KULLIYYAH OF INFORMATION AND COMMUNICATION TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE

FINAL YEAR PROGRESS REPORT

PROJECT ID

1215R

PROJECT TITLE

DEEP LEARNING-BASED EVALUATION OF THE RELATIONSHIP BETWEEN MANDIBULAR THIRD MOLAR AND MANDIBULAR CANAL ON CBCT

STUDENT(S)

1. MUHAMMAD SYAFIQ FARHAN BIN MOHD FARIDZ (2011871)
2. MUHAMMAD ASYRAF BIN ANAHAR (2027199)

SUPERVISOR

ASSISTANT PROF. DR. NORLIA MD YUSOF

JANUARY 2024

SEMESTER 1 2023/2024

FINAL YEAR PROGRESS REPORT

PROJECT ID

1215R

PROJECT TITLE

DEEP LEARNING-BASED EVALUATION OF THE RELATIONSHIP BETWEEN MANDIBULAR THIRD MOLAR AND MANDIBULAR CANAL ON CBCT

PROJECT CATEGORY

RESEARCH

by

MUHAMMAD SYAFIQ FARHAN BIN MOHD FARIDZ (2011871)

MUHAMMAD ASYRAF BIN ANAHAR (2027199)

SUPERVISED BY

ASSISTANT PROF. DR. NORLIA MD YUSOF

In partial fulfillment of the requirement for the

Bachelor of Computer Science

Kuliyyah of Information and Communication Technology

International Islamic University Malaysia

ABSTRACT

This research focuses on the innovative use of deep learning, particularly convolutional Neural Networks (CNNs) such as VGG-16 and ResNet-50, for the examination of Cone-Beam computed Tomography (CBCT) images in dental diagnostics. The project seeks to improve the identification of dental problems, especially in the Mandibular Third Molar and Mandibular canal, by utilizing Generative Adversarial Networks (GANs) to address the challenge of limited data availability. CBCT imaging offers essential three-dimensional dental data, but its precise analysis might be difficult due to complex anatomical structures. GANs are crucial in enhancing the training dataset by producing synthetic CBCT images, hence enhancing the accuracy of detection. In addition, this research investigates the complex interconnections among different dental features within CBCT images. Gaining a comprehensive understanding of these relationships is crucial for developing diagnostic models that are more robust and effective. The acquired insights possess the capacity to completely transform dental diagnostics, enhancing accuracy and effectiveness. This project aims to improve patient care by automating the interpretation of CBCT images, hence lowering the workload of healthcare personnel and minimizing the need for manual analysis. This study makes a valuable contribution to the progress of healthcare technology and the empowerment of dental professionals. To summarise, this research is groundbreaking in its application of deep learning, GANs, and sophisticated neural networks in the processing of CBCT images for dental diagnosis. The relevance of this technology lies in its ability to enhance patient outcomes, simplify treatment planning, and minimize the need for manual intervention, thereby paving the way for the future of data-driven dental healthcare.

TABLE OF CONTENTS

CHAPTER ONE
INTRODUCTION

1.1 Project Overview

Our research targets the prevalent issue of class imbalance in dental diagnostic datasets, particularly evident in CBCT images. The disparate distribution of data poses significant challenges for training effective deep learning models. To address this, we employ a strategic integration of GANs with CNN models, aiming to boost accuracy. GANs play a pivotal role by generating synthetic data that closely mimics real dental conditions, effectively addressing imbalances and furnishing a balanced training dataset. This augmentation enhances the CNN models' ability to identify dental abnormalities, especially those with limited examples in the original dataset. While our research is an ongoing process with continual enhancements, we are dedicated to exploring other data augmentation techniques to further refine our approach and extend the capabilities of dental diagnostics.

1.2 Problem Statement

Our research centers on the critical issue of class imbalance in dental diagnostics, posing challenges for the optimal functioning of CNN models. The uneven distribution of classes in our dataset hampers the accurate detection of dental problems in CBCT images. The prevalence of certain issues and the scarcity of others lead to biased models and diminished accuracy. This issue has far-reaching consequences for dental healthcare, potentially resulting in inaccurate diagnoses and delayed interventions. Our proposed solution involves integrating GANs with CNN models to address class imbalance by introducing artificially generated data. This rebalancing process aims to enhance the capabilities of CNN models, ensuring accurate identification of dental issues in CBCT images, and contributing to the overarching goal of advancing dental diagnostics in clinical settings.

1.3 Project Objectives

The initiative aims to achieve a variety of objectives that are diverse and complex:

1. Imbalance Mitigation: Fix the dataset's class imbalance to improve model training.
2. Generative Adversarial Network (GAN) Integration: Using GAN to generate synthetic images to correct the imbalance.
3. Model Enhancement: Integrate GAN-augmented data into existing CNN models.
4. Performance Evaluation: Extract insights into the effectiveness of the GAN-augmented CNN models

1.4 Significance of Project

The effective implementation of this study has great importance in the field of automated dental diagnostics. The potential impact of this technology is vast, as it promises to bring about a new era of improved healthcare outcomes, more effective treatment planning, and a significant decrease in the need for manual interpretation of CBCT images.

1. **Enhanced Diagnosis Accuracy:** The project's results have the potential to greatly enhance the precision of dental issue identification. To achieve a more accurate and consistent diagnosis of dental diseases, it is necessary to tackle the problem of class imbalance in the dataset and enhance the precision of CNN models. Consequently, this leads to faster and more accurate diagnoses for patients, minimizing the possibility of errors and guaranteeing prompt medical treatment.
2. **Efficient Treatment Strategizing:** Within the field of dentistry, the precision of diagnosis is essential in choosing the most appropriate treatment methods. Enhancing the capacity of our models to identify and categorize dental problems in CBCT images empowers dental professionals with important knowledge. This allows them to develop treatment strategies with a greater level of certainty, resulting in enhanced and personalized patient care.
3. **Reduced Manual Interpretation:** An important advantage of our research is its capacity to decrease the heavy dependence on manual interpretation of CBCT images. The current procedure of manually examining and evaluating these photographs is

laborious and prone to human errors. Our initiative simplifies and optimizes dental diagnostics, resulting in a streamlined workflow for healthcare providers and reducing the risk of misinterpretation. This eventually leads to improved patient care.

4. **Advancing Dental Healthcare:** In addition to its immediate uses, the effective implementation of this study adds to the overall progress of dental healthcare. It is in line with the increasing tendency towards data-driven decision-making in healthcare, utilizing state-of-the-art technologies to enhance patient outcomes. As the research progresses and its methods are improved, it establishes a standard for future advancements in automated dental diagnostics.

