Dermatology AI: Identifying Skin Injuries and Determining Burn Degrees using Deep Learning

Abdullaziz Alosayl [411109523],
Mohammed Alrushud[411107962],
Anas Shawky Alarbied [411116054],
Abdullah Huthlul Alhuthlul [411107932]

A project report submitted to Qassim University in partial fulfilment
of the requirements for the degree
of BSc in Computer Science
at the Department of Computer Science
College of Computer
Qassim University
Saudi Arabia

October 2023

Supervisor Dr. Mohammed Alsuhaibani

Abstract

The Abstract of the report should be written here, it should provide a short summary of the work encompassing no more than 300 words.

Dedication

We would like to dedicate this project to ....

Acknowledgements

The Acknowledgements section may be used to thank your supervisor, family, research funding bodies, or any other applicable individuals or institutions.

Declaration

declaration

Contents

List of Tables

List of Figures

List of Algorithms

Listings

Chapter 1

Introduction

1.1 Introduction

Skin burns, resulting from exposure to extreme temperatures, are among the most common and painful of all injury types. Globally, millions suffer from burn injuries annually, with a significant proportion experiencing long-term physical and psychological consequences. Timely and accurate diagnosis of burn severity is not just pivotal but is often the linchpin in ensuring optimal therapeutic outcomes [1].

However, a notable challenge in the field of burn treatment is the diagnosis of burn severity. Traditionally, healthcare professionals have depended on visual examinations to measure burn severity. This approach involves judging the burn's appearance to estimate its depth and extent. Despite its widespread use, this method has limitations, it's subjective, varies based on the examiner's experience, and lacks a standardized assessment protocol. Consequently, this can lead to divergent treatment recommendations and inconsistency in patient care, sometimes resulting in sub-optimal outcomes or complications [2].

Therefore, the imperative for a more objective, reliable, and scalable solution becomes clear. This project thus proposes a transformative approach: leveraging Artificial Intelligence (AI) and Machine Learning (ML) in particular for burn analysis and detection. Central to our proposal is the adaptation of deep learning techniques, such as YOLO [3] algorithm designed for object detection. By utilising such approaches, our project aims to discern burns in visual media, be it images for example, and provide a reliable severity assessment.

1.2 Project Scope

The primary objective of this project is to develop an AI-based model that specializes in distinguishing between first-degree, second-degree, and third-degree burns [4]. To ensure the model's precision and reliability, we will train the model on a diverse dataset, encompassing diverse burn images from various patient demographics, burn sources, and stages of injury. This dataset will serve as the foundation upon which our ML model will be trained, enabling it to understand the visual characteristics unique to each burn degree.

Once trained, as usual with ML-powered models, the model will be adept at processing and categorizing the severity of burns in new, unseen images, positioning it as a valuable adjunct to healthcare professionals in their assessment and diagnosis of burn injuries.

1.3 Motivation

ML-based systems have demonstrated the ability to improve the accuracy and expediency of burn detection, thereby ensuring the timely provision of essential care to patients [5, 6]. These systems alleviate the dependence on human expertise, particularly in remote or underserved regions, effectively addressing healthcare disparities. With their objectivity and consistency, ML-based systems offer assessments that mitigate the inherent risks of bias and diagnostic errors. Furthermore, these systems hold promise for early detection, facilitating timely interventions and fostering improved patient outcomes.

Our main drive to develop a ML-based model for burn detection comes from a few key reasons. Firstly, we want to offer a faster way to check burns. Secondly, this tool can help doctors give advice over the phone or online, especially in places where getting to a hospital is hard. Lastly, it is also a useful aid for those who might find it hard to see burns clearly or can't easily move around. Overall, by using ML for burn care, we believe both patients and healthcare workers can benefit greatly.

1.4 Aim and Objectives

1.4.1 Aim

Our primary goal is to develop an AI system that utilizes deep learning methodologies to detect and categorize burn injuries. While there are various algorithms available, we are initially exploring the capabilities of the You Only Look Once (YOLO) [7] model as an exemplary object detection model. This exploration will guide us in determining the most optimal approach for our specific needs in burn identification from images, ensuring both efficiency and precision.

