diagnosis of Chest X-rays. This dataset consists of 112,120 frontal-view chest X-ray images extracted from 30,805 unique patients spanning 14 different classes of thoracic pathologies including Cardiomegaly, Consolidation, Edema, Emphysema, Effusion, Fibrosis, Hernia, Infiltration, Nodule, Mass, Pleural_thickening, Pneumonia, and Pneumothorax (Wang et al., 2017). Accordingly, this dataset is considered suitable for weakly supervised learning with labels that have acceptable accuracy for research efforts.

3.1.3 Data preprocessing and transformation

Data pre-processing is important and is an on-going research area in machine learning. A well-balanced dataset is critical for obtaining a much more accurate model on image recognition tasks. According to Garcia et al., (2015), data pre-processing may include data cleaning, integration, transformation, and reduction operations. Quality decisions depend on quality data which leverages the ability to transform data into a form that can be efficiently and accurately processed by computers. Figure 15 shows the distribution of all the diagnoses associated with the Chest X-rays. As shown in Figure 15, the classes in the Chest X-ray dataset is heavily imbalanced, a term known as *linear imbalance* where minority classes are almost equal, and a big gap exists between majority and minority classes. For example, the majority class has over 60,000 images classified as *No finding* at the expense of the rare classes such as Pneumonia, requiring adjustments to balance the distributions across the other classes.

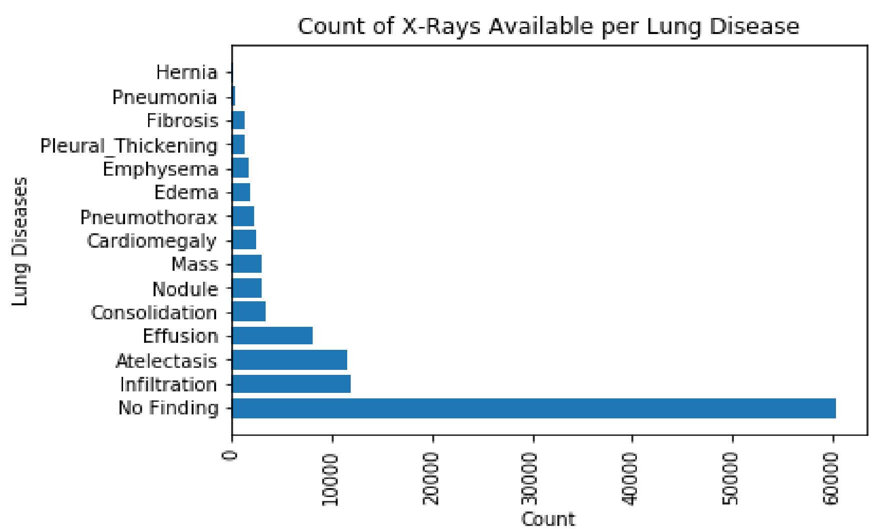


Figure 15. The probability distribution of Chest X-ray classes

To overcome this problem of class imbalance, we applied the undersampling method on the *No Finding* class to standardize the count around the mean. After the data preprocessing step, the results are displayed in Figure 16 where the goal is to distribute the weights across the classes to minimize the large differences between the majority and minority class.

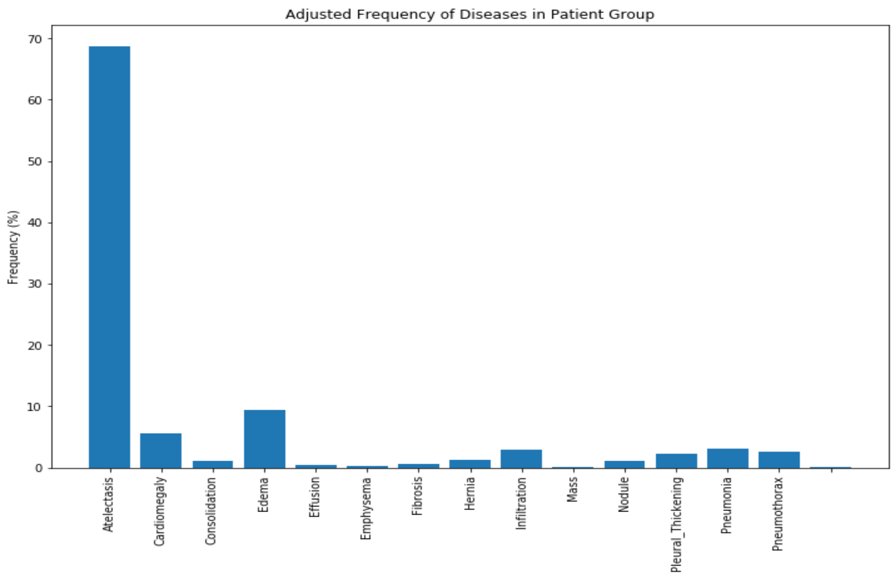


Figure 16. Adjusted distribution after pre-processing

In Figure 17, the sample images show the different types of thoracic pathologies with their associated labels.

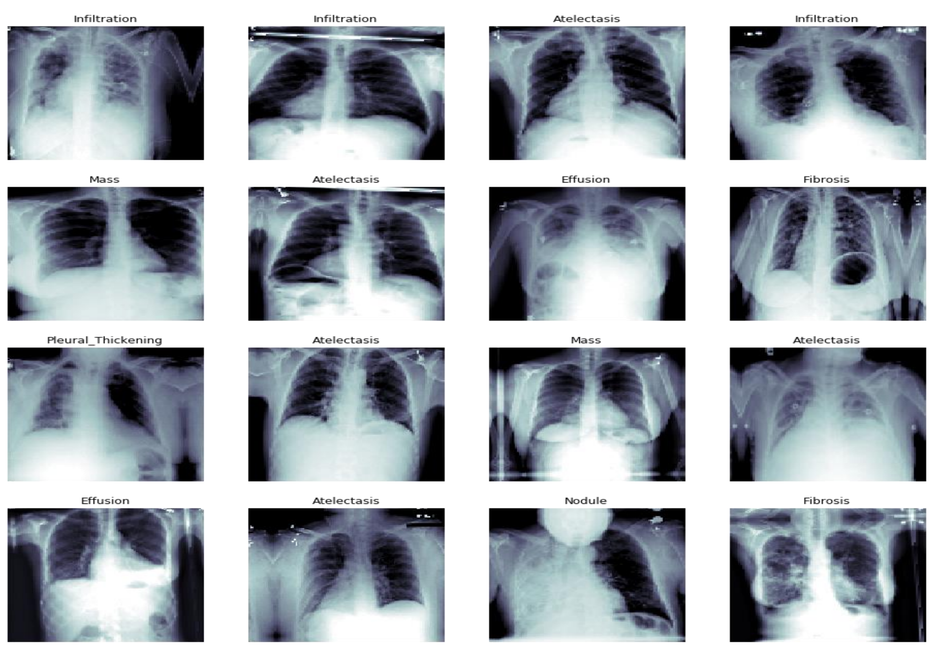


Figure 17. Images showing types of Chest X-ray diagnoses

In data transformation, a Keras library *flow_from_dataframe* was used to binarize the labels to binary vectors from the categorical data by applying one hot encoder. The reason for using a one-hot encoder is to transform the categorical variables into a form that the machine algorithms can understand (e.g. into 1 and 0s). Afterward, the images were resized to a dimension of 112 x112 pixels to match the input layer of the pretrained model and assist with minimizing the overhead during the training step thus improving the overall learning speed. Additional preprocessing steps included normalization and standardization of training and validation datasets.

3.2 Deep CNN phase

For our experimental setup, we developed our model using stacked CNNs, pooling, and fully connected layers. Next, we prepared our novel deep model by training it on the CIFAR-10 dataset as a pretrained network. Many CNN architectures developed for different deep learning tasks have been proposed in the literature (Szegedy et al., 2015). In this work, we use our novel architecture with additional structural and functional improvements by introducing a DenseNet module (DIM module) into the network. The DIM modules provide an efficient yet simple block that maximizes on multiscale representations of learned features.

3.3 Implementation details

We developed our implementation strategy for our proposed model in Python using Keras (Chollet & others, 2015), a deep learning framework with TensorFlow (Abadi et al., 2016) as the backend to implement the network architecture. Following prior work, we trained our network using stochastic gradient descent (SGD) with weight decay of 0.0001, a momentum of 0.9 (He et al., 2016; Szegedy, Vanhoucke, et al., 2016), and a mini-batch was set to 1024. Unlike Adam and RMSprop optimizers which tend to converge faster, SGD was selected for its better generalization properties (Luo et al., 2019). All experiments were conducted using model and data-parallelism. The setup included a multiple GPU environment running Windows 10 operating system, utilizing dual NVIDIA TITAN V 12GB/16GB GPU for training and transfer learning with Intel core i7 processor to facilitate computation more rapidly. Additionally, more rigorous training of the proposed method was deployed on

Lawrence Supercomputer running CentOS Linux operating system which comprised of over 2,000 CPU cores, 1.5TB memory and multiple GPU accelerators (2x NIVIDIA Tesla P100 16GB/1x NVIDIA Tesla V100 32GB). By using this GPU configuration platform, we implemented a pretrained DIM network on CIFAR-10 and re-trained the network on the Chest X-ray dataset using fine-tuning approaches. The transfer learning approach involved jointly training the pretrained model with our new classifier on Chest X-ray images and later fine-tuned the higher layers to find the optimal cut-off layers. Moreover, we considered an adaptive learning schedule where each iteration of the learning process used 200 epochs with a decreasing learning rate schedule of 5% for every 10 epochs.

4.0 EXPERIMENTS AND RESULTS

4.1 Overview

This section discusses the transfer learning approach, experiments, and metrics used to measure the performance of the proposed DINET. Section 4.2 explains the transfer learning approach selected for the medical image classification task. Section 4.3 and section 4.4 describes the experiments and evaluation used in this work. The results are presented in Section 4.5. We trained and fine-tuned three versions of the DIM network as shown in Table 16. Notably, all the experiments were fine-tuned based on the pretrained model trained on the CIFAR-10 dataset.

4.2 Transfer learning method

In transfer learning approaches, the encoded information of the DCNN residing at the lower level layers is responsible for detecting low-level features such as colors, visual edges, contours, shapes, and textures which are universal across most image recognition problems. Likewise, the higher-level layers detect more complex abstract concepts and objects (such as “human eye”, or “human ear”). Therefore, the lower level layers were trained in a feed-forward fashion to discover the universal features that will be reused during the training of the medical images (target domain). In this work, we adopted transfer learning with a fine-tuning approach, a popular strategy for model reuse where the weights of the pretrained model were adjusted by unfreezing the higher-level layers of the network for training on the medical image classification task. The common practice is to remove the last layer (SoftMax layer) of the pretrained model and replace it with a new SoftMax layer relevant to the problem under investigation. The new classifier is then retrained on top of the network with the new dataset (medical images). Furthermore, the other vanilla practice involves freezing the weights of the lower-level layers of the pretrained network, this is because basic features that are relevant to our problem such as edges, and curves are already captured. By unfreezing higher-level layers of the pretrained network for training or jointly training with the classifier, and continuing with backpropagation, the network can focus on learning specific data-centric features of our medical images. Figure 18 illustrates the transfer learning workflow for common practices when training a pretrained network (mathworks, 2020).

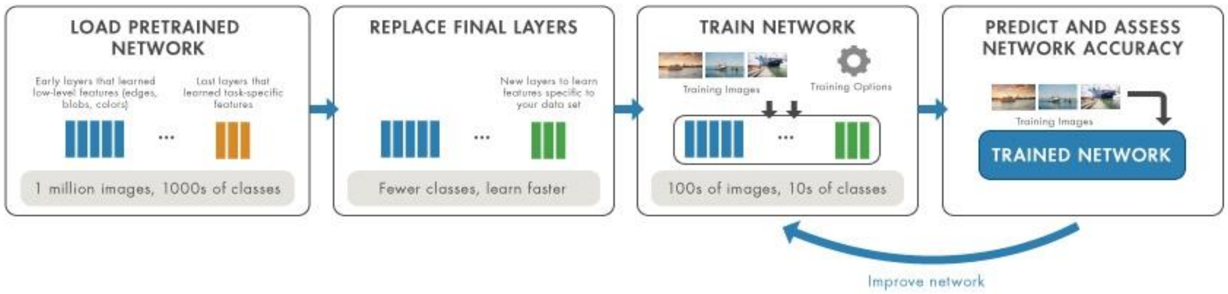


Figure 18. Transfer learning workflow in training a network

4.3 Experiments

We set up several experiments, training the pretrained model for the three DINET versions, where the DIM modules were strategically positioned at different parts (top, middle, and bottom) of the proposed architecture. The original size of the CIFAR-10 images was 32 x 32. The images were then up-sampled resulting in image sizes of 112 x 112. Similarly, the medical images were resized and cropped resulting in an image size of 112 x 112. For data augmentation, we employed the following variations: size rescale, rotation, width, and height range, shift range, zoom range, horizontal flip, and fill mode.

In experiment (1), the DIM ver1 network (top part) was then pretrained on the CIFAR-10 dataset (*source domain*) in preparation to learn our medical images (*target domain*). Once the pretrained model was trained until convergence, the next step was to save on disk the performance of the model together with the trained weights in preparation for the next step of transferring the model to the problem of interest (medical image classification). The fully connected layers of the pretrained model were truncated, and a new classifier defined on top of the network. We used dropout (Srivastava et al., 2014) regularization with a ratio of 0.25 in the fully connected layers and created a SoftMax layer in the new classifier. We also added a batch normalization layer consistent with current practice to facilitate accelerated learning rates during the training of the network. Next, the new classifier was trained with the medical image dataset to ensure that minimal error signal was propagated throughout the network during the training process. Lastly, the fine-tuning step was applied, where several blocks of the pretrained network (higher layers) were unfrozen and retrained together with our medical images while freezing the rest of the network. The process was iterated until the best performing OCL was identified and the results are presented in Table 16. The structure of the network architecture is illustrated in Appendix 3.

In experiment (2), we experimented with DIM ver2 and DIM ver3. First, the DIM module was strategically positioned in the middle of the network for DIM ver2. The training procedure followed in the experiment (1) was implemented for DIM ver 2 as well. Second, the DIM module was also strategically positioned at the bottom of the network and we called this DIM ver3. Similarly, the training procedure used for DIM ver2 was utilized here as well. The results for both experiments are shown in Table 16. In summary, the main application of our novel method was in the use of transfer learning systems in the diagnosis of medical images, more specifically, Chest X-ray images.

4.4 Evaluation procedures and techniques

To assess our proposed network, we empirically examine the effectiveness of knowledge transfer through incremental fine-tuning to evaluate how learned feature representations mitigate the problem of vanishing gradients and identify the OCL that achieves performance improvements in the multi-class classification problem. Also, evaluation metrics like classification losses (categorical cross-entropy loss) metrics were used. The network was incrementally fine-tuned beginning with the top layer to the next few upper ones. With the GPU capabilities, each training scenario took hours to train with fine-tuning of different higher-level layers of the network to determine the OCL.

4.5 Results

Table 16 presents the results of our experiments on the performance of different DIMx versions. The results reveal that when our proposed model was trained using variants of DIMx modules strategically positioned in the network, DIM v1 produced better performance while alleviating the problem of vanishing gradients. The training and loss graphs are depicted in Figure 19 to Figure 21.