1.5 Project Schedule

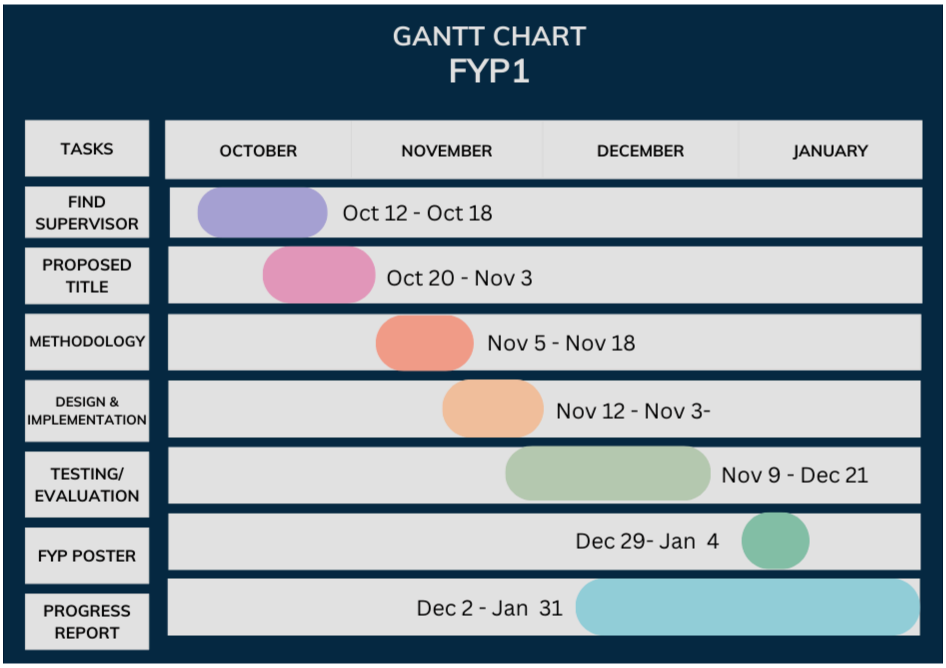

Figure 1: Gantt Chart of CNN Model Development.

CHAPTER TWO

REVIEW OF PREVIOUS WORK

2.1 Introduction

An assessment of the correlation between mandibular third molars and the mandibular canal is crucial for dental diagnosis and surgical planning. The accuracy and efficiency of these evaluations have been greatly improved by recent developments in deep learning. This section examines pivotal research that has made significant contributions to this particular project.

2.2 Key Studies of Deep Learning-Based Evaluation of Mandibular Third Molar and Mandibular Canal Relationship

2.2.1 Article 1

Study by Buyuk et al. (2022). Titled "A Fused Deep Learning Architecture for the Detection of the Relationship between the Mandibular Third Molar and the Mandibular Canal" : This work involved the development of a sophisticated deep learning system that was able to accurately classify the connection between the mandibular third molar and mandibular canal using orthopantomography. The technique utilized a hybrid architecture that included a U-Net-like structure for image segmentation with AlexNet for image classification. This model demonstrated exceptional accuracy and was highly efficient in aiding dental practitioners in identifying anatomical variances, a critical aspect in preventing surgical problems (Buyuk et al., 2022).

2.2.2 Article 2

Study by Zhu et al. (2021). Titled "Artificial Intelligence Model to Detect Real Contact Relationship between Mandibular Third Molars and Inferior Alveolar Nerve Based on Panoramic Radiographs" : The objective of this study was to create a deep-learning model called MM3-IANnet, which can accurately evaluate the actual contact between mandibular third molars and the inferior alveolar nerve using panoramic radiographs. The objective of the model was to minimize the occurrence of pseudo-contact interference and decrease the reliance on CBCT use. The MM3-IANnet demonstrated exceptional accuracy, making it a very promising instrument for dentists to identify actual contact connections and perhaps minimize the likelihood of nerve injury during dental operations (Zhu et al., 2021).

2.2.3 Article 3

Study by Maruta et al. (2023). Titled "Automatic Machine Learning-based Classification of Mandibular Third Molar Impaction Status" : The study aimed to evaluate the effectiveness of automated machine learning (ML) in classifying the impaction status of mandibular third molars. Using a dataset of 1864 images, researchers annotated impaction patterns using Pell and Gregory as well as Winter classifications. Data augmentation techniques, including rotation and flip, were applied, and VGG-16 was employed for ML. The augmented dataset-trained model showed strong performance in Pell and Gregory Class classification, achieving an accuracy of 0.8609, macro F1-score of 0.7624, and macro AUC of 0.9334. Similar results were obtained for position categorization. The ML model demonstrated notable efficacy in determining impaction status, potentially benefiting dentists in image evaluation and improving patient care and informed consent in oral surgery (Maruta et al., 2023).

2.2.4 Article 4

Study by Choi et al. (2022). Titled: "Artificial Intelligence in Positioning between Mandibular Third Molar and Inferior Alveolar Nerve on Panoramic Radiography" : The paper presents an AI model designed to determine the positional relationship between the mandibular third molar (M3) and the inferior alveolar nerve (IAN) in panoramic radiographs, crucial for surgical extractions. The AI model outperformed oral and maxillofacial surgery (OMFS) professionals in accurately detecting contact and bucco-lingual location between M3 and IAN, showing higher accuracy and significant agreement. The study suggests that AI can aid clinicians in M3 treatment decisions, potentially reducing the need for more radiation-intensive techniques like cone-beam computerized tomography (CBCT). However, it emphasizes the need for additional data and model enhancements (Choi et al., 2022).

CHAPTER THREE

METHODOLOGY

The main objective of this research was to improve the identification and classification of dental problems in CBCT images by employing sophisticated deep-learning methods. We employed Convolutional Neural Networks (CNNs), namely VGG-16 and ResNet-50 architectures, together with the novel application of Generative Adversarial Networks (GANs) to enhance our data augmentation techniques. This section delineates the execution and incorporation of various approaches.

We employed the VGG-16 architecture, which is well-known for its deep layers and compact, 3x3 convolutional filters. This design is particularly effective at extracting complex features from images. Our study involved utilizing CBCT pictures as input for the VGG-16 model. Each layer of the model successively processed the images, extracting and enhancing important features necessary for identifying and classifying different dental problems. The utilization of VGG-16 enabled us to extract intricate features from the CBCT images, which were crucial for precise dental diagnosis.