1.4.2 Objectives

Data Collection

The objective of data collection is to gather a large and diverse dataset of burn-related images and videos. This involves conducting thorough research and sourcing data from various reliable and relevant sources. The collected data should include examples of different types of burns, such as first-degree, second-degree, and third-degree burns.

Model Development

The objective of model development is to design and develop the architecture of the burn detection model. This includes selecting an appropriate deep-learning framework and defining the network architecture. The model should be capable of accurately detecting and classifying burns in images and videos. The architecture may involve using convolutional neural networks (CNNs) [8].

Model Training and Evaluation

Once the model architecture is defined, the objective is to train the model using the collected dataset. This involves feeding the dataset into the model and optimizing its parameters to improve its performance. The training process aims to enable the model to learn patterns and features that distinguish different types of burns. After training, the model needs to be evaluated using appropriate evaluation metrics, such as accuracy, precision, recall, and score.

User Interface

Ensuring our model is user-friendly is essential. We aim to design an intuitive web-based interface that integrates with our model. This approach aims to make the tool accessible and effective for both medical professionals and the wider public, facilitating rapid and accurate burn detection.

1.5 Project Plan and Schedule

Our plan for the project started with deciding on the appropriate idea for our project. Next, we ensured the idea's feasibility by searching for the appropriate datasets and tools. After that, we began writing an introduction and a literature review for the

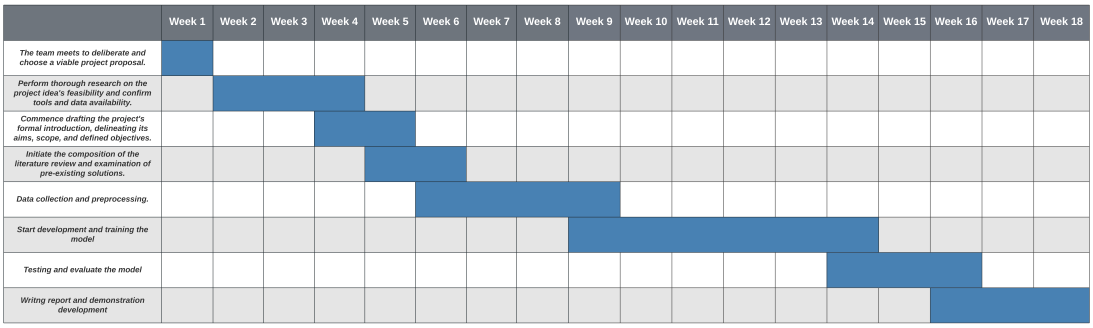

Figure 1.1: Time line

project. Following this, we initiated the work by collecting the datasets and commencing the development and testing of the model. Finally, we will compile a report summarizing our work. The detailed timeline of our project is illustrated in Figure 1.1.

1.6 Outline of The Project

After this introductory chapter, we will write three more chapters in this report as follows:

- **Chapter 2: Literature Review**. It will contain an introduction to this chapter, background information on the project's components, a section on existing related systems, and a concluding section summarizing the chapter.

- **Chapter 3: Problem Analysis**. This chapter will also include an introduction. We will proceed to specify a problem and provide a solution. Following that, we will analyze the system by collecting, analyzing, and organizing the requirements. Afterward, we will implement and evaluate a plan for the system, selecting technical tools and languages that are appropriate. Finally, we will provide a summary of the chapter.

- **Chapter 4: System Design**. Similar to the previous two chapters, it will begin with an introduction. It will then include the System Design Specification, followed by the drawing of the system architecture. Subsequently, it will feature pseudocode, algorithms, or scenarios of the model. Finally, it will conclude with a summary.

Chapter 2

Literature Review

This chapter provides some background research on the project and examines some previous work.