Table 16. Performance of different DIMx versions

DIMx version	Fine-tuning layer (OCL)	epochs	Parameters	Accuracy	Loss
DIM v1	Conv_152	200	6,195,615	0.6871	1.7961
DIM v1	Conv_154	200	6,195,615	0.6851	1.8665
DIM v1	Conv_155	200	6,195,615	0.6851	1.8076
DIM v2	Conv_152	200	6,001,210	0.5844	1.84
DIM v3	Conv_264	200	6,004,095	0.5899	1.70

The results in Table 16 are from a multi-class classification of 15 thoracic pathologies. The classification accuracy that gave the best balance between bias and variance trade-off was 68.71%. After each fine-tuning step, we identified the OCL (see Table 16) that gave the best model performance while balancing the overfitting and underfitting problem, a concept known as the bias-variance trade-off. We believe these results are significant and promising on medical image tasks showing that with correct hyper-parameter tuning on our proposed model, transfer learning approaches can produce respectable performances while mitigating the problem of vanishing gradients.

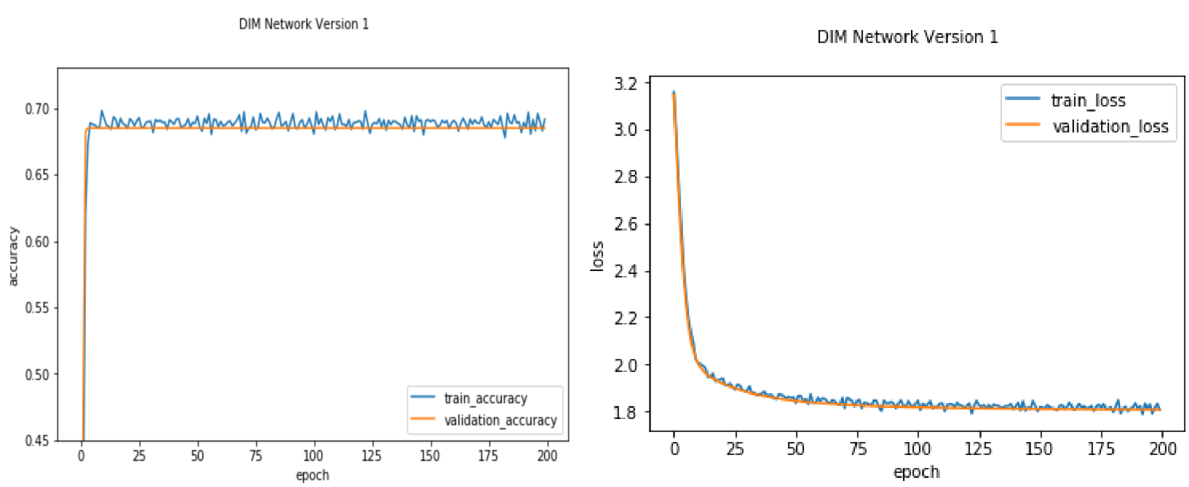


Figure 19. DIM v1 fine-tuned layer i=97

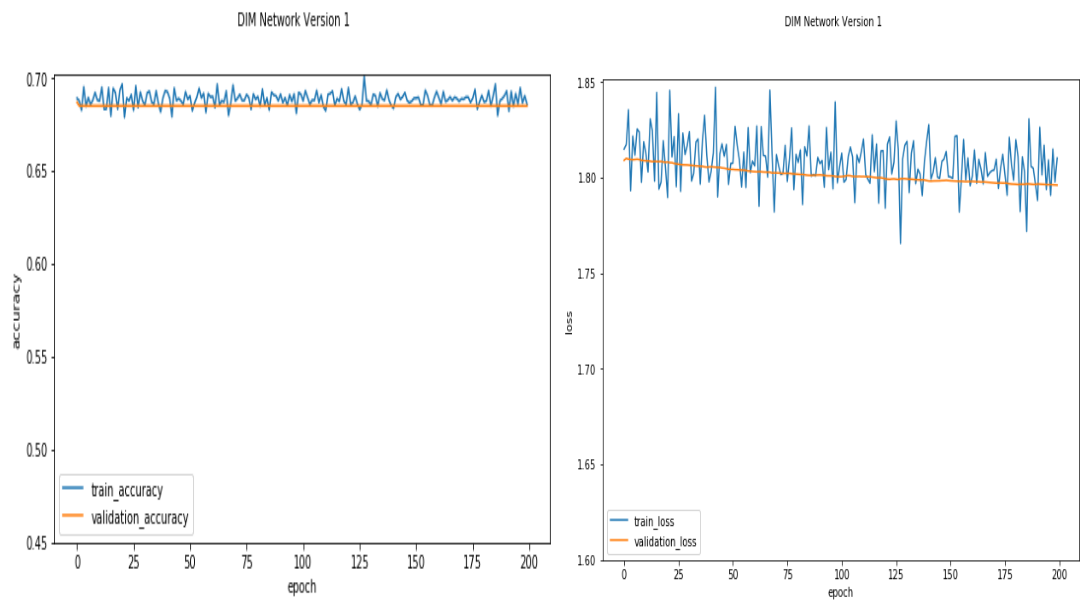


Figure 20. DIM v1 fine-tuned layer i=82

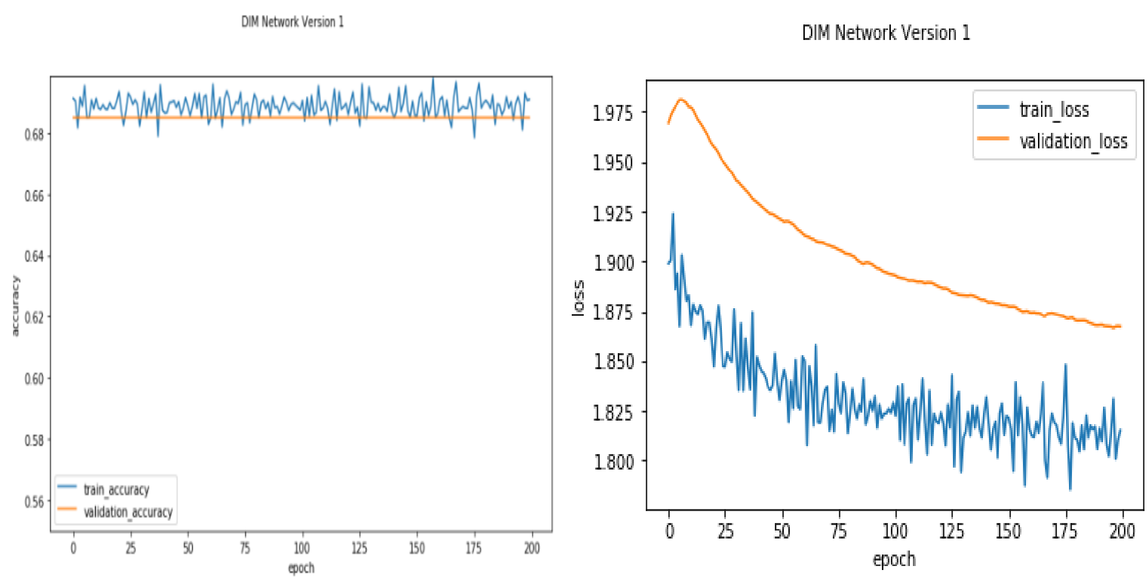


Figure 21. DIM v1 fine-tuned layer i=90

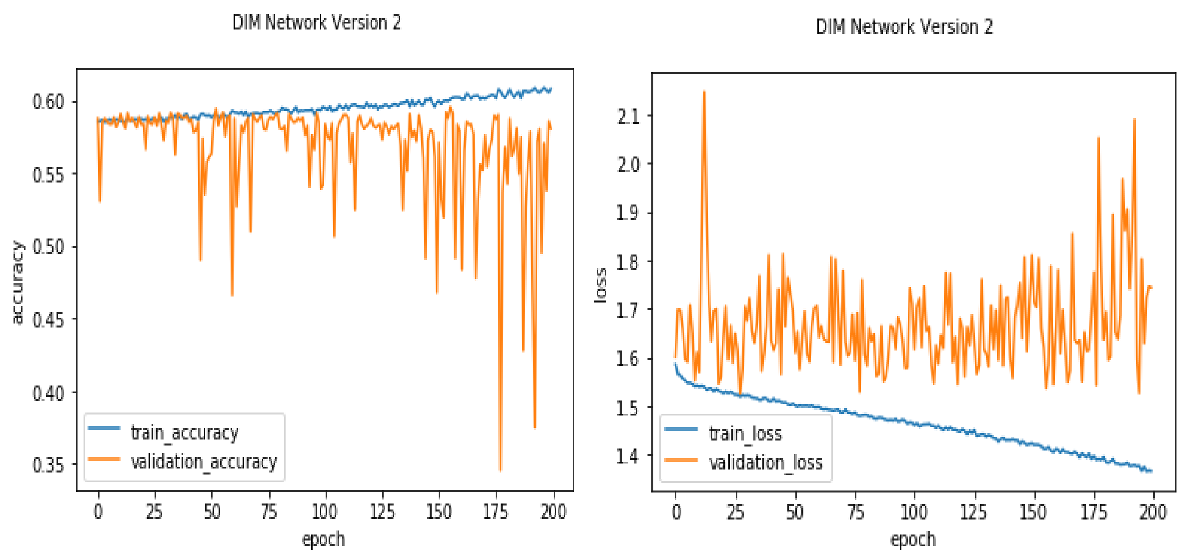


Figure 22 DIM v2 fine-tuned layer i=95

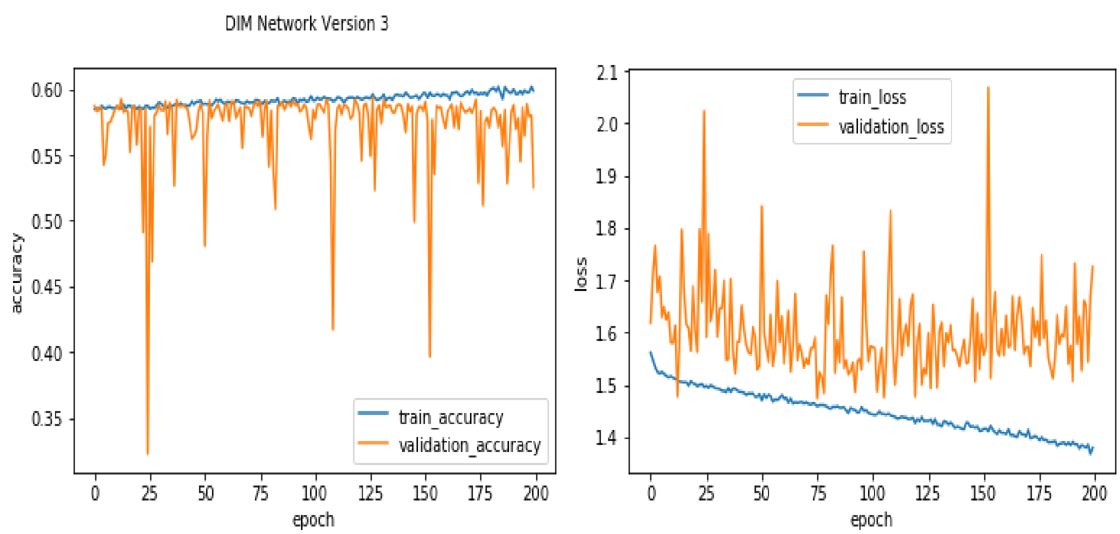


Figure 23. DIM v3 fine-tuned layer i=97

The results of experiment 2 are shown in Figure 22 and Figure 23 respectively.

5.0 DISCUSSION

5.1 Overview

The recent achievements of leveraging transfer learning in reusing the weights from pretrained models and re-training the networks on a new problem of interest have received increased attention in the medical imaging domain. TL is currently very attractive in deep learning because of its ability to train deep neural networks where scarcity of very-large labeled data such as medical images still exists. However, fine-tuning pretrained models on similar datasets have proven to be successful in addressing the challenges with limited data. The proposed deep model has shown that it can increase learned characteristics of the medical image dataset thereby avoiding the effort to train the network from scratch. In particular, the transfer learning approach we used enhanced the robustness of the model by alleviating the problem of vanishing gradients and therefore, improving the convergence of the model. In other words, the convergence concept infers that neither significant change in the error is observed nor any performance increase is reported.

Another observation was that retraining a few of the batch normalization layers jointly with the new classifier improved accuracy and minimized the overfitting problem. Also, we noted that when we introduce a batch normalization layer on top of our classifier, the efficiency of the learning rate improved thus leading to better model performance. This finding can be observed from the experiments performed on DIM v1 where we achieved better results than all the other DIMx versions while mitigating the problem of vanishing gradients. Extraction of knowledge from pretrained models using natural images seems a viable strategy compared to the traditional method of training from scratch especially instances where large labeled domain-specific datasets are limited. Also, the optimal cut-off layers were identified and summarized in Table 16.

Recent studies by Esteva et al., (2017) and Gulshan et al., (2016) where they used Inception V3 demonstrated the success of using pretrained models on natural images and used the deep model to fine-tune on medical data. In this work, we also showed that using transfer learning from nonmedical images that are considerably different from Chest X-rays, can be successfully employed to learn or generalize contextual information from medical images and achieve comparable classification performance. Therefore, this approach to transfer learning could be a viable solution when the availability of domain-specific datasets and computational resources are limited or severely constrained.

6.0 CONCLUSION, LIMITATIONS, AND FUTURE WORK

6.1 Conclusion

In this work, we proposed a novel deep model called the "DenseBlock-Inception Network" (DINET) network to demonstrate the effectiveness of transfer learning systems on medical images and address the problem of scarcity in having domain expertise in the medical field, a major challenge in medical image analysis. The innovation of DINET is its ability to combine the parameter efficiency of DenseBlocks and the feature map extraction abilities of Inception networks while mitigating the vanishing gradient problem. The experiments have empirically confirmed that performance improvements can be achieved when DIM modules are strategically positioned in the network. The proposed deep model was able to successfully perform the classification of medical images from non-image data using a supervised inductive transfer learning approach.

Furthermore, we demonstrated that with transfer learning, the use of non-medical data is a relevant strategy that can be leveraged to train and fine-tune deep neural networks to extract feature maps in the medical images. However, there is more room for improvement, for example, combining features extracted from multi-task and multi-source approaches and then jointly training with the medical images. The results achieved by our proposed model show promising results in the classification of thoracic diseases. Healthcare practitioners such as radiologists can potentially benefit from these DCNN models which can provide an automatic diagnosis of diseases thereby saving time and reduce the workload required to analyze medical images thus shifting their focus on other high-level risk medical images.

Finally, the findings in this work have implications for theory and practice. First, our findings confirm that the effectiveness of using transfer learning systems for medical image tasks while leveraging pretrained models using natural images is achievable. This was successfully established by Shin et al., (2016b) who demonstrated the beneficial aspects of transfer learning systems in transmission of learned representations from natural images that are disparate to medical datasets. Second, while transfer learning systems can offer satisfactory performances with the correct amount of hyperparameter tuning, applying different kernel tricks could potentially help to extract relevant amounts of contextual information and discriminative features in solving various medical image recognition tasks. To this end, we have demonstrated that pretrained models in combination with transfer learning is a relevant strategy for generating learned representations in the Chest X-ray Images and can potentially overcome the problem of limited availability of large labeled data (*domain-specific*).

6.2 Limitations

Despite promising performance results from our proposed deep model, there exist limitations in this work. Training deep neural networks often require very large datasets which in most cases is difficult to access or acquire. While transfer learning may help resolve the problem of having large amounts of labeled data, training requires tedious amounts of time and technology infrastructure (e.g. more GPUs) to evaluate a robust model. Although we achieved the research objectives in this work, due to memory limitations and time-constraint, the CheXpert dataset (Irvin et al., 2019) was not used in this work to further evaluate the generalizability of our deep model.