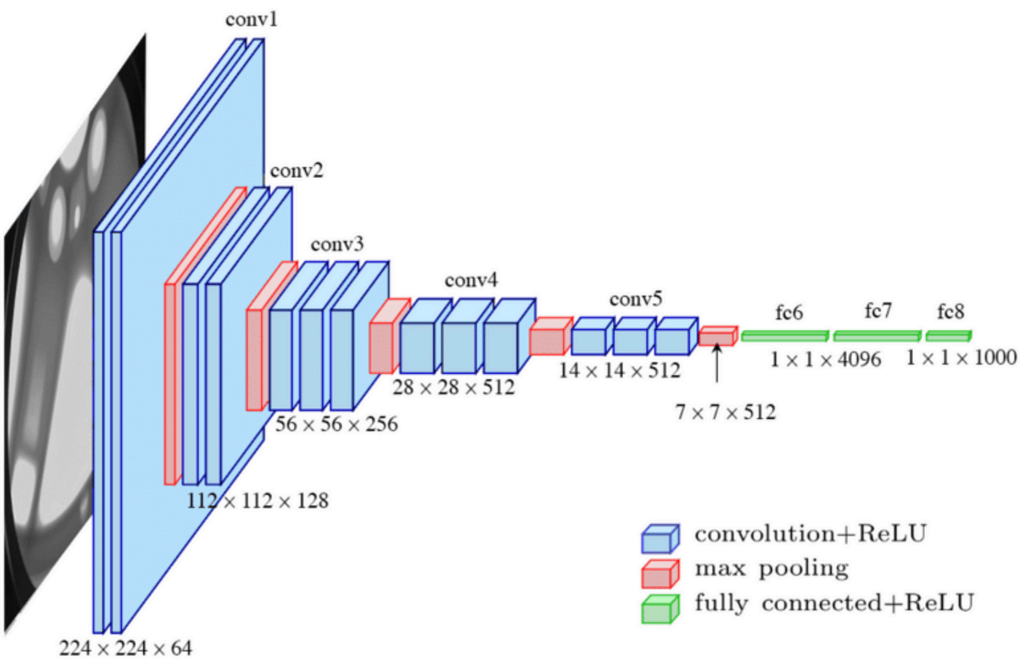

Figure 2: VGG-16 Architecture

We also utilised the ResNet-50 model in combination with VGG16. ResNet-50 stands out due to its capacity to efficiently train neural networks with greater depth, owing to its unique skip connections that tackle the prevalent issue of disappearing gradients in deep neural networks. This architectural design allowed us to utilize a more profound and intricate understanding of the CBCT images. The inclusion of skip connections in ResNet-50 facilitated the acquisition of residual functions, hence assuring efficient training of the network and enabling it to scrutinize more intricate characteristics present in dental images.

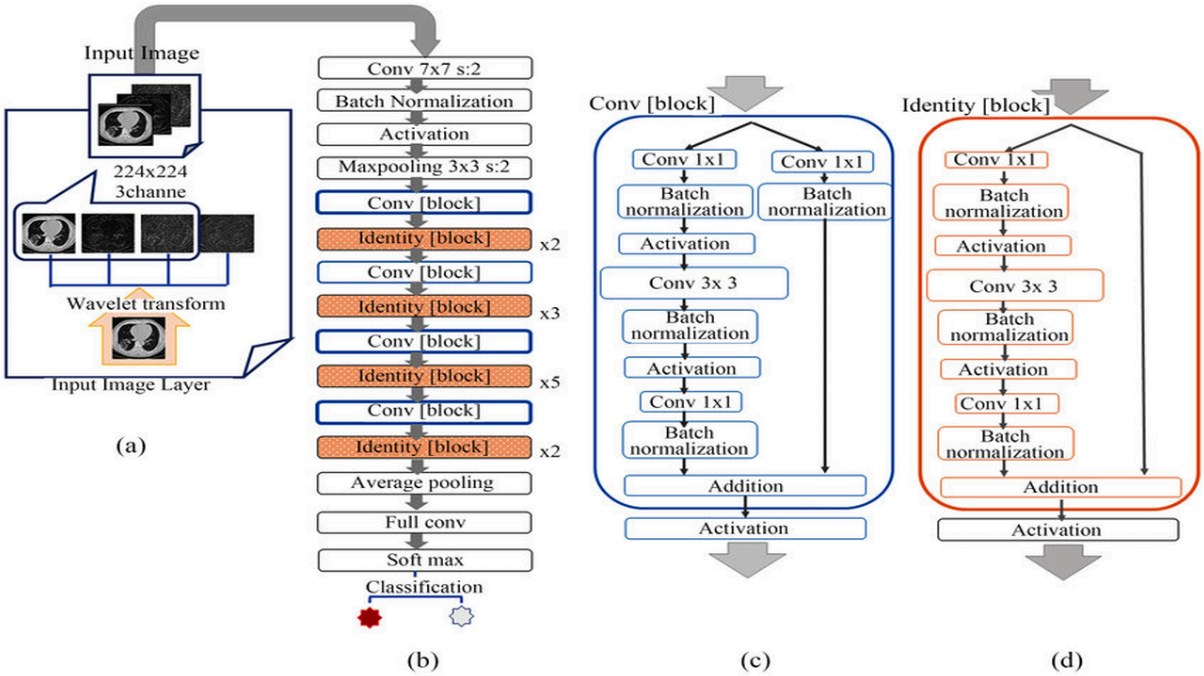

Figure 3: ResNet-50 Architecture

To tackle the issue of restricted data availability, GANs were utilized. GANs were essential in producing artificial yet realistic CBCT images, which were employed to enhance the training dataset. The GANs comprised a generator, responsible for generating new CBCT images, and a discriminator, tasked with evaluating the quality of these images. The iterative process between the generator and discriminator facilitated the creation of synthetic pictures of high quality. These images were subsequently employed to improve the training dataset for both the VGG-16 and ResNet-50 models.