2.1 Introduction

ANAS: This chapter provides an overview of existing research on automated skin burn classification. The review focuses on studies that have explored machine learning and image analysis techniques for accurate classification. Various algorithms, including CNNs and SVMs, have been utilized, along with the integration of clinical data. The aim is to improve the efficiency and effectiveness of burn classification, ultimately enhancing patient care and outcomes. The review identifies gaps in current research and highlights areas for further investigation in this field.

2.2 Background

2.2.1 Skin Burns

AZIZ: Burns can be categorized into three degrees based on their severity. First-degree burns are superficial and only affect the outermost layer of the skin, causing redness, pain, and mild swelling. These burns heal within a week without leaving scars. Second-degree burns extend beyond the epidermis and affect the underlying dermis. They are characterized by blistering, intense pain, redness, and swelling. Healing for second-degree burns may take several weeks and can result in scarring. Third-degree burns are the most severe, penetrating through all skin layers and affecting underlying tissues. They appear charred or white and may cause numbness due to nerve damage. Immediate medical attention is required for third-degree burns, often involving skin grafting for proper healing. Accurate assessment of burn severity is crucial for determining appropriate treatment. AI-based solutions can assist in this assessment by providing precise and consistent evaluations based on factors such as burn depth and extent.

2.2.2 Artificial Intelligence and Machine Learning in Skin Burns

AZIZ: approaches to burn injury evaluation rely on healthcare professionals visually examining the affected skin to estimate the severity and extent of burns based on their visual characteristics. These methods do have some shortcomings, though, like a reliance on personal knowledge and inconsistent treatment suggestions. The use of machine learning (ML) and artificial intelligence (AI) techniques for analyzing and diagnosing skin burns is on the rise as a solution to these problems. These solutions' ability to provide accurate and consistent assessments of burn severity using AI and ML will improve diagnostic procedures, which will ultimately result in better patient outcomes.

2.2.3 Image classification

AZIZ: Image classification is a complex task, crucial in domains like stock market prediction, weather forecasting, and medical diagnosis. It involves multiple factors and techniques to enhance accuracy. Challenges arise from data properties and model capabilities. Classification can be done through supervised and unsupervised methods, with the widely used Support Vector Machine (SVM) as a popular learning model.[9]

2.3 Existing Related Systems

2.3.1 generative adversarial network

ANAS: In this study Abdelhalim et al. [10] , focuses on data augmentation techniques to enhance skin disease detection using images. It proposes a method called SPGGAN for data augmentation in dermoscopy images. The SPGGAN generates synthetic skin lesion images to augment the training dataset, improving the performance of automated skin lesion detection systems. Experimental results show the effectiveness of the proposed method compared to traditional data augmentation approaches.

2.3.2 image classification in the medical field

MK: In recent years the advancements in image classification technology was used for the goal of fast and accurate diagnosis of medical conditions ,aiming to help radiologists, physicians, or researchers in accurate disease detection from X-rays, MRIs, CT scans,and images to ultimately improve patient care and health.

MK: Mateen et al. [11] , the paper focuses on automating diabetic retinopathy (DR) diagnosis from fundus images. their approach combines a Gaussian mixture model (GMM) for precise region segmentation, the VGGNet [12] for high-dimensional feature extraction, and a combination of Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) for efficient feature selection. Using a Kaggle dataset of 35,126 fundus images, their study demonstrates a better performance of the VGG-19 Deep Neural Network (DNN) in comparison to other models like AlexNet [13] and SIFT [14], while also showcasing reduced computational time.

MK: The authors in this study [15] presented an automated approach for detecting COVID-19 cases in X-ray images with deep neural networks, they introduced the DarkCovidNet model which is inspired by the Darknet-19 [16] architecture. They used publicly available datasets, including COVID-19 X-ray images and the ChestX-ray8 database, their model achieved notable accuracy, sensitivity, specificity, and F1-score values, demonstrating its potential as a valuable tool that can assist radiologists and healthcare professionals in efficient COVID-19 diagnosis.

ANAS: In this study Aishwarya et al. [17] , focuses on developing an automated system for diagnosing skin cancer using the Yolo deep neural network. The study utilizes a dataset of skin cancer lesion images and trains the Yolo model to classify and detect different types of skin cancer. The performance of YoloV3 and YoloV4 models is evaluated, showing high classification scores and outperforming existing methods.