Data imbalance and the quality of data was another observed limitation. A noisy dataset could negatively impact the performance of deep learning models. Moreover, unbalanced data may lead to biases, especially during the learning process. Despite our efforts to use available techniques to address the class imbalances, the data quality may have impacted the performance of the model to accurately classify all the thoracic pathologies. Researchers widely believe that the medical image dataset has several problematic issues relating to unlabeled images and noise which could potentially prevent deep neural networks from achieving desired performance capabilities. To motivate future research advancements, other datasets could be made available to support the evaluation of different DCNN network architecture performances, generalizability, and effectiveness of transfer learning systems.

6.3 Future work

As mentioned previously, this dissertation focused on the implementation and evaluation of a novel deep model to evaluate the effectiveness of transfer learning systems on medical images. Future work may focus on improving the effectiveness of transfer learning by training with more data to improve the discriminative features and increase learned interactions on new data to solve medical image recognition tasks. Potential improvements may also focus on using different class balancing techniques such as Synthetic Minority Over-sampling Technique (SMOTE) (Chawla et al., 2002) to address the issues present in many deep learning datasets of under-sampling and over-sampling in classes.

In future research, we may consider other approaches of transfer learning like multi-task learning (Caruana, 1997; Y. Zhang & Yang, 2018), multi-source learning (Christodoulidis et al., 2017), and multi-source domain adaptation (Sun et al., 2015; S. Zhao et al., 2020) to examine in much more depth, the effectiveness of the cross-domain transfer of features and the effectiveness of fine-tuning of pretrained models on other classes of deep neural networks. In the literature survey, we observed the limited application of unsupervised transfer learning which could be another attractive research area for researchers. Moreover, the impact of data quality and dataset size on transfer learning systems may be investigated further to measure the effectiveness of the underlying functions from the input (*source domain*) to outputs (*target domain*).

Ultimately, we believe that input from domain experts should be included in the evaluation efforts to enhance the design, development, and implementation of deep neural networks for solving medical imaging problems. Finally, although some of the issues continue to exist, we believe our findings may further create and drive opportunities to motivate future research directions on medical image tasks.

PART III: DOMAIN ADAPTATION USING AN AUTOENCODER ON MEDICAL IMAGES

1.0 INTRODUCTION

Recently, deep neural networks, especially CNNs have achieved remarkable success in computer vision problems such as segmentation tasks. However, technical issues may make effective application of deep learning algorithms challenging since they are highly reliant on huge volumes of annotated data. While large volumes of data are usually generated and remain unlabeled in most fields, the human cost associated with annotating data with labels is still prohibitive, especially in the medical domain. To tackle these challenges of acquiring labeled data, solutions in literature have been proposed to overcome unlabeled data scenarios, for example, domain adaptation. Domain adaptation is an important paradigm in machine learning that has in recent times received tremendous attention in computer vision and natural language processing domains. Pan and Yang, (2010) stated domain adaptation as “the difference between the marginal probability distributions of source and target data” (same tasks but different domains) and classified it under the transductive transfer learning approach. In other words, domain adaptation can be defined as a transfer learning scenario where the sample and label spaces are the same but only the probability distributions are different (Kouw & Loog, 2019).

With transductive transfer learning, the target tasks are similar, but the source domains are dissimilar. Because of its ability to transfer knowledge from labeled data (source domain) to unlabeled data (target domain), domain adaptation provides a cost-effective solution in adapting deep learning algorithms to new application domains that suffer from the unavailability of unlabeled data. Besides, the popularity and immense success of CNNs in image recognition tasks (LeCun et al., 2010), have enabled the exploitation of different transfer learning approaches, including the domain adaptation paradigm. Recently, transfer learning domain adaptation has found success in medical imaging tasks. For instance, Chen et al., (2018), proposed a novel approach based on unsupervised domain adaptation leveraging

GANs for lung segmentation in chest X-rays. Furthermore, Ghafoorian et al., (2017), applied domain adaptation and transfer learning leveraging CNNs for brain lesion segmentation achieving a dice score of 0.76. To this end, domain adaptation can be considered as an instance of transfer learning and an extensive listing of domain adaptation materials is collated and maintained by Zhao, (2020). For a comprehensive review of domain adaptation for visual tasks, please refer to Csurka, (2017) and Patel et al.,(2015b); single-source unsupervised domain adaptation (Wilson & Cook, 2020); and multi-source domain adaptation (Sun et al., 2015).

In this work, we mainly focus on the unsupervised domain adaptation where we train an autoencoder with labeled source data and use the probability distribution to learn the unlabeled target data (M. Long et al., 2015). Consequently, we aim to classify our medical image dataset, specifically Chest X-rays through fine-tuning technique. To achieve this goal, we adopt the classification scheme in transductive transfer learning (Pan & Yang, 2010) using a domain adaptation paradigm to harmonize the learning of the different distributions between the source domain (labeled) and the target domain (unlabeled). To explore this problem, we aim to examine the effectiveness of transfer learning using U-Net architecture on our medical images. The main idea is to extract local and contextual information for the segmentation task while ensuring less error in the prediction of the unlabeled medical images.

The rest of this work is organized as follows. In section 2, the background literature is provided. Section 3 covers the methodology and datasets adopted in this work. In section 4, the evaluation of the experiments is formulated, and the results reported. Discussions of the experiments are elaborated and presented in Section 5. Finally, the conclusion, limitations, and future work are summarized in Section 6.

2.0 BACKGROUND AND MOTIVATION

Artificial intelligence (AI) is rapidly transforming and revolutionizing our day-to-day lives, in almost all sectors like transportation, healthcare, and manufacturing industries. With AI leveraging advancements in machine learning and deep learning, continued enhancement in computing is creating a paradigm shift towards automation of tasks and improvements in decision-making processes. In the past few years, breakthroughs in deep neural networks trained on very large datasets (ImageNet), (Krizhevsky et al., 2012), have paved the way for the development of even deeper networks (Simonyan & Zisserman, 2015). The success of deep convolutional networks have surpassed the state-of-the-art benchmarks in many visual recognition tasks, for example, in the work by Girshick et al., (2014) where they improved the existing mean average precision (mAP) on a very large dataset by 30% using domain-specific fine-tuning approaches. Although, many hospitals generate vast amounts of clinical images stored in their digital repositories, acquiring annotated large-scale medical image datasets for the most common tasks in medical image analysis: classification and image segmentation, remains a challenge.

In the domain of medical imaging, deep learning technologies are rapidly finding applications in various medical image analysis tasks due to the availability of imaging data. Nevertheless, except for the recently released ChestX-ray 14 (Wang et al., 2017) and ChestXpert (Irvin et al., 2019) datasets, considered as some of the largest publicly available database of labeled datasets, there still exist difficulties in finding huge amounts of labeled medical images which is fundamental for training in deep learning. As has previously been reported in the literature, several issues exist when sourcing for labeled medical images: (1) Lack of specialized tools for annotations; (2) Limited domain expertise to label images; (3) Very expensive to have radiologists for annotations of medical images; and (4) Inconsistencies when labeling images, thus creating a barrier especially when the goal is to train an effective DL model for medical image analysis (Sahiner et al., 2019). These underlying issues are emblematic of the medical image domain exacerbated by the difficulties in accessing annotated medical images and the presence of distribution shifts especially when data is from the same source, persists (Ciga et al., 2019). The distribution shift is a phenomenon in machine learning where the training examples differ from real-world examples thereby affecting the ML algorithms to draw wrong conclusions.

Currently, unsupervised learning is still in nascent research stages when compared to supervised and semi-supervised learning approaches. Although unsupervised learning has gained importance over the last few years, especially in addressing the issue of unlabeled datasets, there is a relatively low number of existing studies towards incorporating transfer learning systems on medical image tasks. Dosovitskiy et al.,(2014), demonstrated the effectiveness of using unsupervised feature learning combined with data augmentation techniques for visual object recognition. A similar approach was used by Ronneberger et al., (2015) where they developed a novel architecture known as U-Net which incorporated skip-connections in the network to utilize spatial and semantic information. The U-Net model was used for biomedical cell segmentation and surpassed the benchmark scores of a previous state-of-the-art method by Ciresan et al.,(2012). Besides, data augmentation has been reported in the literature as techniques that increase the amount of limited labeled images (Christ et al., 2017; Milletari et al., 2016). Nonetheless, one of the most popular applications of unsupervised learning is an autoencoder. Autoencoders are a class of neural networks that aims to replicate their inputs to be equal to their outputs. In this regard, U-Net is also known as autoencoders based on the functionality of an autoencoder. Towards this end, autoencoders are designed as an efficient dimensionality reduction technique that works very well by preserving information from original input and removes noise from the data in the process.

2.1 Semantic segmentation in medical image analysis

Recently, the World Health Organization declared the Coronavirus disease 2019 (COVID-19) outbreak a global pandemic (WHO, 2020b). The COVID-19 infections occur in the respiratory system of the human body, more specifically in the lower lobes of the lung area (Bernheim et al., 2020). As a result, the computer vision and medical imaging community have witnessed an unprecedented amount of research effort devoted to understanding, exploring, and in combating the novel coronavirus through innovative DL methods with U-Net leading at the forefront. At the same, over 4500 research publications have been identified that implement AI and DL methods to diagnose the novel coronavirus (Bullock et al., 2019). Accordingly, the U-Net architecture has continued to receive renewed and sustained interests especially in the wake of the COVID-19 pandemic. Table 17 summarizes and illustrates the latest applications of the U-Net model in the diagnosis of coronavirus disease.

Table 17. Current application of U-Net architectures in COVID-19 diagnosis

Author	Architecture	Imaging modality	Performance	Remarks
Chen et al.,(2020)	UNet++	CT	Sensitivity (100%);Specificity(93.55%); Accuracy(95,24%)	Over 40,000 CT images were used to diagnose COVID-19 and were able to identify correctly 51 out of 106 patients with pneumonia
Gaál et al.,(2020)	Attention U-Net	X-rays	Dice score of 97.5%	Used JSRT and Shenzhen dataset to diagnose COVID-19 applying contrast limited adaptive histogram equalization (CLAHE)
Gozes et al.,(2020)	U-Net	CT	AUC(0.996); Sensitivity (98.2%);Specificity(92.2%);	Used Resnet-50 as a base model with U-Net architecture as a segmentor model for COVID-19 diagnosis.
Jin et al.,(2020)	3D U-Net++	CT	AUC 0.991, specificity 0.922 and sensitivity 0.974	Used transfer learning on ResNet-50 and applied 3D U-Net++ as a segmentor model for diagnosing COVID-19
Li et al., (2020)	U-Net	CT	Sensitivity (90%);Specificity(96%);	ResNet50 was used to extract volumetric information with U-Net as the segmentor model
Zheng et al.,(2020)	UNET	CT	0.959 ROC AUC	Detection of COVID-19 using 3D CNN

Advances in neural network developments especially autoencoders based on CNN breakthroughs have shown great promise towards incorporating domain-specific knowledge into DCNNs specifically for semantic segmentation tasks of medical images. In the medical imaging domain, U-Net architectures and its variant have been successful in many image segmentation tasks across the human anatomy system, for example, brain tumor (Dong et al., 2017), bones (Zeng et al., 2017), kidney (Çiçek et al., 2016), prostate (Yu et al., 2017), lung nodule (Setio et al., 2017), skin lesion (B. S. Lin et al., 2017), liver (Christ et al., 2017) and histology (Sirinukunwattana et al., 2016), etc.

2.2 U-Net network architecture

In the last few years, the U-Net (Ronneberger et al., 2015) has gained a lot of attention as a successful encode-decoder network. The U-Net architecture is synonymously associated with an encoder-decoder architecture based on the improvements of the works by (J. Long et al., 2015) on Fully Convolutional Networks (FCNs) which was very successful for image segmentation tasks exceeding approaches in the period it was developed. FCNs have no dense layers resulting in a reduced number of parameters and less computation time required. Dense layers also known as fully connected layers are parts of a CNN that take the results from the convolution and pooling operations to inform the classification decision of an object. Additionally, prior work by Badrinayarayanan et al.,(2017) made improvements on the FCNs by proposing a novel architecture known as SegNet which comprised of 13 layers of in each of the encoder and decoder networks. Another seminal contribution towards the advancement of FCNs included the works of Chen et al.,(2016) who proposed a DeepLab system incorporating an innovative *atrous algorithm* inside the network. These studies were instrumental in image segmentation tasks and paved the way for the development of the currently popular U-Net architecture.

The U-Net architecture takes the form of a "U" shape hence its name. It is widely considered as one of the first and most common methods for semantic medical image segmentation. According to Ronneberger et al., (2015), this architecture comprises three parts: Contraction, bottleneck, and the expansion parts. In the contraction part (encoder or downsampling path), analogous to the traditional CNN network, often takes in the input and outputs the feature maps through several convolutional (contraction blocks) operations via down-sampling thus preserving the contextual information. In between the contraction and the expansion path lies the bottleneck part. This part of the architecture accelerates the convolutional operations to increase the number of feature maps. At each step of the expansion path (decoder or upsampling path), a 2x2 convolution is used to half the number of feature maps via the upsampling layer to combine localization and contextual information critical for the precise prediction of segmentation maps. In other words, the decoder takes the feature maps (*output of encoder*) as input and estimates the best representations of the original input or the predicted output. Perhaps, one of the main innovations of U-Net is the introduction of skip connections. A diagram of the U-Net network architecture is illustrated in Figure 24 (Tsang, 2019).

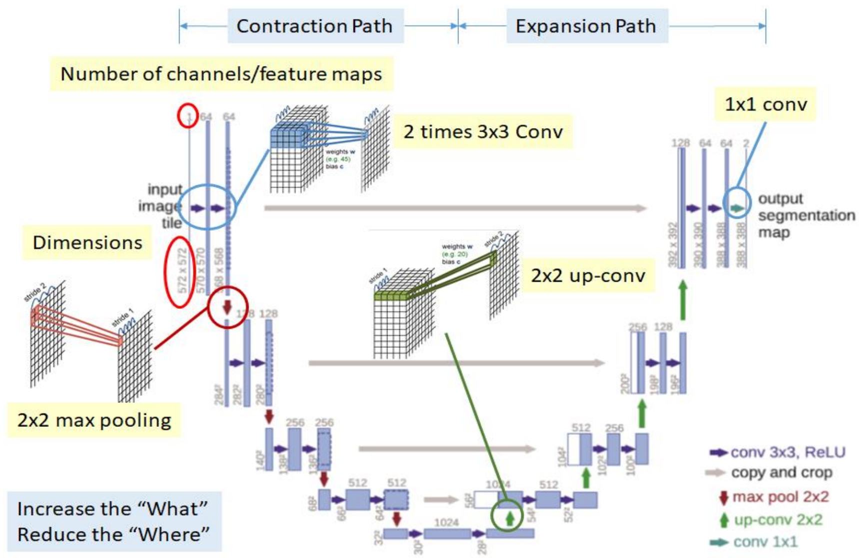


Figure 24. U-Net architecture

Although U-Net has found great success in the medical imaging domain, other achievements of applications in computer vision tasks have been documented in the literature (R. Li et al., 2017; Z. Zhang et al., 2018). The success and popularity of U-Net architecture can be attributed to several key observations. First, the model requires very few labeled datasets for training, unlike other DL models. Second, it lacks a fully connected layer making it attractive, and highly customizable to have multiple classes. Another benefit of the U-Net model is that there is no specific requirement to have a fixed size of an input image meaning different image sizes can be used to extract features.