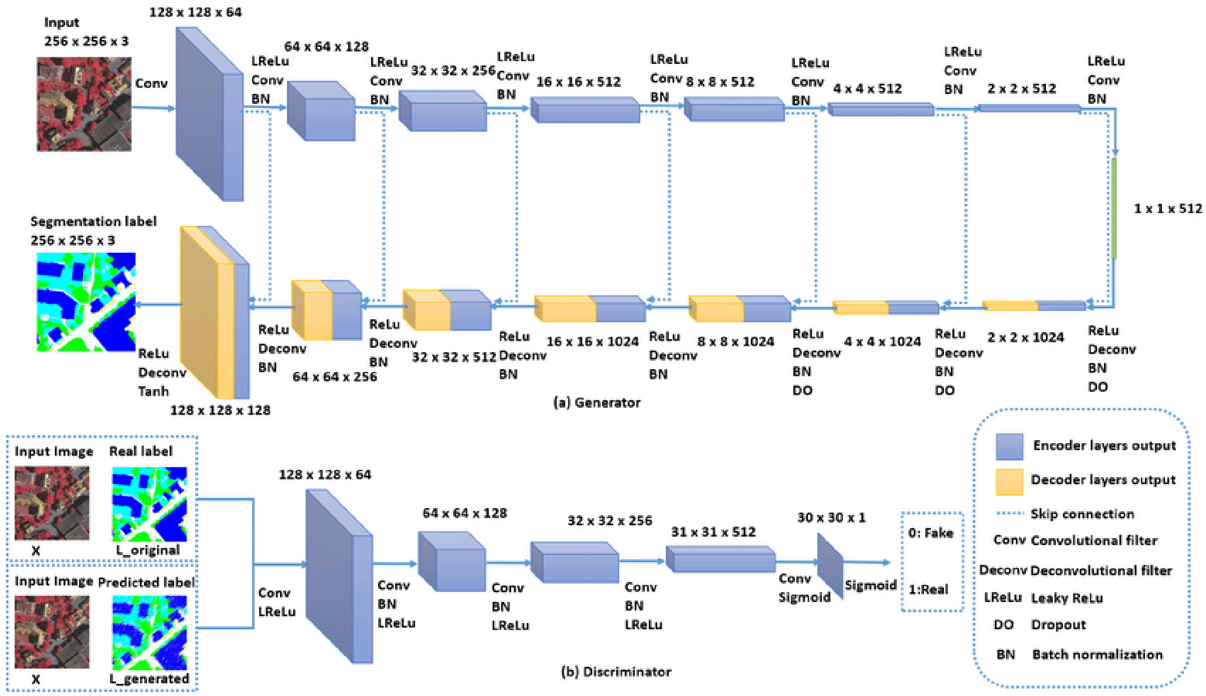

Figure 4: Generative Adversarial Network (GAN) Architecture

The incorporation of three methodologies—VGG-16, ResNet-50, and GANs—provided a holistic approach to our study of CBCT images. GANs facilitated the availability of a wide range of real and artificial images for training, enabling the VGG-16 and ResNet-50 models to gain knowledge from a diverse array of dental problems. By employing a multifaceted strategy, we were able to conduct a comprehensive and robust analysis, resulting in enhanced precision and dependability in identifying dental problems in CBCT images.

3.1 Design

3.1.1 Flowchart

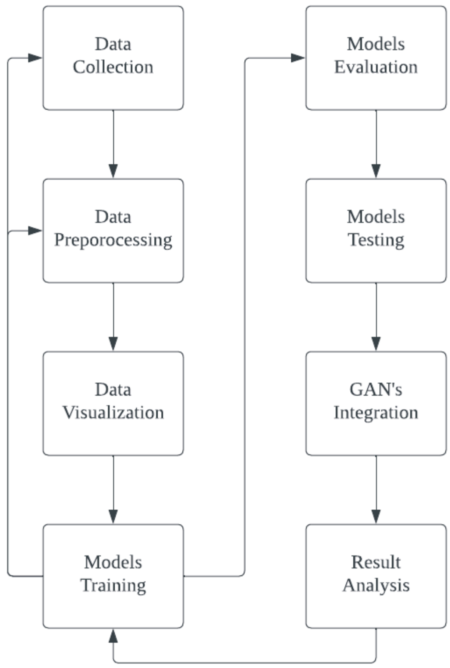

Figure 5: Flowchart of Resnet-50 and VGG-16 model development.

3.1.2 System Architecture

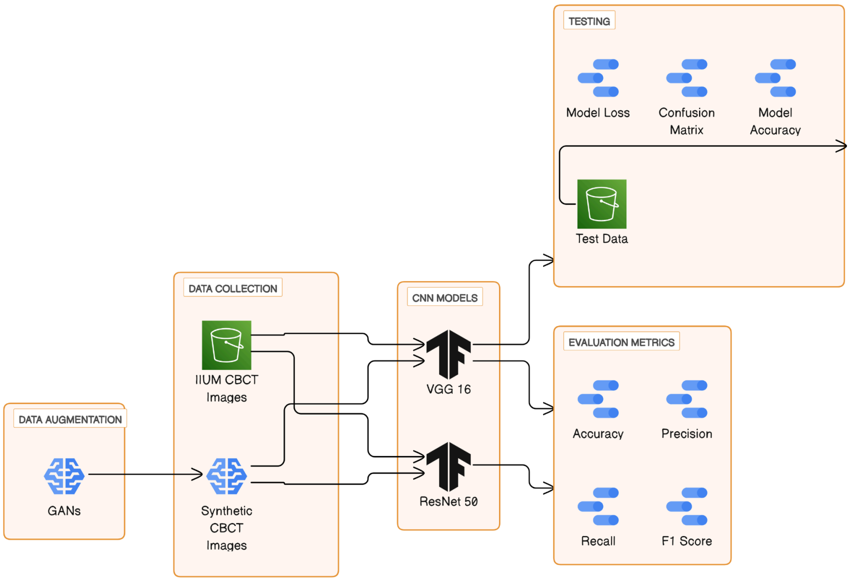

Figure 6: System Architecture of Resnet-50 and VGG-16 model development.

3.2 Data Collection

3.2.1 Source of Datasets

The datasets for this research were acquired from Dr. Badrul at IIUM, Kuantan, which facilitated the procurement of diverse Cone-Beam Computed Tomography (CBCT) images related to dental diagnostics.

3.2.2 Types of CNN Images

The dataset comprises five distinct types of CBCT images, each representing different perspectives: distoangular, horizontal, mesioangular, transverse, and vertical. These variations in image types ensure a comprehensive representation of dental conditions for training and evaluation.

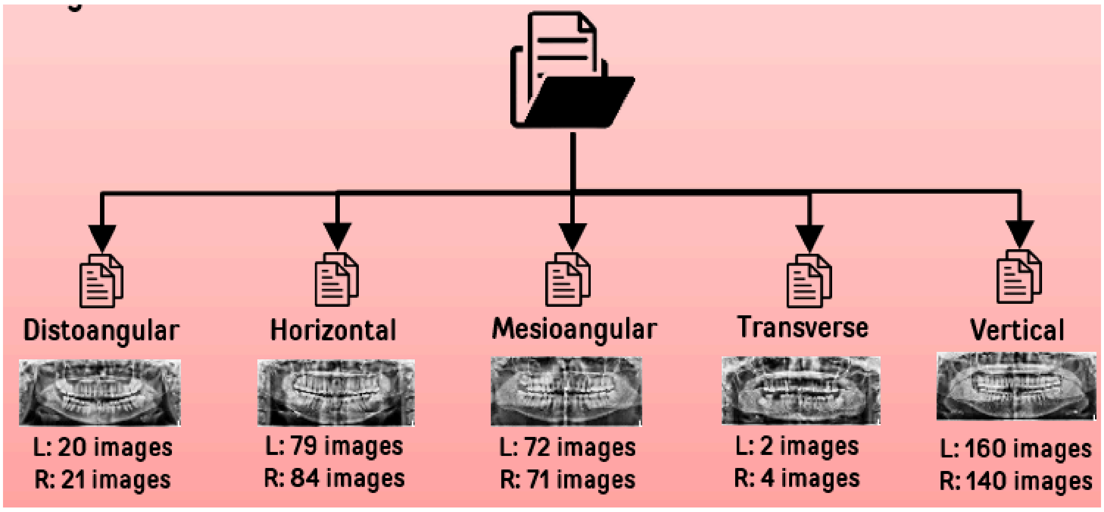

Figure 7: Initial Number of Images for Each Folder/Class

3.3 Testing/Evaluation

3.3.1 Testing Goals

The overarching goals of the testing phase include:

- Evaluate the functionality of the automated diagnostics system in terms of image classification and dental condition identification.
- Measure the speed and resource utilization during system inference.
- Assess the impact of GANs on data augmentation and system generalization.