AZIZ: ALKolifi ALEnezi [18] build system employs image processing and machine learning to detect, extract, and classify skin diseases. It follows a process involving preprocessing to enhance image quality, feature extraction through resizing and CNN-based feature extraction, and classification using a multiclass SVM to identify various skin diseases. The system effectively identifies Eczema, Melanoma, and Psoriasis with 100% accuracy using a dataset of 100 images (80 for training and 20 for validation).

AH: In this research of Rostami et al. [19] introduces an ensemble Deep Convolutional Neural Network (DCNN) classifier, merging patch-based classification with AlexNet, for image-wise classification of diverse wound types, addressing the challenges of manual wound classification with advanced Artificial Intelligence techniques. The study employs an ensemble (DCNN) classifier, combining patch-wise and image-wise classifications through Multilayer Perceptron (MLP), utilizing various deep architectures with transfer learning, and object localization algorithms, achieving superior multi-class wound image classification performance compared to methods discussed in the literature. The proposed ensemble DCNN-based classifier demonstrated high accuracy, achieving maximum values of 96.4% for binary and 91.9% for 3-class wound image classification, outperforming other deep classifiers.

2.3.3 image classification in skin burns

ANAS: In this study E.B. et al. [20] , conducts a comparative study of different segmentation algorithms in the classification of human skin burn depth. Eight hybrid segmentation algorithms are studied on a dataset of burn images categorized into three burn classes. The performance of the algorithms is evaluated by calculating the number of correctly segmented images for each burn depth. The study highlights the importance of using a good segmentation algorithm for accurate classification of skin burn depth.

AZIZ: Rangel-Olvera and Rosas-Romero [4] used sparse representation with over-redundant dictionaries to detect and classify burnt areas in color images,two methods were used to build dictionaries for burn severity classes: direct collection of feature vectors from patches in various images and locations, and collection of feature vectors followed by dictionary learning accompanied by K-singular value decomposition.The image dataset included 33 first-degree burns, 26 second-degree burns, and 56 third-degree burns, with a diversity of background artifacts,they achieved 95.65 percent sensitivity and 94.02 percent precision. Color and texture were important features, and a shadowed skin dictionary reduced false positives.

AH: The study of Suha and Sanam [21] proposes a Deep Convolutional Neural Network (DCNN) approach for real-time detection and categorization of burn severity, outperforming traditional methods by incorporating transfer learning, fine-tuning, and multiple convolutional layers for feature extraction and classification. The study compares (DCNN) by its effectiveness with traditional methods incorporating digital image processing and machine learning classifiers. And reveals that (DCNN) model's superior accuracy of 95.63% in detecting burn severity, outperforming traditional classifiers. Augmented data improved performance across models, underscoring the significance of dataset size, demonstrating the potential of intelligent technologies in burn assessment and treatment.

AH: The paper of Abubakar et al. [22] addresses the challenges of accurate burn assessment by proposing machine learning techniques, specifically utilizing transfer learning, to discriminate burns from injured skin, aiming to overcome the limitations of human expertise and high assessment costs. The paper employs transfer learning with pre-trained deep learning models, ResNet50, ResNet101, and ResNet152, utilizing both fine-tuning and off-the-shelf features approaches to classify skin burn images and other injuries, enhancing discrimination accuracy. The study's approach, employing transfer learning with fine-tuned ResNet50, ResNet101, and ResNet152 models using Google

Collab's GPU hardware, achieved exceptional accuracy of approximately 99.9% in discriminating burnt skin and injured skin, utilizing 200 training epochs and ensuring diverse data representation patterns.