Inspired by the latest literature in understanding more about domain adaptation, transfer learning, and the recent improvements in the use of autoencoders, we aim to examine the transferability of features within the transductive transfer learning approach on medical images. Therefore, through this work, we believe that transfer learning with fine-tuning on medical images leveraging domain adaptation can significantly increase learned feature interactions inside the networks for various performance vision tasks thus improving the goal of finding an optimal model that is robust and can perform across multiple but similar domains.

3.0 METHODOLOGY

In this section, we discuss the steps undertaken to train an autoencoder for the Chest X-ray disease segmentation of organs. Section 3.1 describes the datasets used in this work. Data pre-processing tasks are explained in section 3.2. Section 3.3 describes the environment used for model deployment.

3.1 Datasets

The primary datasets used in this work for both training and validation include: (1) The Japanese Society of Radiological Technology (JSRT) (Shiraishi et al., 2000) and, (2) The Montgomery and Shenzhen dataset (Candemir et al., 2014; Jaeger et al., 2014). The two datasets are publicly available and contain organ segmentation masks that have been carefully reviewed by radiologists. For our evaluation, we used the ChestX-ray14 dataset (Wang et al., 2017). These datasets are widely considered as the gold standard for Chest X-ray images.

The *JSRT dataset* comprises a total of 247 images with 154 of those images containing lung nodules. The dimensions of the X-ray images are 2048 x 2048 pixels with 12-bit grayscale intensity values. This dataset contains both the lung and heart segmentation masks for training purposes.

The *Montgomery dataset* consists of 138 Chest X-rays, from which 80 X-rays denote the healthy patients and the rest of the images represent patients with tuberculosis. The images are extracted from the Department of Health and Human Services, Montgomery County, Maryland, USA. The dimensions of the images vary from images with a resolution of 4020 x 4892 pixels and other images with 4892 x 4020 pixels consisting of 12-bit grayscale intensity values. Unlike the JSRT dataset, this dataset only contains the lung segmentation masks. All the images are in portable network graphics (*.png) format. Figure 25 illustrates two examples of Chest X-ray images from JSRT and Montgomery datasets (Dai et al., 2017). The upper half (two images) shows the JSRT Chest X-ray images while the bottom half (two images) represents images from the Montgomery Dataset.

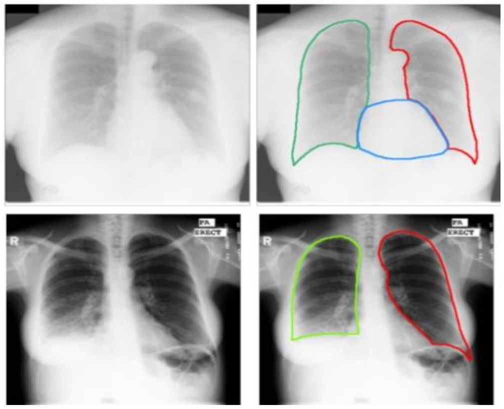


Figure 25. Sample images of JSRT and Shenzhen Chest X-rays

The *Shenzhen dataset* includes a total of 662 Chest X-ray images, from which 326 represent the healthy patients while the remaining images are of patients with tuberculosis. While the images are of high resolution, they vary in sizes but have an approximate resolution of 3000 x 3000 pixels with 8-bit grayscale intensity values. This dataset was collected from Shenzhen No. 3 People's Hospital in Shenzhen, China. Further, the Chest X-ray images were manually and carefully segmented by Stirenko et al.,(2018).

The *ChestX-ray14 dataset* is a large publicly available dataset containing a total of 112,120 frontal-view chest X-ray images extracted from 30,805 unique patients. The Chest X-ray images characterize 14 different classes of thorax diseases including Cardiomegaly, Consolidation, Edema, Emphysema, Effusion, Fibrosis, Hernia, Infiltration, Nodule, Mass, Pleural_thickening, Pneumonia, and Pneumothorax. The dimensions of the images are 1024 x 1024 pixels with 8-bit grayscale intensity values.

3.2 Data preprocessing

The X-ray images are usually in black and white format (i.e. grayscale form) containing a low amount of light representing the intensity of information which can be a difficult task to analyze. To tackle this problem, the contrast was limited to the adaptive histogram equalization (CLAHE) technique to ensure optimally balanced images. In other words, images that have high contrast ratios tend to lose subtle information. The goal is to have an image that contains the right number of contours which in effect removes some noise from the image and can improve the model's performance. Next, we combined both the JSRT and Shenzhen datasets, however, the masks were unequal to the total number of images. To resolve this problem, we performed a 1 to 1 mapping between the masks and the images. Further, a sanity check was conducted to ensure that the data correspondence reflected the mapping between the masks and images. Finally, the images were cropped and resized to a resolution of 512 x 512 pixels and mapped to the input layer of the U-Net architecture before the training step. Other preprocessing procedures that were introduced to improve the visualization of the chest radiographs included bone suppression techniques.

3.3 Environment

The U-Net network architecture was implemented using open-source deep learning libraries like Keras (Chollet & others, 2015) with TensorFlow (Abadi et al., 2016) as the back end. For the model training, we used Adam (Kingma & Lei, 2015) optimizer. The dataset was split in the 80:10:10 (train-test-validation) ratio using the Scikit machine learning library. The experiments and analyses were conducted on a 12 core Intel(R) Core(TM) CPU i7 with Dual NVIDIA TITAN V GPU consisting of a total of 28 GB memory.

4.0 EXPERIMENT AND RESULTS

4.1 Training and optimization

In this experiment, we implemented the U-Net architecture introduced by Ronneberger et al., (2015). We used the JSRT and Shenzhen datasets as inputs to the network for training purposes. The images were resized to 512 x 512 pixels and normalized. Afterward, the combined dataset was split into 8:1:1 ratio for training, validation, and test portions respectively. Next, the training process was initialized using Adam optimizer with a learning rate of 2e-4 and trained the datasets. We performed a multi-class image segmentation thereby adopting two evaluation metrics, for example, Dice Score and accuracy measures. The Dice score is a common metric that is widely used for binary segmentation tasks in the medical domain because of its robustness to class imbalance (Fidon et al., 2018). It can be expressed as follows:

$$\frac{2|P \cap G|}{|P|+|G|} = \frac{2|TP|}{2|TP|+|FP|+|FN|}$$

Where P is the pixels to be predicted in the segmentation mask for class G while G is the pixels of the ground truth mask belonging to the same segmentation class (Dai et al., 2017). Early stop mechanism was utilized on the validation set while where the training of the model stopped if no further improvement of the model was realized. Concurrently, we closely monitored the values of the Dice coefficient as well on the training and validation sets. To examine the accuracy of the model, the final model was saved and used to predict the Chest X-ray datasets. Figure 26 shows the visualization correspondence of images from the training and test.

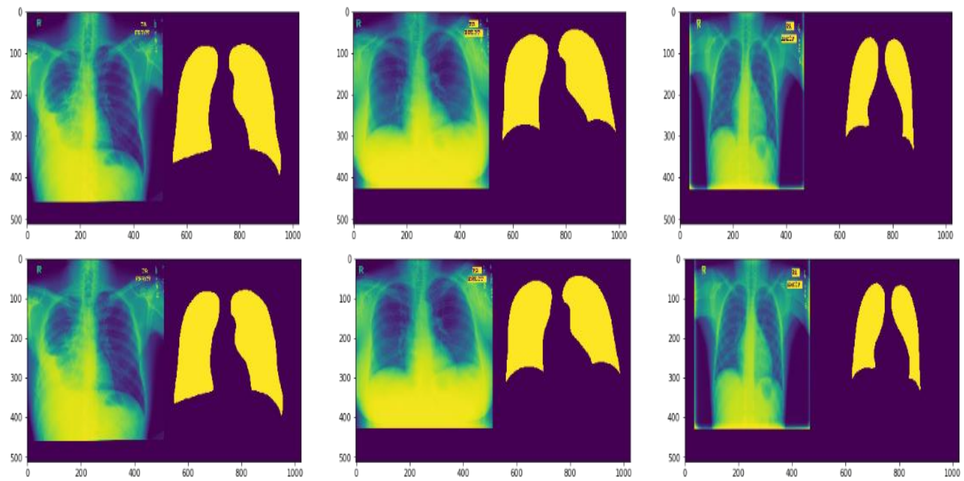


Figure 26. Training and validation results

Several training iterations (50 and 100 epochs) were performed on the training and validation datasets following the early stop mechanism protocol. The results from the accuracy and loss values are shown in Figure 27 and Figure 28. In this training phase, we achieved a Dice score of over 96% with an accuracy of about 98%. We found that applying data augmentation techniques such as the horizontal and vertical orientation of the images as well as increasing the contrast or brightness of the image during the training process had a marginal effect on the performance of the model.

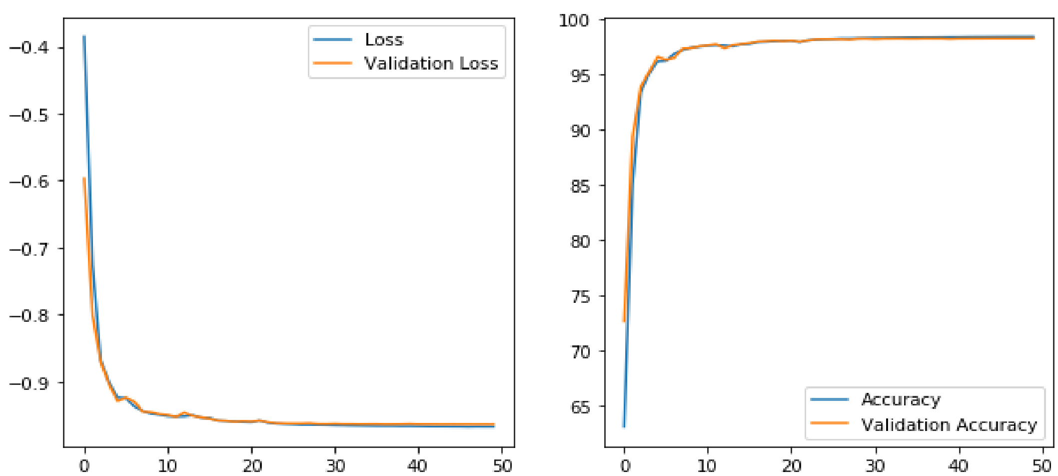


Figure 27. Training and validation accuracy @ 50 epochs

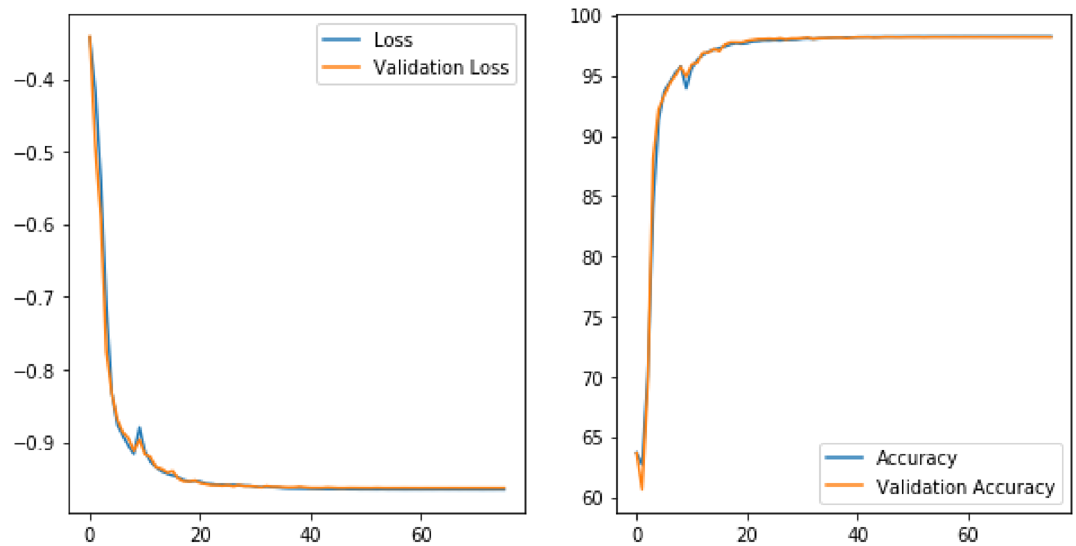


Figure 28. Training and validation accuracy @ 100 epochs

The final model was used on the test dataset and the results are shown in Figure 26. The evaluation results in Figure 29 shows the comparison between the ground truth and the predicted image. To visualize the predictions of the Chest X-ray dataset, we used sequential colormaps such as *bone* to represent much more information at pixel level intensity. Similarly, Figure 30 shows the results of the final model when predicted on the Chest X-ray dataset.

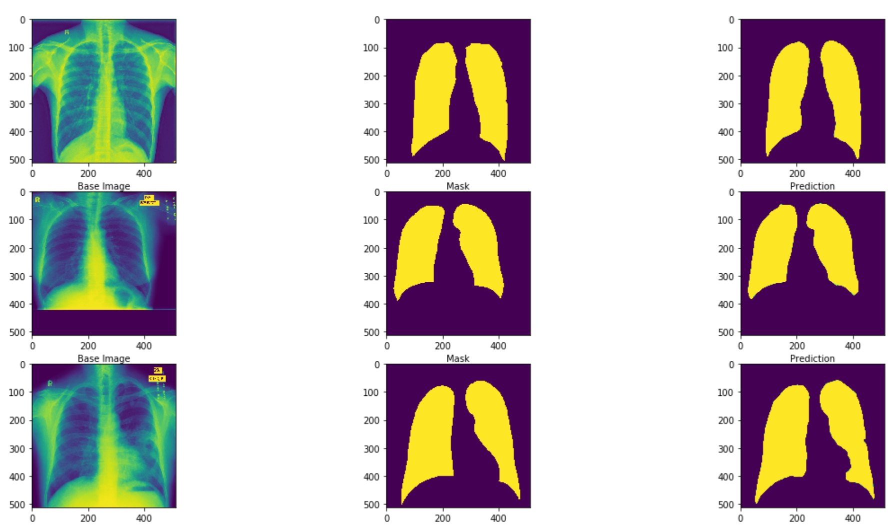


Figure 29 Results of the test model on training and validation set

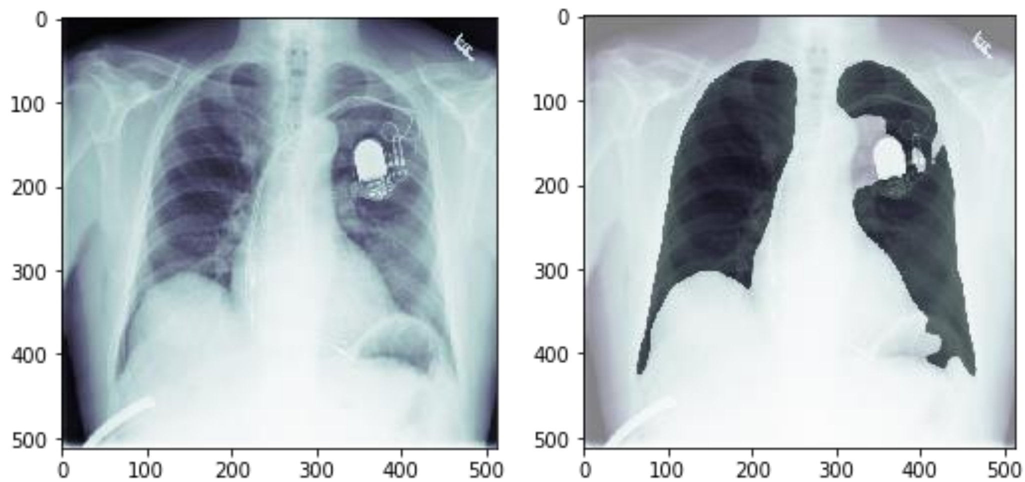


Figure 30. Prediction on the Chest X-ray dataset

4.2 Fine-tuning

To investigate the effectiveness of transfer learning techniques on the U-Net architecture, we split the network into two sections (contracting and expansion parts) for fine-tuning. First, we froze the expansion path (includes 10 convolution layers) and trained the network on the unfrozen part of the contraction path (includes 10 convolutions). The process was repeated for several convolutional blocks of the contraction path while noting the model performances. Each of the blocks for fine-tuning included two convolutional layers. Second, the same procedure was performed on the unfrozen expansion path by freezing the contraction path while fine-tuning several convolutional blocks with the JSRT and Shenzhen datasets. This experiment aimed to determine whether the shallow layers (contracting path) and the deeper layers (expansion path) offer significant differences in performance. Lastly, the prediction was also performed on the Chest X-ray to analyze the accuracy of the fine-tuned models. The findings indicated that indeed the higher-level layers of the U-Net architecture contributed towards marginal improvement in performance while we also observed accelerated training speeds when fine-tuning with different higher layers.