3.3.2 GANs Testing

This testing type will assess the impact of GANs on the system's ability to generalize to new, synthetic data. The effectiveness of data augmentation techniques facilitated by GANs will be evaluated, and the system's performance when presented with augmented images will be measured.

3.3.3 Test Data

A separate set of test images that were not used during the training or validation phases will be utilized. The test dataset will be designed to ensure diversity in dental conditions and image perspectives.

3.3.4 Testing Metrics

Standard evaluation metrics, including accuracy, precision, recall, and F1 score, will be used to measure the performance of the automated dental diagnostics system.

3.3.5 Evaluation Metrics of ResNet-50 Before Implementation of GAN

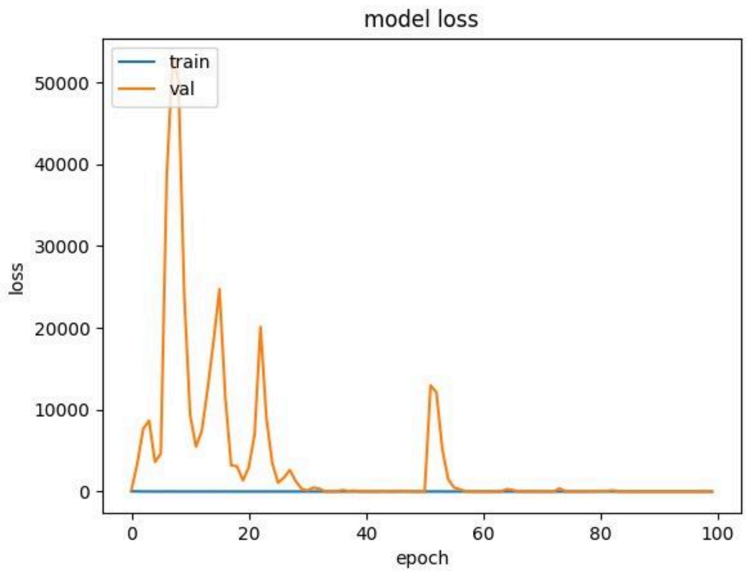

Figure 8: Model loss of ResNet-50

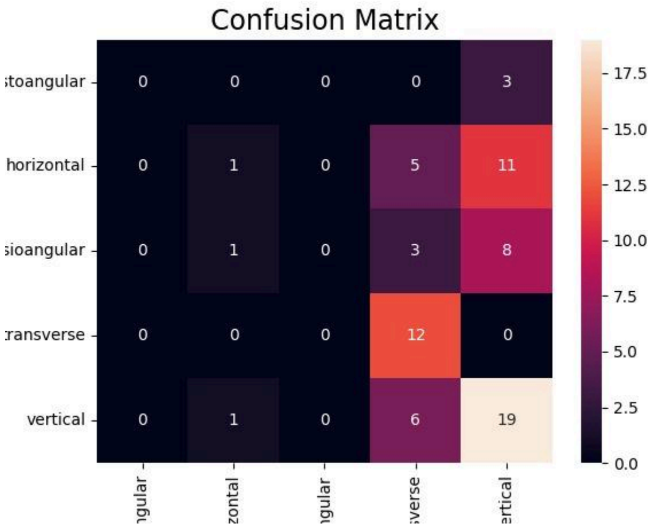

Figure 9: Confusion Matrix of ResNet-50

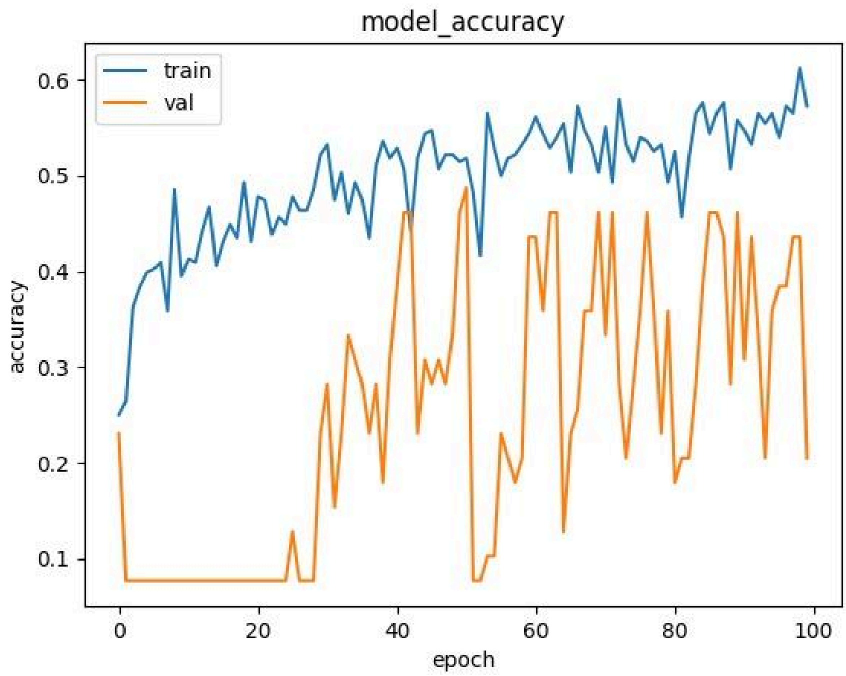

Figure 10: Model accuracy of ResNet-50

The training accuracy (blue line) in Figure 9 increases steadily over time and stabilizes around 0.9, indicating that the model is learning well from the training data. However, the validation accuracy (orange line) fluctuates greatly and remains low (around 0.4), suggesting that the model may not be generalizing well to unseen data.