2.3.4 Man vs. Machine: comparison Techniques in Machine Learning Algorithms

AZIZ: Haenssle et al. [23] a modified Google Inception v4 CNN was trained on dermoscopic images and corresponding diagnoses. Dermatologists evaluated images at two levels: Level-I and Level-II. In Level-I, dermatologists correctly identified 86.6% of melanomas and 71.3% of benign lesions. In Level-II, their accuracy improved to 88.9% for melanomas and 75.7percent for benign lesions. The CNN outperformed dermatologists, with higher specificity (82.5%) in both levels and a greater area under the ROC curve (0.86 vs. 0.79).

AH: This paper of Rao et al. [24] presents the Global Filter Network GFNet, a computational efficient alternative to self-attention and MLP models, utilizing a 2D discrete Fourier transform by replacing self-attention layers in transformer architectures with a 2D discrete Fourier transform, element-wise multiplication with global filters, and inverse Fourier transform, evaluated on ImageNet and various datasets, demonstrating superior accuracy/complexity trade-offs, robustness, and generalization ability. The paper showcases GFNet's competitive edge in image classification by outperforming various architectures on ImageNet and demonstrating superior adversarial robustness and generalization abilities across multiple datasets, positioning it as a compelling alternative to transformer-style models and CNNs.

AH: Chaganti et al. [25] explores image classification, beginning with traditional machine learning like Support Vector Machines (SVM) despite the prevalence of Neural Networks (NN). The study initially utilized Support Vector Machines (SVM) for image classification, achieving 93% accuracy on a small dataset which reduced to 82% with data augmentation. Transitioning to Convolutional Neural Networks (CNN) resulted in an impressive 93.57% accuracy on the same dataset, emphasizing the superior potential of deep learning methods, facilitated by various data augmentation techniques.

MK: In [26] the authors introduce CheXNet which is a robust 121-layer convolutional neural network (CNN) based on the DenseNet[27] architecture, with the aim of identifying pneumonia from chest X-rays and surpassing the diagnostic capabilities of practicing radiologists. They used the ChestX-ray14 dataset which includes over 100,000 annotated X-ray images, CheXNet also pinpoints pertinent areas most indicative of pneumonia within the image. When it was evaluated against the annotations of four radiologists that had 4, 7, 25, and 28 years of experience, CheXNet achieved an impressive F1 score of 0.435, and it outperformed the average of the radiologist's performance.

2.4 Contribution

2.5 Summary

MK: * In this chapter we have given a background on the limitations of conventional burn assessment techniques and the potential benefits of AI and ML-based solutions, we also briefly explain the various degrees of burn severity and underline the importance of accurate assessment for an effective treatment , then we give a few examples on some existing work done in image classification in the medical field , in skin burn injuries and a comparison between machine learning techniques and some dermatologists experts ....to be continued.

Chapter 3

Problem Analysis

This chapter is about the problem analysis.

Chapter 4

System Design

This chapter examines the design of the project.

Bibliography

[1] John P Abraham, Brian D Plourde, Lauren J Vallez, Brittany B Nelson-Cheeseman, John R Stark, Ephraim M Sparrow, and John M Gorman. Skin burns. *Theory and Applications of Heat Transfer in Humans*, 2:723–739, 2018.

[2] PN Kuan, S Chua, EB Safawi, HH Wang, and W Tiong. A comparative study of the classification of skin burn depth in human. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 9(2-10):15–23, 2017.

[3] Peiyuan Jiang, Daji Ergu, Fangyao Liu, Ying Cai, and Bo Ma. A review of yolo algorithm developments. *Procedia Computer Science*, 199:1066–1073, 2022.

[4] Brenda Rangel-Olvera and Roberto Rosas-Romero. Detection and classification of burnt skin via sparse representation of signals by over-redundant dictionaries. *Computers in Biology and Medicine*, 132:104310, 2021. ISSN 0010-4825. doi: https://doi.org/10.1016/j.compbiomed.2021.104310. URL https://www.sciencedirect.com/science/article/pii/S0010482521001049.

[5] Nehemiah T. Liu and Jose Salinas. Machine learning in burn care and research: A systematic review of the literature. *Burns*, 41(8):1636–1641, 2015. ISSN 0305-4179. doi: https://doi.org/10.1016/j.burns.2015.07.001. URL https://www.sciencedirect.com/science/article/pii/S0305417915002004.