5.0 DISCUSSION

Currently, the most common imaging modality is CT and X-rays. These images provide the mechanism to rapidly diagnose and accurately evaluate lung-related diseases. In this work, we demonstrated the application of a transductive transfer learning approach where we have similar source and domain data for image analysis leveraging an autoencoder for domain adaptation. This approach to transfer learning is highly beneficial in instances where small amounts of labeled images are available for training. Furthermore, the lack of domain expertise in the medical imaging field, makes it an attractive proposition as well. In our experiments, we selected U-Net as the autoencoder for training and evaluation of the selected datasets. The results in Figure 27 and Figure 28, suggest that potential high performance can be achieved with a moderate amount of training epochs (approximately 30-60).

The training behavior of the U-Net architecture on the medical datasets shows a good balance between overfitting and underfitting, a phenomenon known as the bias-variance trade-off. In Figure 27, the model shows properties of convergence meaning the error moves closer to the local or global minimum. This property sometimes is referred to as *convexity* meaning the results follow a curved shape much like the exterior of a circle which represents the ideal scenario for model performance as a result of various hyperparameter tuning steps.

In the computer vision field, fine-tuning is currently one of the popular techniques used in transfer learning. We noticed that fine-tuning the deeper layers of the U-Net architecture yielded a marginal increase in performance which may provide opportunities for further investigations on performance improvements. However, in a recent work by Amiri et al.,(2020), a contrary observation was made when they fine-tuned a U-Net architecture on ultrasound images. The results showed that fine-tuning of the shallow layers surpassed the results from the fine-tuning of deeper layers. The findings in our work are consistent with the literature where performance increase can be observed when fine-tuning higher layers of a deep neural network rather than the juvenile layers (Choudhary & Hazra, 2019; J. Long et al., 2015).

6.0 CONCLUSION, LIMITATIONS AND FUTURE WORK

6.1 Conclusion

Autoencoders have presented a very promising opportunity to work on unlabeled datasets which is essential in the real-world. Notably, U-Net architecture, a very popular autoencoder in the medical imaging community, has proven its remarkable performance on many medical imaging tasks. In this work, we demonstrated that deep feature transfer can occur successfully when using transfer learning approaches leveraging an autoencoder (U-Net) while ensuring high model performance. Moreover, we selected two datasets (JSRT and Shenzhen datasets) for training and validation while the Chest X-ray dataset was used for evaluation. It is worth noting that the autoencoder was able to segment the Chest X-ray images even with low light intensity present across the images. While additional research would be desirable in improving the performance of the model even further, we achieved satisfactory performance using the U-Net based autoencoder in the experiment. The findings have great potential in the diagnosis of Chest X-ray related diseases via segmentation tasks.

6.2 Limitations

Despite the promising potential for autoencoders to overcome the problem of unlabeled datasets, this work had some limitations. First, the issue of data scarcity and class imbalance. The datasets we used for training and validation suffered from the issue of class imbalance. We had to apply data pre-processing tricks to match the images and masks thus reducing the total number of images available for training. Relevant images and masks may have been excluded during preprocessing which could affect the overall performance of the model. Also, domain-specific datasets currently remain limited for training purposes, therefore, we are restricted in using only publicly available labeled datasets to train many other medical images. Second, the issue of balancing the trade-offs between finding an effective model and computational efficiency remains a challenge. While determining an effective model requires rigorous experiments, the computational requirement for each step may affect experiments that require huge amounts of data. This technical issue deserves further investigation in future work.

6.3 Future work

Future research may focus on exploiting other methods such as attention U-Net adversarial architectures for Chest X-ray segmentations (Gaál et al., 2020) and other medical images to enhance consistency of the segmentation outputs. Other variants of U-Net, for example, U-Net++ (Z. Zhou et al., 2020) and MultiResUNET (Ibtehaz & Rahman, 2020) may be explored in the future to investigate multiscale features of chest X-rays in image segmentation tasks. Also, we aim to employ the multi-task learning technique to further improve the robustness and accuracy of the model.

The use of pretrained models is a popular technique in transfer learning. In this regard, a future research perspective may include using pretrained models like VGG-16 (Simonyan & Zisserman, 2015) and ResNet-50 (He et al., 2016) as base models to pretrain on the medical images using U-Net as the segmentor model. Additionally, we may extend this work by exploring recent works on light architectures by Google, for example, MobileNetV2 (Sandler et al., 2019), plugged in as an encoder part of the U-Net architecture.

In essence, with fewer ground truth images available for training, we would focus on generalizing the proposed model on other medical imaging tasks (e.g. skin lesions) from different imaging modalities such as CT scans. Also, the use of hybrid U-Net architectures may shed deeper insights on model performance, generalizability, and interpretability of medical imaging tasks thus enhancing the accurate diagnosis of diseases in clinical practice. Finally, unsupervised domain adaptation is a promising field that deserves further investigation to evaluate the performance of deep neural networks on unlabeled data in target domains.

REFERENCES

A. Romero Lopez, X. Giro-i-Nieto, J. Burdick, & O. Marques. (2017). Skin lesion classification from dermoscopic images using deep learning techniques. *2017 13th IASTED International Conference on Biomedical Engineering (BioMed)*, 49–54. https://doi.org/10.2316/P.2017.852-053

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., … Zheng, X. (2016). *TensorFlow: A system for large-scale machine learning*. 21.

Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2020). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *ArXiv:2003.13815 [Cs, Eess, Stat]*.

Agarwal, N. (2019). *Exploring Action Recognition without using Deep Learning*. 3.

Agrawal, T., Gupta, R., & Narayanan, S. (2019). On evaluating CNN representations for low resource medical image classification. *ArXiv:1903.11176 [Cs, Eess, Stat]*.

Altaf, F., Islam, S. M. S., Akhtar, N., & Janjua, N. K. (2019). Going Deep in Medical Image Analysis: Concepts, Methods, Challenges and Future Directions. *ArXiv:1902.05655 [Cs]*. http://arxiv.org/abs/1902.05655

Amin, J., Sharif, M., Yasmin, M., Saba, T., Anjum, M. A., & Fernandes, S. L. (2019). A New Approach for Brain Tumor Segmentation and Classification Based on Score Level Fusion Using Transfer Learning. *JOURNAL OF MEDICAL SYSTEMS*, *43*(11). https://doi.org/10.1007/s10916-019-1453-8

Amiri, M., Brooks, R., & Rivaz, H. (2020). Fine tuning U-Net for ultrasound image segmentation: Which layers? *ArXiv:2002.08438 [Cs, Eess]*. http://arxiv.org/abs/2002.08438

Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical Image Analysis using Convolutional Neural Networks: A Review. *Journal of Medical Systems*, *42*(11), 226. https://doi.org/10.1007/s10916-018-1088-1

Apostolopoulos, I. D., & Mpesiana, T. A. (2020). Covid-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*. https://doi.org/10.1007/s13246-020-00865-4

Awate, G. (2019). Detection of Alzheimers Disease from MRI using Convolutional Neural Networks, Exploring Transfer Learning And BellCNN. *ArXiv:1901.10231 [Cs, Eess]*.

Azizpour, H., Razavian, A. S., Sullivan, J., Maki, A., & Carlsson, S. (2016). Factors of Transferability for a Generic ConvNet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(9), 1790–1802. https://doi.org/10.1109/TPAMI.2015.2500224

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(12), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615

Baltruschat, I. M., Nickisch, H., Grass, M., Knopp, T., & Saalbach, A. (2018). Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification. *ArXiv:1803.02315 [Cs]*. http://arxiv.org/abs/1803.02315

Bar, Y., Diamant, I., Wolf, L., Lieberman, S., Konen, E., & Greenspan, H. (2015). Chest pathology detection using deep learning with non-medical training. *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 294–297. https://doi.org/10.1109/ISBI.2015.7163871

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, *5*(2), 157–166.

Bernheim, A., Mei, X., Huang, M., Yang, Y., Fayad, Z. A., Zhang, N., Diao, K., Lin, B., Zhu, X., Li, K., Li, S., Shan, H., Jacobi, A., & Chung, M. (2020). Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection. *Radiology*, *295*(3), 200463. https://doi.org/10.1148/radiol.2020200463

Bullock, J., Cuesta-Lazaro, C., & Quera-Bofarull, A. (2019). XNet: A convolutional neural network (CNN) implementation for medical X-Ray image segmentation suitable for small datasets. *Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 69. https://doi.org/10.1117/12.2512451

Burdick, J., Marques, O., Weinthal, J., & Furht, B. (2018). Rethinking Skin Lesion Segmentation in a Convolutional Classifier. *JOURNAL OF DIGITAL IMAGING*, *31*(4), 435–440. https://doi.org/10.1007/s10278-017-0026-y

Byra, M., Sznajder, T., Korzinek, D., Piotrzkowska-Wroblewska, H., Dobruch-Sobczak, K., Nowicki, A., & Marasek, K. (2018). Impact of ultrasound image reconstruction method on breast lesion classification with neural transfer learning. *ArXiv:1804.02119 [Physics]*.

Candemir, S., Jaeger, S., Palaniappan, K., Musco, J. P., Singh, R. K., Zhiyun Xue, Karargyris, A., Antani, S., Thoma, G., & McDonald, C. J. (2014). Lung Segmentation in Chest

Radiographs Using Anatomical Atlases With Nonrigid Registration. *IEEE Transactions on Medical Imaging*, *33*(2), 577–590. https://doi.org/10.1109/TMI.2013.2290491

Caruana, R. (1997). *Multitask Learning*. 35.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Chen, C., Dou, Q., Chen, H., & Heng, P.-A. (2018). Semantic-Aware Generative Adversarial Nets for Unsupervised Domain Adaptation in Chest X-ray Segmentation. *ArXiv:1806.00600 [Cs]*.

Chen, J., Wu, L., Zhang, J., Zhang, L., Gong, D., Zhao, Y., Hu, S., Wang, Y., Hu, X., Zheng, B., Zhang, K., Wu, H., Dong, Z., Xu, Y., Zhu, Y., Chen, X., Yu, L., & Yu, H. (2020). *Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: A prospective study* [Preprint]. Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.02.25.20021568

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2016). Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. *ArXiv:1412.7062 [Cs]*. http://arxiv.org/abs/1412.7062

Chen, Q., Hu, S., Long, P., Lu, F., Shi, Y., & Li, Y. (2019). A Transfer Learning Approach for Malignant Prostate Lesion Detection on Multiparametric MRI. *TECHNOLOGY IN CANCER RESEARCH & TREATMENT*, *18*. https://doi.org/10.1177/1533033819858363

Choi, J. Y., Yoo, T. K., Seo, J. G., Kwak, J., Um, T. T., & Rim, T. H. (2017). Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database. *PLOS ONE*, *12*(11). https://doi.org/10.1371/journal.pone.0187338

Chollet, F., & others. (2015). *Keras*. https://keras.io

Choudhary, P., & Hazra, A. (2019). Chest disease radiography in twofold: Using convolutional neural networks and transfer learning. *EVOLVING SYSTEMS*. https://doi.org/10.1007/s12530-019-09316-2

Christ, P. F., Ettlinger, F., Grün, F., Elshaera, M. E. A., Lipkova, J., Schlecht, S., Ahmaddy, F., Tatavarty, S., Bickel, M., Bilic, P., Rempfler, M., Hofmann, F., Anastasi, M. D., Ahmadi, S.-A., Kaissis, G., Holch, J., Sommer, W., Braren, R., Heinemann, V., & Menze, B. (2017). Automatic Liver and Tumor Segmentation of CT and MRI Volumes using Cascaded Fully Convolutional Neural Networks. *ArXiv:1702.05970 [Cs]*. http://arxiv.org/abs/1702.05970

Christodoulidis, S., Anthimopoulos, M., Ebner, L., Christe, A., & Mougiakakou, S. (2017). Multi-source Transfer Learning with Convolutional Neural Networks for Lung Pattern Analysis. *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS*, *21*(1), 76–84. https://doi.org/10.1109/JBHI.2016.2636929

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *ArXiv:1606.06650 [Cs]*. http://arxiv.org/abs/1606.06650

Ciga, O., Chen, J., & Martel, A. (2019). Multi-layer Domain Adaptation for Deep Convolutional Networks. *ArXiv:1909.02620 [Cs]*. http://arxiv.org/abs/1909.02620

Ciresan, D., Giusti, A., Gambardella, L. M., & Schmidhuber, J. (2012). *Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images*. 9.

Cook, D., Feuz, K. D., & Krishnan, N. C. (2013). Transfer learning for activity recognition: A survey. *Knowledge and Information Systems*, *36*(3), 537–556. https://doi.org/10.1007/s10115-013-0665-3

Csurka, G. (2017). Domain Adaptation for Visual Applications: A Comprehensive Survey. *ArXiv:1702.05374 [Cs]*. http://arxiv.org/abs/1702.05374

Dai, W., Doyle, J., Liang, X., Zhang, H., Dong, N., Li, Y., & Xing, E. P. (2017). SCAN: Structure Correcting Adversarial Network for Organ Segmentation in Chest X-rays. *ArXiv:1703.08770 [Cs]*. http://arxiv.org/abs/1703.08770

Dawud, A. M., Yurtkan, K., & Oztoprak, H. (2019). Application of Deep Learning in Neuroradiology: Brain Haemorrhage Classification Using Transfer Learning. *COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE*. https://doi.org/10.1155/2019/4629859

Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li, & Li Fei-Fei. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Doersch, C. (2016). Tutorial on Variational Autoencoders. *ArXiv:1606.05908 [Cs, Stat]*. http://arxiv.org/abs/1606.05908

Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., & Darrell, T. (2014). *DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition*. 9.

Dong, H., Yang, G., Liu, F., Mo, Y., & Guo, Y. (2017). Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. *ArXiv:1705.03820 [Cs]*. http://arxiv.org/abs/1705.03820

Dosovitskiy, A., Springenberg, J. T., Riedmiller, M., & Brox, T. (2014). *Discriminative Unsupervised Feature Learning with Convolutional Neural Networks*. 9.