3.3.6 Evaluation Metrics of VGG-16 Before Implementation of GAN

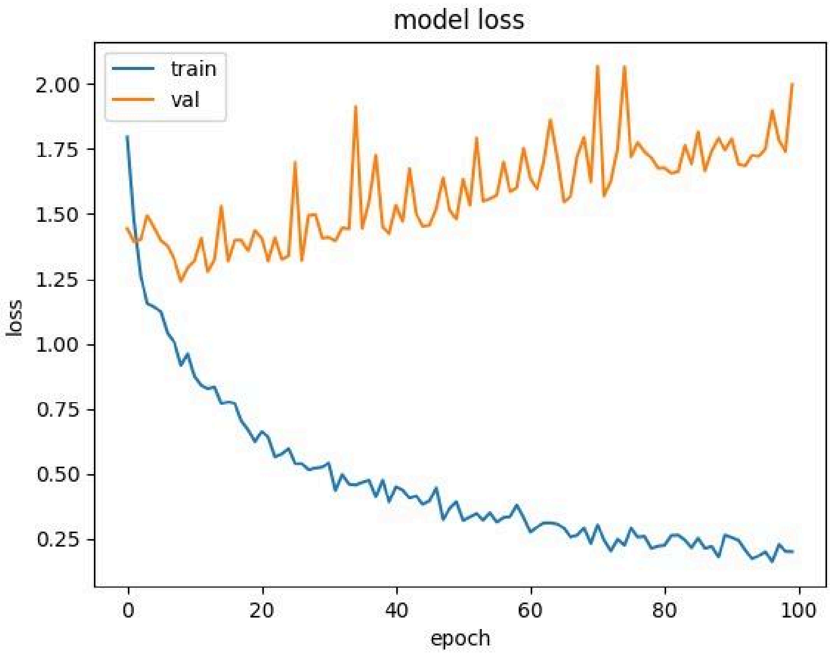

Figure 11: Model loss of VGG-16

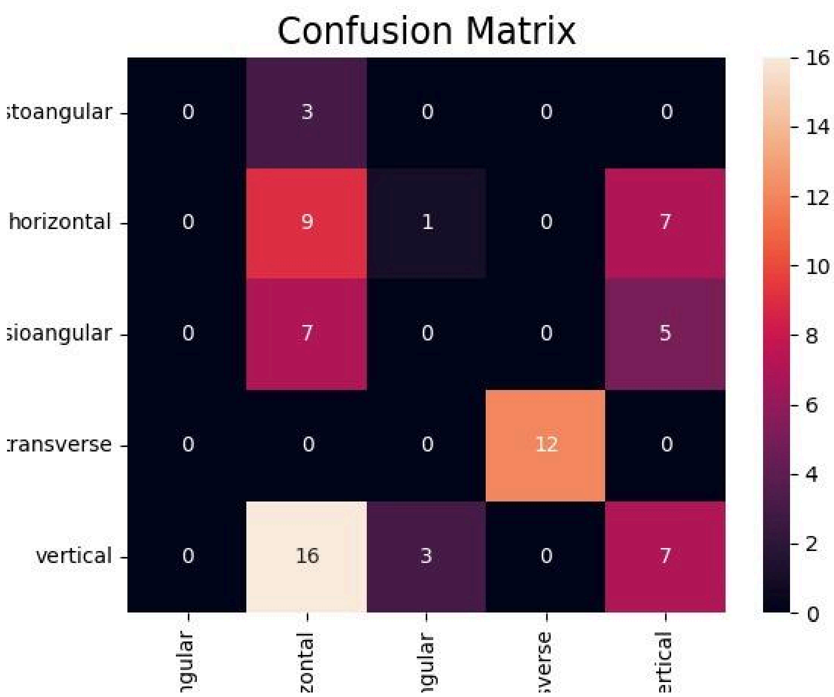

Figure 12: Confusion matrix of VGG-16

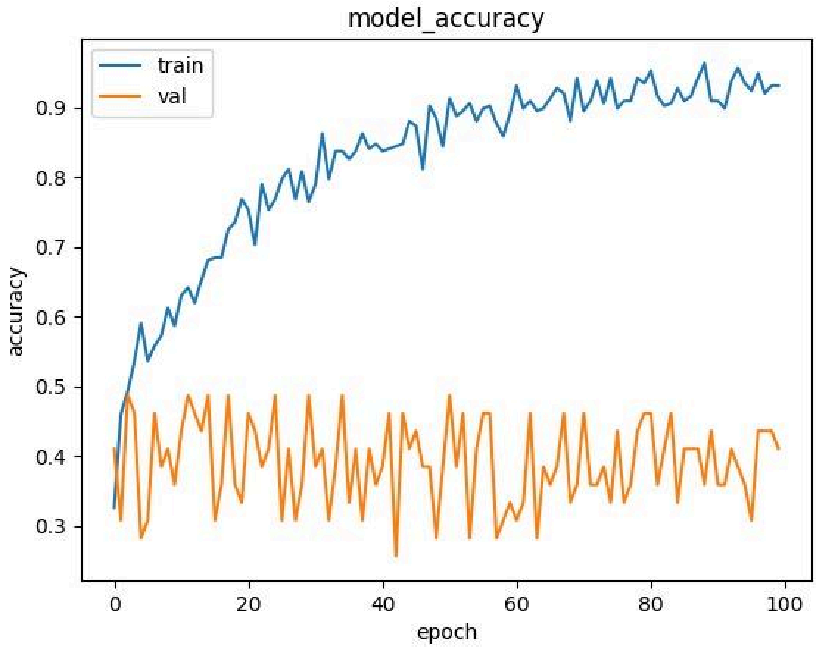

Figure 13: Model accuracy of VGG-16

It's important to note that the blue line (training accuracy) will almost always increase over time, as the model is designed to learn from the training data. However, the orange line (validation accuracy) is a better indicator of how well the model is generalizing to new data. Ideally, we want to see both lines increasing over time, but if the validation accuracy is not increasing, it's a sign that the model may need to be adjusted.