[6] Dhirendra Yadav, Ashish Sharma, Madhusudan Singh, and Ayush Goyal. Feature extraction based machine learning for human burn diagnosis from burn images. *IEEE Journal of Translational Engineering in Health and Medicine*, PP:1–1, 07 2019. doi: 10.1109/JTEHM.2019.2923628.

[7] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[8] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[9] D. Lu and Q. Weng. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5):823–870, 2007. doi: 10.1080/01431160600746456. URL https://doi.org/10.1080/01431160600746456.

[10] Ibrahim Saad Aly Abdelhalim, Mamdouh Farouk Mohamed, and Yousef Bassyouni Mahdy. Data augmentation for skin lesion using self-attention based progressive generative adversarial network. *Expert Systems With Applications*, 165, 2021. ISSN 0957-4174.

[11] Muhammad Mateen, Junhao Wen, Nasrullah, Sun Song, and Zhouping Huang. Fundus image classification using vgg-19 architecture with pca and svd. *Symmetry*, 11(1):1, 2018.

[12] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.

[14] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60:91–110, 2004.

[15] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U Rajendra Acharya. Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in biology and medicine*, 121:103792, 2020.

[16] Joseph Redmon and Ali Farhadi. Yolo9000: better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7263–7271, 2017.

[17] N Aishwarya, K Manoj Prabhakaran, Frezewd Tsegaye Debebe, M Sai Sree Akshitha Reddy, and Posina Pranavee. Skin cancer diagnosis with yolo deep neural network. *Procedia Computer Science*, 220:651 – 658, 2023. ISSN 18770509.

[18] Nawal Soliman ALKolifi ALEnezi. A method of skin disease detection using image processing and machine learning. *Procedia Computer Science*, 163:85–92, 2019. ISSN 1877-0509. doi: https://doi.org/10.1016/j.procs.2019.12.090. URL https://www.sciencedirect.com/science/article/pii/S1877050919321295. 16th Learning and Technology Conference 2019Artificial Intelligence and Machine Learning: Embedding the Intelligence.

[19] Behrouz Rostami, DM Anisuzzaman, Chuanbo Wang, Sandeep Gopalakrishnan, Jeffrey Niezgoda, and Zeyun Yu. Multiclass wound image classification using an ensemble deep cnn-based classifier. *Computers in Biology and Medicine*, 134:104536, 2021.

[20] Safawi E.B., Wang H.H., Chua Stephanie, and Kuan P.N. A comparative study of segmentation algorithms in the classification of human skin burn depth. *International Journal on Advanced Science, Engineering and Information Technology*, 10:145, 2020. ISSN 2460-6952.

[21] Sayma Alam Suha and Tahsina Farah Sanam. A deep convolutional neural network-based approach for detecting burn severity from skin burn images. *Machine Learning with Applications*, 9:100371, 2022.

[22] Aliyu Abubakar, Mohammed Ajuji, and Ibrahim Usman Yahya. Comparison of deep transfer learning techniques in human skin burns discrimination. *Applied System Innovation*, 3(2):20, 2020.

[23] Holger Haenssle, Christine Fink, R Schneiderbauer, Ferdinand Toberer, Timo Buhl, A Blum, A Kalloo, A Hassen, Luc Thomas, A Enk, and Lorenz Uhlmann. Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of oncology : official journal of the European Society for Medical Oncology*, 29, 05 2018. doi: 10.1093/annonc/mdy166.

[24] Yongming Rao, Wenliang Zhao, Zheng Zhu, Jiwen Lu, and Jie Zhou. Global filter networks for image classification. *Advances in neural information processing systems*, 34:980–993, 2021.

[25] Sai Yeshwanth Chaganti, Ipseeta Nanda, Koteswara Rao Pandi, Tavva GNRSN Prudhvith, and Niraj Kumar. Image classification using svm and cnn. In *2020 International conference on computer science, engineering and applications (ICCSEA)*, pages 1–5. IEEE, 2020.

[26] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017.

[27] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.