Elangovan, A., & Jeyaseelan, T. (2016). Medical imaging modalities: A survey. *2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)*, 1–4. https://doi.org/10.1109/ICETETS.2016.7603066

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Farabet, C., Couprie, C., Najman, L., & LeCun, Y. (2013). Learning Hierarchical Features for Scene Labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1915–1929. https://doi.org/10.1109/TPAMI.2012.231

Feng, X., Bernard, M. E., Hunter, T., & Chen, Q. (2020). Improving accuracy and robustness of deep convolutional neural network based thoracic OAR segmentation. *PHYSICS IN MEDICINE AND BIOLOGY*, *65*(7). https://doi.org/10.1088/1361-6560/ab7877

Feng, Y., Lighter, D., Zhang, L., Wang, Y., & Dehghani, H. (2020). Application of deep neural networks to improve diagnostic accuracy of rheumatoid arthritis using diffuse optical tomography. *QUANTUM ELECTRONICS*, *50*(1), 21–32. https://doi.org/10.1070/QEL17177

Fidon, L., Li, W., Garcia-Peraza-Herrera, L. C., Ekanayake, J., Kitchen, N., Ourselin, S., & Vercauteren, T. (2018). Generalised Wasserstein Dice Score for Imbalanced Multi-class Segmentation using Holistic Convolutional Networks. *ArXiv:1707.00478 [Cs], 10670*, 64–76. https://doi.org/10.1007/978-3-319-75238-9_6

G. Liang, H. Hong, W. Xie, & L. Zheng. (2018). Combining Convolutional Neural Network With Recursive Neural Network for Blood Cell Image Classification. *IEEE ACCESS, 6*, 36188–36197. https://doi.org/10.1109/ACCESS.2018.2846685

Gaál, G., Maga, B., & Lukács, A. (2020). Attention U-Net Based Adversarial Architectures for Chest X-ray Lung Segmentation. *ArXiv:2003.10304 [Cs, Eess]*. http://arxiv.org/abs/2003.10304

García, S., Luengo, J., & Herrera, F. (2015). *Data Preprocessing in Data Mining* (Vol. 72). Springer International Publishing. https://doi.org/10.1007/978-3-319-10247-4

Ghafoorian, M., Mehrtash, A., Kapur, T., Karssemeijer, N., Marchiori, E., Pesteie, M., Guttmann, C. R. G., de Leeuw, F.-E., Tempany, C. M., van Ginneken, B., Fedorov, A., Abolmaesumi, P., Platel, B., & Wells III, W. M. (2017). Transfer Learning for Domain Adaptation in MRI: Application in Brain Lesion Segmentation. *ArXiv:1702.07841 [Cs], 10435*, 516–524. https://doi.org/10.1007/978-3-319-66179-7_59

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 580–587. https://doi.org/10.1109/CVPR.2014.81

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1). MIT press Cambridge.

Goodfellow, Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Networks. *ArXiv:1406.2661 [Cs, Stat]*. http://arxiv.org/abs/1406.2661

Gozes, O., Maayan, F.-A., Hayit, G., Patrick D, B., Huangqi, Z., & Wenbin, J. (2020). *Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis*. https://arxiv.org/ftp/arxiv/papers/2003/2003.05037.pdf

Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., & Yang, Y. (2018). Diagnose like a Radiologist: Attention Guided Convolutional Neural Network for Thorax Disease Classification. *ArXiv:1801.09927 [Cs]*. http://arxiv.org/abs/1801.09927

Guan, X., Lee, J., Wu, P., & Wu, Y. (2018). Machine Learning for Exam Triage. *ArXiv:1805.00503 [Cs]*. http://arxiv.org/abs/1805.00503

Guendel, S., Grbic, S., Georgescu, B., Zhou, K., Ritschl, L., Meier, A., & Comaniciu, D. (2018). Learning to recognize Abnormalities in Chest X-Rays with Location-Aware Dense Networks. *ArXiv:1803.04565 [Cs]*. http://arxiv.org/abs/1803.04565

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., & Webster, D. R. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, *316*(22), 2402. https://doi.org/10.1001/jama.2016.17216

Gupta, O., & Raskar, R. (2018). Distributed learning of deep neural network over multiple agents. *ArXiv:1810.06060 [Cs, Stat]*. http://arxiv.org/abs/1810.06060

Gupta, V., Demirer, M., Bigelow, M., Little, K. J., Candemir, S., Prevedello, L. M., White, R. D., O'Donnell, T. P., Wels, M., & Erdal, B. S. (2019). Performance of a Deep Neural Network Algorithm Based on a Small Medical Image Dataset: Incremental Impact of 3D-to-2D Reformation Combined with Novel Data Augmentation, Photometric Conversion, or Transfer Learning. *JOURNAL OF DIGITAL IMAGING*. https://doi.org/10.1007/s10278-019-00267-3

H. Oliveira, & J. dos Santos. (2018). Deep Transfer Learning for Segmentation of Anatomical Structures in Chest Radiographs. *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 204–211. https://doi.org/10.1109/SIBGRAPI.2018.00033

H. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, & R. M. Summers. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Transactions on Medical Imaging*, *35*(5), 1285–1298. https://doi.org/10.1109/TMI.2016.2528162

Hagos, M. T., & Kant, S. (2019). *Transfer Learning based Detection of Diabetic Retinopathy from Small Dataset*. 9.

Hariharan, B., Arbeláez, P., Girshick, R., & Malik, J. (2014). Simultaneous Detection and Segmentation. *ArXiv:1407.1808 [Cs]*. http://arxiv.org/abs/1407.1808

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hermessi, H., Mourali, O., & Zagrouba, E. (2018). Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain. *NEURAL*

COMPUTING & APPLICATIONS, *30*(7, SI), 2029–2045. https://doi.org/10.1007/s00521-018-3441-1

Hermessi, H., Mourali, O., & Zagrouba, E. (2019). Deep feature learning for soft tissue sarcoma classification in MR images via transfer learning. *EXPERT SYSTEMS WITH APPLICATIONS*, *120*, 116–127. https://doi.org/10.1016/j.eswa.2018.11.025

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *ArXiv:1704.04861 [Cs]*. http://arxiv.org/abs/1704.04861

Hu, Q., Whitney, H. M., & Giger, M. L. (2019). Transfer Learning in 4D for Breast Cancer Diagnosis using Dynamic Contrast-Enhanced Magnetic Resonance Imaging. *ArXiv:1911.03022 [Physics]*.

Huang, G., Liu, Z., Maaten, L. van der, & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2261–2269. https://doi.org/10.1109/CVPR.2017.243

Huh, M., Agrawal, P., & Efros, A. A. (2016). What makes ImageNet good for transfer learning? *ArXiv:1608.08614 [Cs]*. http://arxiv.org/abs/1608.08614

Hussein, S., Kandel, P., Bolan, C. W., Wallace, M. B., & Bagci, U. (2019). Lung and Pancreatic Tumor Characterization in the Deep Learning Era: Novel Supervised and Unsupervised Learning Approaches. *IEEE TRANSACTIONS ON MEDICAL IMAGING*, *38*(8), 1777–1787. https://doi.org/10.1109/TMI.2019.2894349

Huynh, B. Q., Li, H., & Giger, M. L. (2016). Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *JOURNAL OF MEDICAL IMAGING*, *3*(3). https://doi.org/10.1117/1.JMI.3.3.034501

Hwang, S., & Kim, H.-E. (2016). Self-Transfer Learning for Fully Weakly Supervised Object Localization. *ArXiv:1602.01625 [Cs]*.

Ibtehaz, N., & Rahman, M. S. (2020). MultiResUNet: Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation. *Neural Networks*, *121*, 74–87. https://doi.org/10.1016/j.neunet.2019.08.025

Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R., Shpanskaya, K., Seekins, J., Mong, D. A., Halabi, S. S., Sandberg, J. K., Jones, R., Larson, D. B., Langlotz, C. P., Patel, B. N., Lungren, M. P., & Ng, A. Y. (2019). CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison. *ArXiv:1901.07031 [Cs, Eess]*. http://arxiv.org/abs/1901.07031

J. Chang, J. Yu, T. Han, H. Chang, & E. Park. (2017). A method for classifying medical images using transfer learning: A pilot study on histopathology of breast cancer. *2017 IEEE 19th International Conference on E-Health Networking, Applications and Services (Healthcom)*, 1–4. https://doi.org/10.1109/HealthCom.2017.8210843

J. Sachdev, S. Shekhar, & S. Indu. (2018). Melanoma Screening Using Deep Neural Networks. *2018 3rd International Conference for Convergence in Technology (I2CT)*, 1–5. https://doi.org/10.1109/I2CT.2018.8529724

J. Sun, & A. Binder. (2017). Comparison of deep learning architectures for H&E histopathology images. *2017 IEEE Conference on Big Data and Analytics (ICBDA)*, 43–48. https://doi.org/10.1109/ICBDAA.2017.8284105

J. Xiong, X. Li, L. Lu, L. H. Schwartz, X. Fu, J. Zhao, & B. Zhao. (2019). Implementation Strategy of a CNN Model Affects the Performance of CT Assessment of EGFR Mutation Status in Lung Cancer Patients. *IEEE ACCESS*, *7*, 64583–64591. https://doi.org/10.1109/ACCESS.2019.2916557

J. Zhou, Z. Li, W. Zhi, B. Liang, D. Moses, & L. Dawes. (2017). Using Convolutional Neural Networks and Transfer Learning for Bone Age Classification. *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, 1–6. https://doi.org/10.1109/DICTA.2017.8227503

Jaeger, S., Candemir, S., Antani, S., Wáng, Y.-X. J., Lu, P.-X., & Thoma, G. (2014). *Two public chest X-ray datasets for computer-aided screening of pulmonary diseases*. 3.

Jin, S., Wang, B., Xu, H., Luo, C., Wei, L., Zhao, W., Hou, X., Ma, W., Xu, Z., Zheng, Z., Sun, W., Lan, L., Zhang, W., Mu, X., Shi, C., Wang, Z., Lee, J., Jin, Z., Lin, M., … Xu, W. (2020). *AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system in four weeks* [Preprint]. Health Informatics. https://doi.org/10.1101/2020.03.19.20039354

Jodeiri, A., Zoroofi, R. A., Hiasa, Y., Takao, M., Sugano, N., Sato, Y., & Otake, Y. (2019). Region-based Convolution Neural Network Approach for Accurate Segmentation of Pelvic Radiograph. *2019 26th National and 4th International Iranian Conference on Biomedical Engineering (ICBME)*, 152–157. https://doi.org/10.1109/ICBME49163.2019.9030401

Jordan, M. I. (1997). Serial Order: A Parallel Distributed Processing Approach. *Advances in Psychology*, *121*, 471–495.

Kang, J., & Gwak, J. (2019). Ensemble of Instance Segmentation Models for Polyp Segmentation in Colonoscopy Images. *IEEE ACCESS*, *7*, 26440–26447. https://doi.org/10.1109/ACCESS.2019.2900672

Karri, S. P. K., Chakraborty, D., & Chatterjee, J. (2017). Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. *BIOMEDICAL OPTICS EXPRESS*, *8*(2), 579–592. https://doi.org/10.1364/BOE.8.000579

Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep Learning Applications in Medical Image Analysis. *IEEE Access*, *6*, 9375–9389. https://doi.org/10.1109/ACCESS.2017.2788044

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y. L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., … Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, *172*(5), 1122-1131.e9. https://doi.org/10.1016/j.cell.2018.02.010

Kerrigan. (2020). *Infographic-Body-Systems-overall.jpg (1200×1598)*. https://www.carolina.com/images/teacher-resources/infographics/infographic-Body-Systems-overall.jpg

Kingma, D. P., & Lei, J. (2015). *Adam: A Method for Stochastic Optimization*. 15.

Kingma, D. P., & Welling, M. (2014). Auto-Encoding Variational Bayes. *ArXiv:1312.6114 [Cs, Stat]*. http://arxiv.org/abs/1312.6114

Kotzias, D., Denil, M., Blunsom, P., & de Freitas, N. (2014). Deep Multi-Instance Transfer Learning. *ArXiv:1411.3128 [Cs, Stat]*. http://arxiv.org/abs/1411.3128

Kouw, W. M., & Loog, M. (2019). An introduction to domain adaptation and transfer learning. *ArXiv:1812.11806 [Cs, Stat]*. http://arxiv.org/abs/1812.11806

Krizhevsky, A. (2009). *Learning Multiple Layers of Features from Tiny Images*. 60.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Kudva, V., Prasad, K., & Guruvare, S. (2019). Hybrid Transfer Learning for Classification of Uterine Cervix Images for Cervical Cancer Screening. *J Digit Imaging*. https://doi.org/10.1007/s10278-019-00269-1

L. M. Valério, D. H. A. Alves, L. F. Cruz, P. H. Bugatti, C. de Oliveira, & P. T. M. Saito. (2019). DeepMammo: Deep Transfer Learning for Lesion Classification of Mammographic Images. *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*, 447–452. https://doi.org/10.1109/CBMS.2019.00093

Lan, L., Ye, C., Wang, C., & Zhou, S. (2019). Deep Convolutional Neural Networks for WCE Abnormality Detection: CNN Architecture, Region Proposal and Transfer Learning. *IEEE ACCESS*, *7*, 30017–30032. https://doi.org/10.1109/ACCESS.2019.2901568

Lao, J., Chen, Y., Li, Z.-C., Li, Q., Zhang, J., Liu, J., & Zhai, G. (2017). A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme. *SCIENTIFIC REPORTS*, *7*. https://doi.org/10.1038/s41598-017-10649-8

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, *1*(4), 541–551.

LeCun, Y., Kavukcuoglu, K., & Farabet, C. (2010). Convolutional networks and applications in vision. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, 253–256. https://doi.org/10.1109/ISCAS.2010.5537907

Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., Bai, J., Lu, Y., Fang, Z., Song, Q., Cao, K., Liu, D., Wang, G., Xu, Q., Fang, X., Zhang, S., Xia, J., & Xia, J. (2020). Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT. *Radiology*, 200905. https://doi.org/10.1148/radiol.2020200905

Li, R., Liu, W., Yang, L., Sun, S., Hu, W., Zhang, F., & Li, W. (2017). DeepUNet: A Deep Fully Convolutional Network for Pixel-level Sea-Land Segmentation. *ArXiv:1709.00201 [Cs]*. http://arxiv.org/abs/1709.00201

Li, Z., Wang, C., Han, M., Xue, Y., Wei, W., Li, L.-J., & Fei-Fei, L. (2018). Thoracic Disease Identification and Localization with Limited Supervision. *ArXiv:1711.06373 [Cs, Stat]*. http://arxiv.org/abs/1711.06373

Liang, S., Zhang, R., Liang, D., Song, T., Ai, T., Xia, C., Xia, L., & Wang, Y. (2018). Multimodal 3D DenseNet for IDH Genotype Prediction in Gliomas. *GENES*, *9*(8). https://doi.org/10.3390/genes9080382

Lin, B. S., Michael, K., Kalra, S., & Tizhoosh, H. R. (2017). Skin Lesion Segmentation: U-Nets versus Clustering. *ArXiv:1710.01248 [Cs]*. http://arxiv.org/abs/1710.01248

Lin, M., Chen, Q., & Yan, S. (2014). Network In Network. *ArXiv:1312.4400 [Cs]*. http://arxiv.org/abs/1312.4400

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. *ArXiv:1411.4038 [Cs]*. http://arxiv.org/abs/1411.4038

Long, M., Cao, Y., Wang, J., & Jordan, M. I. (2015). *Learning Transferable Features with Deep Adaptation Networks*. 9.