CHAPTER FOUR

RESULTS & DISCUSSION

4.1 Evaluation of GAN

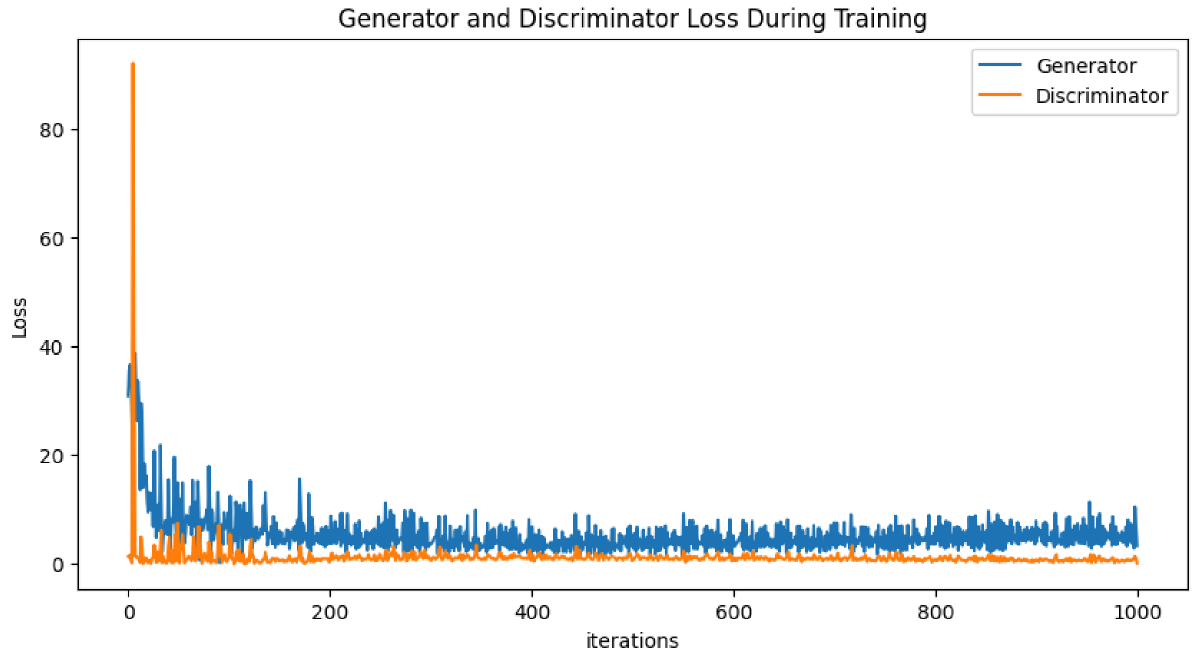

Figure 14: Model accuracy of VGG-16

At the beginning of the training, the discriminator's loss is high because it’s learning to differentiate between real and generated images. As training progresses, both losses stabilize, indicating that the generator is producing more realistic images that fool the discriminator, and they are in a sort of equilibrium. The generator’s loss starts very high but decreases rapidly, then fluctuates around lower values. The discriminator’s loss also starts relatively high but stabilizes quicker than the generator’s loss. Both losses seem to stabilize after approximately 200 iterations, with minor fluctuations.

4.2 Evaluation of ResNet-50 After Implementation of GAN

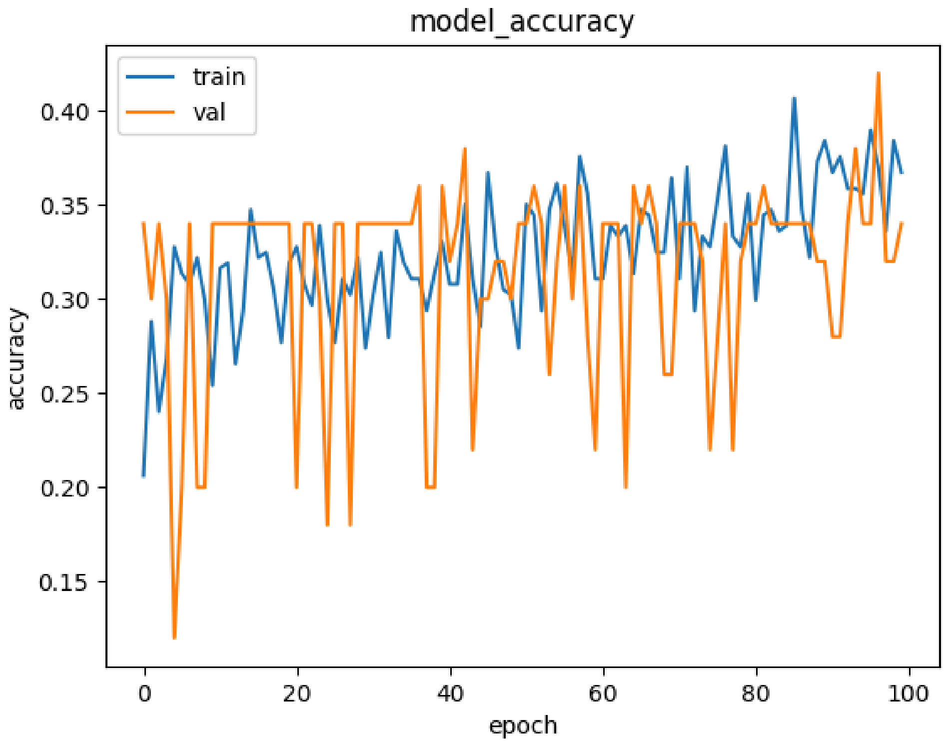

Figure 15: Model accuracy of VGG-16

The training accuracy (blue line) increases steadily over time, indicating that the model is learning from the training data. However, the validation accuracy (orange line) fluctuates and does not show a consistent increase, suggesting that the model may not be generalizing well to unseen data. This could be due to overfitting, where the model is too complex and has learned to fit the training data too closely, resulting in poor performance on new data.

4.3 Evaluation of VGG-16 After Implementation of GAN

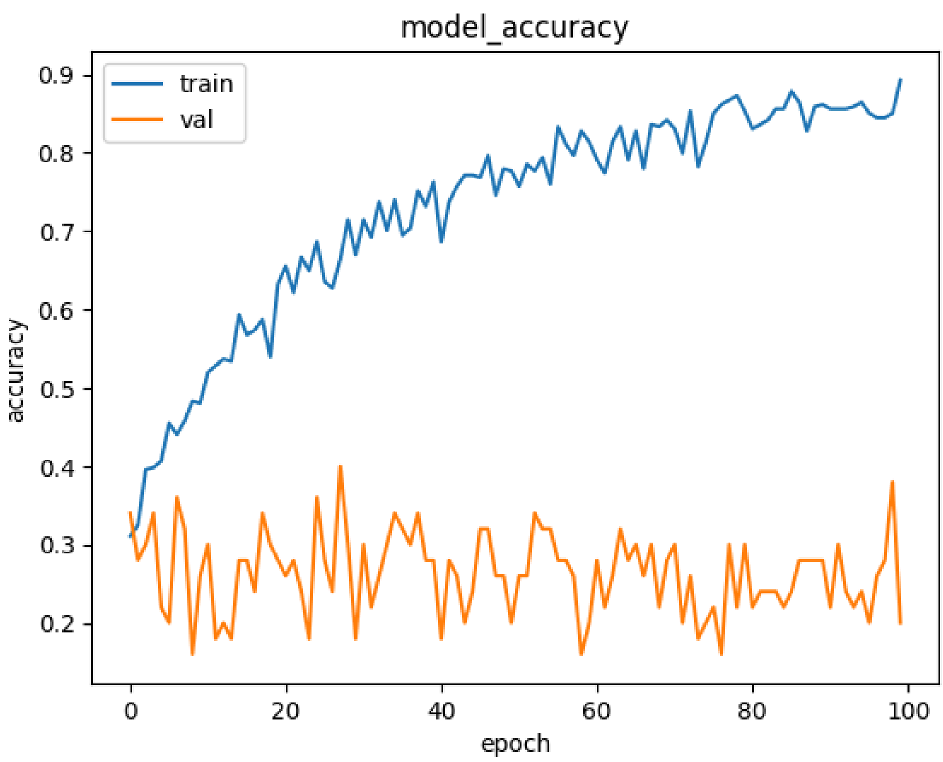

Figure 16: Model accuracy of VGG-16

The training accuracy (blue line) is relatively stable, fluctuating between approximately 0.35 and 0.4. The validation accuracy (orange line) is more volatile, with significant fluctuations but generally increasing trend. Compared to the previous graph you sent, the training accuracy seems to be similar, while the validation accuracy is more stable and shows a consistent increase over time, indicating that the model is generalizing better to unseen data. This is a good sign that the GAN has helped improve the performance of the model.

4.4 Conclusion

Comparing the model accuracy before and after the implementation of GAN for both CNN models, it seems that the training accuracy of the models is generally high and stable, while the validation accuracy fluctuates significantly and does not show a consistent increase. This suggests that the models may be **overfitting** to the training data and **not generalizing well** to new data.

CHAPTER FIVE

FUTURE WORKS

5.1 Future Works

1) **Exploring different image classification models like Pix2Pix and CycleGAN** adds a fresh perspective to our study in computer vision. These models, famous for their image-to-image translation using generative adversarial networks (GANs), provide a unique approach. While Pix2Pix is great for paired data, CycleGAN extends its utility to unpaired datasets. This inclusion allows us to compare their performance with traditional classification architectures.

2) **Emphasizing the use of larger image sizes (e.g., 300 x 300 pixels) during training and testing is crucial.** The standard 128 x 128 pixels resolution might lose important features during downscaling. By opting for larger images, we're aiming to preserve intricate details, reducing the risk of information loss and giving our model a better shot at recognizing complex patterns.

3) **Diversifying our convolutional neural network (CNN) architectures with ResNet101, VGG19, and DenseNet is a strategic move.** ResNet101, with its deep residual learning, is expected to capture detailed hierarchical features. VGG19, being simpler yet deep, offers a different perspective, and DenseNet's dense connectivity structure ensures efficient information flow. This variety in CNN models allows us to evaluate their impact on accuracy, robustness, and generalization in our image classification tasks.

REFERENCES

1) Buyuk, C., Akkaya, N., Arsan, B., Unsal, G., Aksoy, S., & Orhan, K. (2022). A fused deep learning architecture for the detection of the relationship between the mandibular third molar and the mandibular canal. Diagnostics, 12(8), 2018. https://doi.org/10.3390/diagnostics12082018

2) Liu, M.-Q., Xu, Z.-N., Mao, W.-Y., Li, Y., Zhang, X.-H., Bai, H.-L., Ding, P., & Fu, K.-Y. (2021). Deep learning-based evaluation of the relationship between mandibular third molar and mandibular canal on CBCT. Clinical Oral Investigations, 26(1), 981–991. https://doi.org/10.1007/s00784-021-04082-5

3) Maruta, N., Morita, K., Harazono, Y., Anzai, E., Akaike, Y., Yamazaki, K., Tonouchi, E., & Yoda, T. (2023). Automatic machine learning-based classification of mandibular third molar impaction status. Journal of Oral and Maxillofacial Surgery, Medicine, and Pathology, 35(4), 327–334. https://doi.org/10.1016/j.ajoms.2022.12.010

4) Zhu, T., Chen, D., Wu, F., Zhu, F., & Zhu, H. (2021). Artificial intelligence model to detect real contact relationship between mandibular third molars and inferior alveolar nerve based on panoramic radiographs. Diagnostics, 11(9), 1664. https://doi.org/10.3390/diagnostics11091664

5) Choi, E., Lee, S., Jeong, E., Shin, S., Park, H., Youm, S., Son, Y., & Pang, K. (2022). Artificial Intelligence in positioning between mandibular third molar and inferior alveolar nerve on panoramic radiography. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-06483-2

APPENDICES

A. GANTT CHART

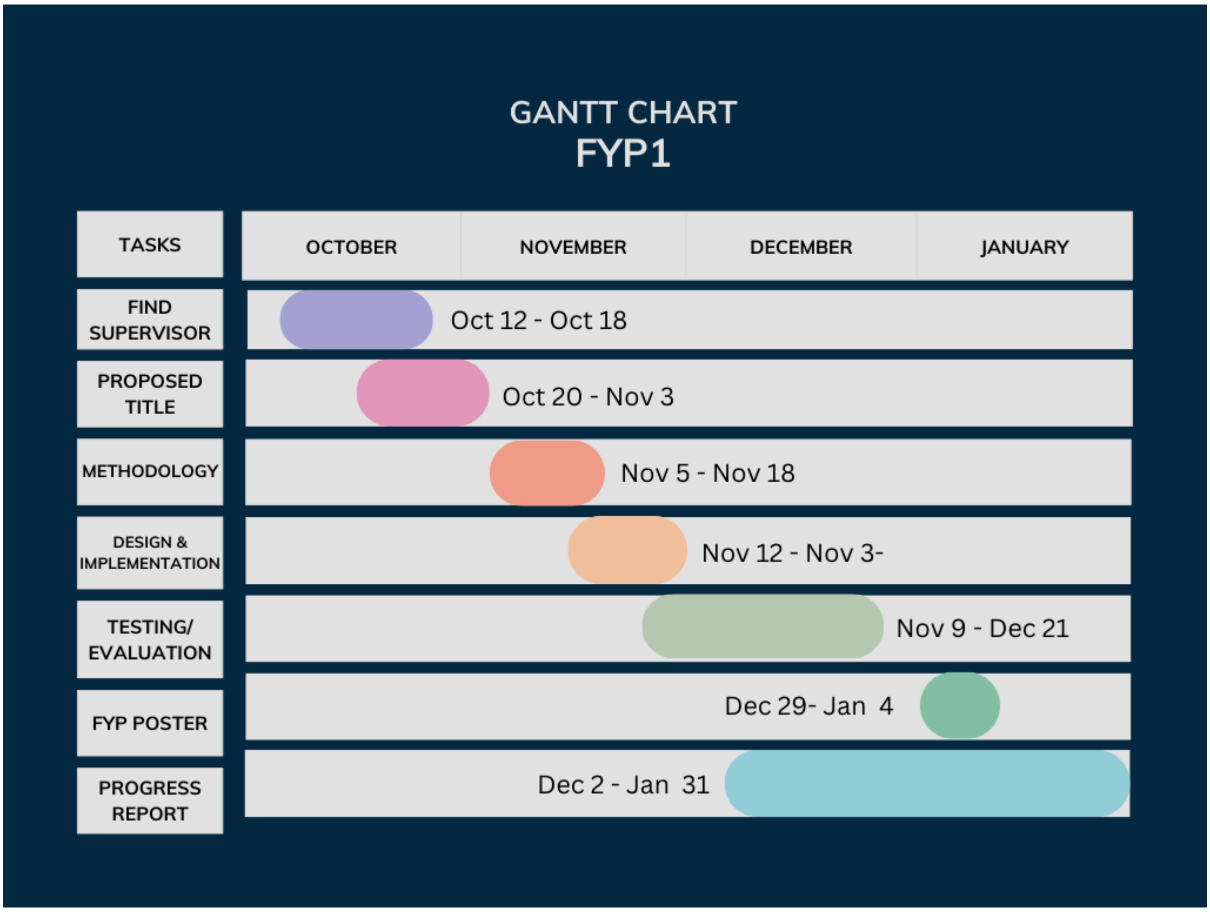

Figure 1: Gantt Chart of CNN model development.