Luo, L., Xiong, Y., Liu, Y., & Sun, X. (2019). Adaptive Gradient Methods with Dynamic Bound of Learning Rate. *ArXiv:1902.09843 [Cs, Stat]*. http://arxiv.org/abs/1902.09843

M. Ahsan, R. Gomes, & A. Denton. (2019). Application of a Convolutional Neural Network using transfer learning for tuberculosis detection. *2019 IEEE International Conference on Electro Information Technology (EIT)*, 427–433. https://doi.org/10.1109/EIT.2019.8833768

M. Ayyar, P. Mathur, R. R. Shah, & S. G. Sharma. (2018). Harnessing AI for Kidney Glomeruli Classification. *2018 IEEE International Symposium on Multimedia (ISM)*, 17–20. https://doi.org/10.1109/ISM.2018.00011

M. Miyagawa, M. G. F. Costa, M. A. Gutierrez, J. P. G. F. Costa, & C. F. F. Costa Filho. (2019). Detecting Vascular Bifurcation in IVOCT Images Using Convolutional Neural Networks With Transfer Learning. *IEEE ACCESS*, *7*, 66167–66175. https://doi.org/10.1109/ACCESS.2019.2918017

Ma, Y., & Peng, Y. (2019). Lymph node detection method based on multisource transfer learning and convolutional neural network. *INTERNATIONAL JOURNAL OF IMAGING SYSTEMS AND TECHNOLOGY*. https://doi.org/10.1002/ima.22371

Mahbod, A., Schaefer, G., Wang, C., Dorffner, G., Ecker, R., & Ellinger, I. (2020). Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification. *Comput Methods Programs Biomed*, *193*, 105475. https://doi.org/10.1016/j.cmpb.2020.105475

mathworks. (2020). *Transfer learning workflow*. https://duckduckgo.com/?q=transfer+learning+workflow&iax=images&ia=images&iai=https%3A%2F%2Fwww.mathworks.com%2Fdiscovery%2Ftransfer-learning%2F_jcr_content%2FmainParsys%2Fimage_226673988.adapt.full.high.jpg%2F1576731630748.jpg

Mazurowski, M. A., Buda, M., Saha, A., & Bashir, M. R. (2018). Deep learning in radiology: An overview of the concepts and a survey of the state of the art. *ArXiv:1802.08717 [Cs, Stat]*. http://arxiv.org/abs/1802.08717

McCann, B., Keskar, N. S., Xiong, C., & Socher, R. (2018). The Natural Language Decathlon: Multitask Learning as Question Answering. *ArXiv:1806.08730 [Cs, Stat]*. http://arxiv.org/abs/1806.08730

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. y. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. *ArXiv:1602.05629 [Cs]*. http://arxiv.org/abs/1602.05629

Menegola, A., Fornaciali, M., Pires, R., Avila, S., & Valle, E. (2016). Towards Automated Melanoma Screening: Exploring Transfer Learning Schemes. *ArXiv:1609.01228 [Cs]*.

Milletari, F., Navab, N., & Ahmadi, S.-A. (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *ArXiv:1606.04797 [Cs]*. http://arxiv.org/abs/1606.04797

Ming Liang, & Xiaolin Hu. (2015). Recurrent convolutional neural network for object recognition. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3367–3375. https://doi.org/10.1109/CVPR.2015.7298958

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Motamed, S., Gujrathi, I., Deniffel, D., Oentoro, A., Haider, M. A., & Khalvati, F. (2019). A Transfer Learning Approach for Automated Segmentation of Prostate Whole Gland and Transition Zone in Diffusion Weighted MRI. *ArXiv:1909.09541 [Cs, Eess, q-Bio]*.

Motozawa, N., An, G., Takagi, S., Kitahata, S., Mandai, M., Hirami, Y., Yokota, H., Akiba, M., Tsujikawa, A., Takahashi, M., & Kurimoto, Y. (2019). Optical Coherence Tomography-Based Deep-Learning Models for Classifying Normal and Age-Related Macular Degeneration and Exudative and Non-Exudative Age-Related Macular Degeneration Changes. *OPHTHALMOLOGY AND THERAPY*, *8*(4), 527–539. https://doi.org/10.1007/s40123-019-00207-y

Nadimi, E. S., Buijs, M. M., Herp, J., Kroijer, R., Kobaek-Larsen, M., Nielsen, E., Pedersen, C. D., Blanes-Vidal, V., & Baatrup, G. (2020). Application of deep learning for autonomous detection and localization of colorectal polyps in wireless colon capsule endoscopy. *COMPUTERS & ELECTRICAL ENGINEERING*, *81*. https://doi.org/10.1016/j.compeleceng.2019.106531

O. Hadad, R. Bakalo, R. Ben-Ari, S. Hashoul, & G. Amit. (2017). Classification of breast lesions using cross-modal deep learning. *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 109–112. https://doi.org/10.1109/ISBI.2017.7950480

Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2014). Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 1717–1724. https://doi.org/10.1109/CVPR.2014.222

Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., & Lerer, A. (2017). *Automatic differentiation in PyTorch*. 4.

Patel, V. M., Gopalan, R., Li, R., & Chellappa, R. (2015a). Visual Domain Adaptation: A survey of recent advances. *IEEE Signal Processing Magazine*, *32*(3), 53–69. https://doi.org/10.1109/MSP.2014.2347059

Patel, V. M., Gopalan, R., Li, R., & Chellappa, R. (2015b). Visual Domain Adaptation: A survey of recent advances. *IEEE Signal Processing Magazine*, *32*(3), 53–69. https://doi.org/10.1109/MSP.2014.2347059

Patrini, I., Ruperti, M., Moccia, S., Mattos, L. S., Frontoni, E., & De Momi, E. (2020). Transfer learning for informative-frame selection in laryngoscopic videos through learned features. *MEDICAL & BIOLOGICAL ENGINEERING & COMPUTING*. https://doi.org/10.1007/s11517-020-02127-7

Paul, R., Hawkins, S. H., Balagurunathan, Y., Schabath, M. B., Gillies, R. J., Hall, L. O., & Goldgof, D. B. (2016). Deep Feature Transfer Learning in Combination with Traditional Features Predicts Survival Among Patients with Lung Adenocarcinoma. *TOMOGRAPHY*, *2*(4), 388–395. https://doi.org/10.18383/j.tom.2016.00211

Pehrson, L., Nielsen, M., & Ammitzbøl Lauridsen, C. (2019). Automatic Pulmonary Nodule Detection Applying Deep Learning or Machine Learning Algorithms to the LIDC-IDRI Database: A Systematic Review. *Diagnostics*, *9*(1), 29. https://doi.org/10.3390/diagnostics9010029

Perez Malla, C. U., Valdes Hernandez, M. del C., Rachmadi, M. F., & Komura, T. (2019). Evaluation of Enhanced Learning Techniques for Segmenting Ischaemic Stroke Lesions in Brain Magnetic Resonance Perfusion Images Using a Convolutional Neura Network Scheme. *FRONTIERS IN NEUROINFORMATICS*, *13*. https://doi.org/10.3389/fninf.2019.00033

Ponzio, F., Urgese, G., Ficarra, E., & Di Cataldo, S. (2019). Dealing with Lack of Training Data for Convolutional Neural Networks: The Case of Digital Pathology. *ELECTRONICS*, *8*(3). https://doi.org/10.3390/electronics8030256

Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., Shyu, M.-L., Chen, S.-C., & Iyengar, S. S. (2018). A Survey on Deep Learning: Algorithms, Techniques, and Applications. *ACM Computing Surveys*, *51*(5), 1–36. https://doi.org/10.1145/3234150

R. V. M. d. Nóbrega, S. A. Peixoto, S. P. P. da Silva, & P. P. R. Filho. (2018). Lung Nodule Classification via Deep Transfer Learning in CT Lung Images. *2018 IEEE 31st*

International Symposium on Computer-Based Medical Systems (CBMS), 244–249. https://doi.org/10.1109/CBMS.2018.00050

Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding Transfer Learning for Medical Imaging. *ArXiv:1902.07208 [Cs, Stat]*.

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. *ArXiv:1711.05225 [Cs, Stat]*. http://arxiv.org/abs/1711.05225

Ramsundar, B., Kearnes, S., Riley, P., Webster, D., Konerding, D., & Pande, V. (2015). Massively Multitask Networks for Drug Discovery. *ArXiv:1502.02072 [Cs, Stat]*. http://arxiv.org/abs/1502.02072

Ravishankar, H., Sudhakar, P., Venkataramani, R., Thiruvenkadam, S., Annangi, P., Babu, N., & Vaidya, V. (2017). Understanding the Mechanisms of Deep Transfer Learning for Medical Images. *ArXiv:1704.06040 [Cs]*.

Rawat, W., & Wang, Z. (2017). Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*, *29*(9), 2352–2449. https://doi.org/10.1162/neco_a_00990

Razavian, A. S., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN Features Off-the-Shelf: An Astounding Baseline for Recognition. *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 512–519. https://doi.org/10.1109/CVPRW.2014.131

Rettner. (2016). *The Human Body: Anatomy, Facts & Functions*. Livescience.Com. https://www.livescience.com/37009-human-body.html

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *ArXiv:1505.04597 [Cs]*. http://arxiv.org/abs/1505.04597

Ruder, S. (2017). An Overview of Multi-Task Learning in Deep Neural Networks. *ArXiv:1706.05098 [Cs, Stat]*. http://arxiv.org/abs/1706.05098

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *ArXiv:1409.0575 [Cs]*. http://arxiv.org/abs/1409.0575

S. Guan, & M. Loew. (2017). Breast Cancer Detection Using Transfer Learning in Convolutional Neural Networks. *2017 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 1–8. https://doi.org/10.1109/AIPR.2017.8457948

S. Suzuki, X. Zhang, N. Homma, K. Ichiji, N. Sugita, Y. Kawasumi, T. Ishibashi, & M. Yoshizawa. (2016). Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis. *2016 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, 1382–1386. https://doi.org/10.1109/SICE.2016.7749265

Sahiner, B., Pezeshk, A., Hadjiiski, L. M., Wang, X., Drukker, K., Cha, K. H., Summers, R. M., & Giger, M. L. (2019). Deep learning in medical imaging and radiation therapy. *Medical Physics*, *46*(1), e1–e36. https://doi.org/10.1002/mp.13264

Samala, R. K., Chan, H.-P., Hadjiiski, L. M., Helvie, M. A., Cha, K. H., & Richter, C. D. (2017). Multi-task transfer learning deep convolutional neural network: Application to computer-aided diagnosis of breast cancer on mammograms. *PHYSICS IN MEDICINE AND BIOLOGY*, *62*(23), 8894–8908. https://doi.org/10.1088/1361-6560/aa93d4

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2019). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *ArXiv:1801.04381 [Cs]*. http://arxiv.org/abs/1801.04381

Sedai, S., Mahapatra, D., Ge, Z., Chakravorty, R., & Garnavi, R. (2018). Deep multiscale convolutional feature learning for weakly supervised localization of chest pathologies in X-ray images. *ArXiv:1808.08280 [Cs, Stat]*. http://arxiv.org/abs/1808.08280

Sekou, T. B., Hidane, M., Olivier, J., & Cardot, H. (2019). From Patch to Image Segmentation using Fully Convolutional Networks—Application to Retinal Images. *ArXiv:1904.03892 [Cs]*.

Sengupta, S., Basak, S., Saikia, P., Paul, S., Tsalavoutis, V., Atiah, F., Ravi, V., & Peters, A. (2019). A Review of Deep Learning with Special Emphasis on Architectures, Applications and Recent Trends. *ArXiv:1905.13294 [Cs, Stat]*. http://arxiv.org/abs/1905.13294

Serte, S., & Demirel, H. (2020). Wavelet-based deep learning for skin lesion classification. *IET IMAGE PROCESSING*, *14*(4), 720–726. https://doi.org/10.1049/iet-ipr.2019.0553

Setio, A. A. A., Traverso, A., de Bel, T., Berens, M. S. N., Bogaard, C. van den, Cerello, P., Chen, H., Dou, Q., Fantacci, M. E., Geurts, B., van der Gugten, R., Heng, P. A., Jansen, B., de Kaste, M. M. J., Kotov, V., Lin, J. Y.-H., Manders, J. T. M. C., Sónora-Mengana, A., García-Naranjo, J. C., … Jacobs, C. (2017). Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Medical Image Analysis*, *42*, 1–13. https://doi.org/10.1016/j.media.2017.06.015

Shie, C.-K., Chuang, C.-H., Chou, C.-N., Wu, M.-H., & Chang, E. Y. (2015). Transfer representation learning for medical image analysis. *Conf Proc IEEE Eng Med Biol Soc*, *2015*, 711–714. https://doi.org/10.1109/EMBC.2015.7318461

Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016a). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *ArXiv:1602.03409 [Cs]*. http://arxiv.org/abs/1602.03409

Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016b). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *ArXiv:1602.03409 [Cs]*. http://arxiv.org/abs/1602.03409

Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K., Matsui, M., Fujita, H., Kodera, Y., & Doi, K. (2000). Development of a Digital Image Database for Chest Radiographs With and Without a Lung Nodule: Receiver Operating Characteristic Analysis of Radiologists' Detection of Pulmonary Nodules. *American Journal of Roentgenology*, *174*(1), 71–74. https://doi.org/10.2214/ajr.174.1.1740071

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, *6*(1), 60. https://doi.org/10.1186/s40537-019-0197-0

Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. *ArXiv:1409.1556 [Cs]*. http://arxiv.org/abs/1409.1556

Sirinukunwattana, K., Pluim, J. P. W., Chen, H., Qi, X., Heng, P.-A., Guo, Y. B., Wang, L. Y., Matuszewski, B. J., Bruni, E., Sanchez, U., Böhm, A., Ronneberger, O., Cheikh,

B. B., Racoceanu, D., Kainz, P., Pfeiffer, M., Urschler, M., Snead, D. R. J., & Rajpoot, N. M. (2016). Gland Segmentation in Colon Histology Images: The GlaS Challenge Contest. *ArXiv:1603.00275 [Cs]*. http://arxiv.org/abs/1603.00275

Sital, C., Brosch, T., Tio, D., Raaijmakers, A., & Weese, J. (2020). 3D medical image segmentation with labeled and unlabeled data using autoencoders at the example of liver segmentation in CT images. *ArXiv:2003.07923 [Cs, Eess]*.

Speicher, A. (2019, September 17). Different Imaging Tests Explained | UVA Radiology. *UVA Radiology and Medical Imaging Blog for Patients*. https://blog.radiology.virginia.edu/different-imaging-tests-explained/

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*. 30.

Stirenko, S., Kochura, Y., Alienin, O., Rokovyi, O., Gordienko, Y., Gang, P., & Zeng, W. (2018). Chest X-Ray Analysis of Tuberculosis by Deep Learning with Segmentation and Augmentation. *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*, 422–428. https://doi.org/10.1109/ELNANO.2018.8477564

Sudheer Kumar, E., & Shoba Bindu, C. (2019). Medical Image Analysis Using Deep Learning: A Systematic Literature Review. In A. K. Somani, S. Ramakrishna, A. Chaudhary, C. Choudhary, & B. Agarwal (Eds.), *Emerging Technologies in Computer Engineering: Microservices in Big Data Analytics* (Vol. 985, pp. 81–97). Springer Singapore. https://doi.org/10.1007/978-981-13-8300-7_8

Sun, S., Shi, H., & Wu, Y. (2015). A survey of multi-source domain adaptation. *Information Fusion*, *24*, 84–92. https://doi.org/10.1016/j.inffus.2014.12.003

Suzuki, A., Sakanashi, H., Kido, S., & Shouno, H. (2018). Feature Representation Analysis of Deep Convolutional Neural Network using Two-stage Feature Transfer -An Application for Diffuse Lung Disease Classification. *ArXiv:1810.06282 [Cs]*.

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2016). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *ArXiv:1602.07261 [Cs]*. http://arxiv.org/abs/1602.07261

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2818–2826. https://doi.org/10.1109/CVPR.2016.308

Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

T. Han, V. X. Nunes, L. F. D. F. Souza, A. G. Marques, I. C. L. Silva, M. A. A, F., J. Sun, & P. P. R. Filho. (2020). Internet of Medical Things-based Deep Learning Techniques for Segmentation of Lung and Stroke Regions in CT Scans. *IEEE ACCESS*, 1–1. https://doi.org/10.1109/ACCESS.2020.2987932

Talo, M. (2019). Automated classification of histopathology images using transfer learning. *ARTIFICIAL INTELLIGENCE IN MEDICINE*, *101*. https://doi.org/10.1016/j.artmed.2019.101743

Thrun, S. (1996). *Is Learning The n-th Thing Any Easier Than Learning The First?* 7.

Tiulpin, A., Melekhov, I., & Saarakkala, S. (2019). KNEEL: Knee Anatomical Landmark Localization Using Hourglass Networks. *ArXiv:1907.12237 [Cs]*.

Tsang, S.-H. (2019, March 20). *Review: U-Net (Biomedical Image Segmentation)*. Medium. https://towardsdatascience.com/review-u-net-biomedical-image-segmentation-d02bf06ca760

U. Gaur, M. Kourakis, E. Newman-Smith, W. Smith, & B. S. Manjunath. (2016). Membrane segmentation via active learning with deep networks. *2016 IEEE International Conference on Image Processing (ICIP)*, 1943–1947. https://doi.org/10.1109/ICIP.2016.7532697

W. O'Quinn, R. J. Haddad, & D. L. Moore. (2019). Pneumonia Radiograph Diagnosis Utilizing Deep Learning Network. *2019 IEEE 2nd International Conference on Electronic Information and Communication Technology (ICEICT)*, 763–767. https://doi.org/10.1109/ICEICT.2019.8846438

Wang, C., Elazab, A., Jia, F., Wu, J., & Hu, Q. (2018). Automated chest screening based on a hybrid model of transfer learning and convolutional sparse denoising autoencoder. *BIOMEDICAL ENGINEERING ONLINE*, *17*. https://doi.org/10.1186/s12938-018-0496-2

Wang, Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3462–3471. https://doi.org/10.1109/CVPR.2017.369

Waymo. (2019). *Technology – Waymo*. Technology. https://waymo.com/tech/

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, *3*(1), 9. https://doi.org/10.1186/s40537-016-0043-6

WHO. (2020a). *The top 10 causes of death*. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

WHO. (2020b). *WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020*. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

WHO, & others. (2020). *Coronavirus disease 2019 (COVID-19): Situation report, 72*.

Wilson, G., & Cook, D. J. (2020). A Survey of Unsupervised Deep Domain Adaptation. *ArXiv:1812.02849 [Cs, Stat]*. http://arxiv.org/abs/1812.02849

Wingate, J., Kollia, I., Bidaut, L., & Kollias, S. (2019). A Unified Deep Learning Approach for Prediction of Parkinson's Disease. *ArXiv:1911.10653 [Cs, Eess, Stat]*.

Wu, Z., Zhao, S., Peng, Y., He, X., Zhao, X., Huang, K., Wu, X., Fan, W., Li, F., Chen, M., Li, J., Huang, W., Chen, X., & Li, Y. (2019). Studies on Different CNN Algorithms for Face Skin Disease Classification Based on Clinical Images. *IEEE ACCESS*, *7*, 66505–66511. https://doi.org/10.1109/ACCESS.2019.2918221

X. Li, T. Pang, B. Xiong, W. Liu, P. Liang, & T. Wang. (2017). Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification. *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 1–11. https://doi.org/10.1109/CISP-BMEI.2017.8301998

X. Xu, J. Lin, Y. Tao, & X. Wang. (2018). An Improved DenseNet Method Based on Transfer Learning for Fundus Medical Images. *2018 7th International Conference on Digital Home (ICDH)*, 137–140. https://doi.org/10.1109/ICDH.2018.00033

Y. A. L. Alsabahi, L. Fan, & X. Feng. (2018). Image Classification Method in DR Image Based on Transfer Learning. *2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1–4. https://doi.org/10.1109/IPTA.2018.8608157

Y. Sawada, & K. Kozuka. (2015). Transfer learning method using multi-prediction deep Boltzmann machines for a small scale dataset. *2015 14th IAPR International Conference on Machine Vision Applications (MVA)*, 110–113. https://doi.org/10.1109/MVA.2015.7153145

Yan, C., Yao, J., Li, R., Xu, Z., & Huang, J. (2018). Weakly Supervised Deep Learning for Thoracic Disease Classification and Localization on Chest X-rays. *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics - BCB '18*, 103–110. https://doi.org/10.1145/3233547.3233573

Yao, L., Prosky, J., Poblenz, E., Covington, B., & Lyman, K. (2018). Weakly Supervised Medical Diagnosis and Localization from Multiple Resolutions. *ArXiv:1803.07703 [Cs]*. http://arxiv.org/abs/1803.07703

Yi, X., Walia, E., & Babyn, P. (2019). Generative Adversarial Network in Medical Imaging: A Review. *Medical Image Analysis*, *58*, 101552. https://doi.org/10.1016/j.media.2019.101552

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). *How transferable are features in deep neural networks?* 9.

Yu, L., Yang, X., Chen, H., Qin, J., & Heng, P.-A. (2017). *Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images*. 7.

Zeng, G., Yang, X., Li, J., Yu, L., Heng, P.-A., & Zheng, G. (2017). 3D U-net with Multi-level Deep Supervision: Fully Automatic Segmentation of Proximal Femur in 3D MR Images. In Q. Wang, Y. Shi, H.-I. Suk, & K. Suzuki (Eds.), *Machine Learning in Medical Imaging* (Vol. 10541, pp. 274–282). Springer International Publishing. https://doi.org/10.1007/978-3-319-67389-9_32

Zhang, J., Li, Q., Caselli, R. J., Ye, J., & Wang, Y. (2017). Multi-task Dictionary Learning based Convolutional Neural Network for Computer aided Diagnosis with Longitudinal Images. *ArXiv:1709.00042 [Cs]*.

Zhang, S., Sun, F., Wang, N., Zhang, C., Yu, Q., Zhang, M., Babyn, P., & Zhong, H. (2019). Computer-Aided Diagnosis (CAD) of Pulmonary Nodule of Thoracic CT Image Using Transfer Learning. *JOURNAL OF DIGITAL IMAGING*, *32*(6), 995–1007. https://doi.org/10.1007/s10278-019-00204-4

Zhang, Y., & Yang, Q. (2018). A Survey on Multi-Task Learning. *ArXiv:1707.08114 [Cs]*. http://arxiv.org/abs/1707.08114

Zhang, Z., Liu, Q., & Wang, Y. (2018). Road Extraction by Deep Residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, *15*(5), 749–753. https://doi.org/10.1109/LGRS.2018.2802944

Zhao, S., Li, B., Reed, C., Xu, P., & Keutzer, K. (2020). Multi-source Domain Adaptation in the Deep Learning Era: A Systematic Survey. *ArXiv:2002.12169 [Cs, Stat]*. http://arxiv.org/abs/2002.12169

Zhao, X. (2020). *A collection of awesome things about domain-adaptation*. https://github.com/zhaoxin94/awesome-domain-adaptation

Zheng, C., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., Liu, W., & Wang, X. (2020). *Deep Learning-based Detection for COVID-19 from Chest CT using Weak Label* [Preprint]. Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.03.12.20027185

Zhou, Bo, Li, Y., & Wang, J. (2018). A Weakly Supervised Adaptive DenseNet for Classifying Thoracic Diseases and Identifying Abnormalities. *ArXiv:1807.01257 [Cs]*. http://arxiv.org/abs/1807.01257

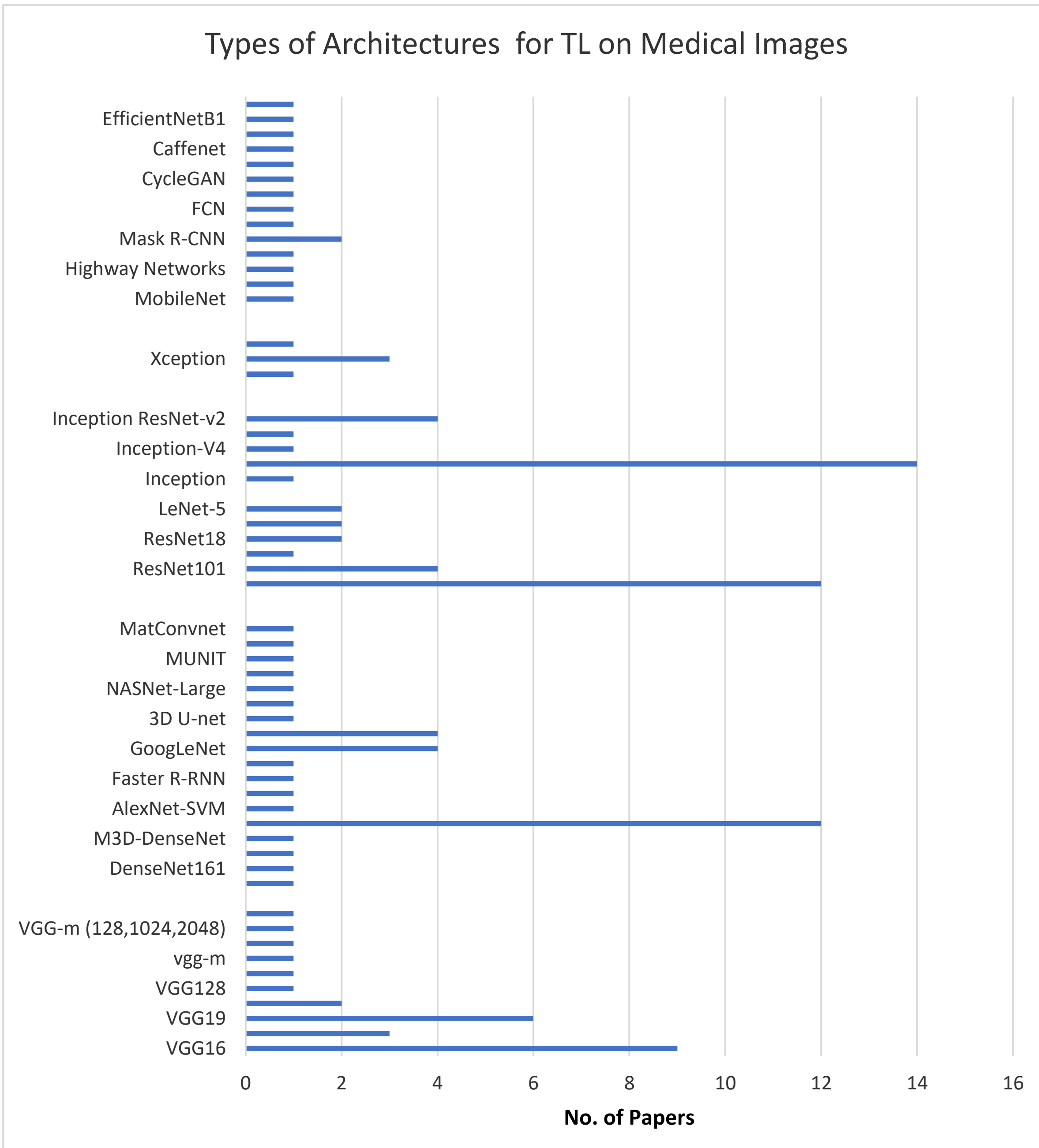
Zhou, Bolei, Lapedriza, A., Khosla, A., Oliva, A., & Torralba, A. (2018). Places: A 10 Million Image Database for Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(6), 1452–1464. https://doi.org/10.1109/TPAMI.2017.2723009

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2020). UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *ArXiv:1912.05074 [Cs, Eess]*. http://arxiv.org/abs/1912.05074

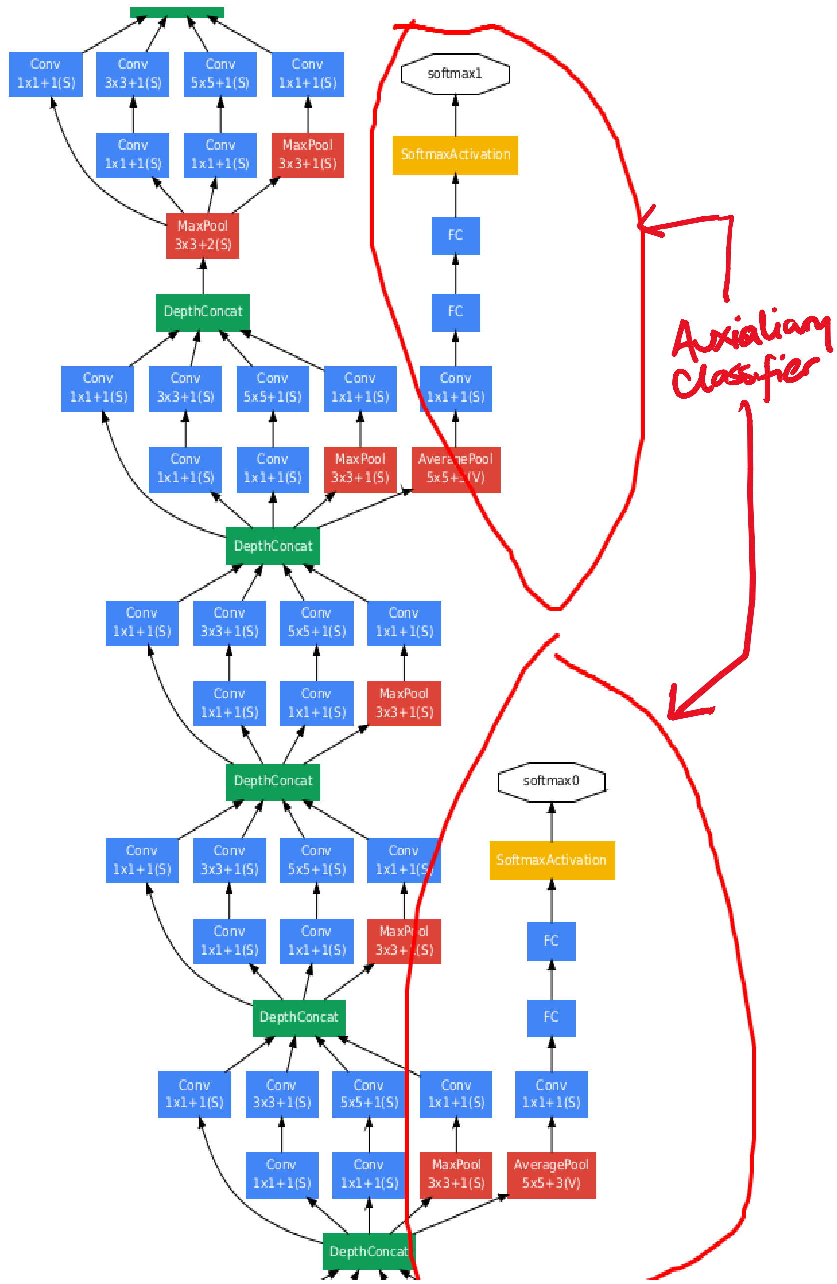
Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2019). A Comprehensive Survey on Transfer Learning. *ArXiv:1911.02685 [Cs, Stat]*. http://arxiv.org/abs/1911.02685

APPENDICES

Appendix 1: Architectures used with TL for Medical Image analysis



Appendix 2: Inception network showing auxiliary classifiers



Appendix 3: DIM V1 Architecture