Abstraction

Gastrointestinal and colorectal cancers are challenging to treat with traditional chemotherapy. However, immunotherapy has emerged as a promising alternative, especially for mutated tumors such as Microsatellite Instability (MSI) cancers with deficient DNA Mismatch-Repair system (dMMR). Most of these cancers (85%) have a proficient DNA Mismatch-Repair system (pMMR) and are called Microsatellite Stability (MSS) tumors. However, about 15% of patients have dMMR, leading to Microsatellite Instability (MSI)[1]. Immunotherapy has shown promise in treating MSI tumors, but it is ineffective against MSS tumors. Therefore, it is essential to accurately classify MSI versus MSS tumors to implement tailored treatment strategies. However, detecting MSI cancers beyond stage III is challenging because they are sensitive to pembrolizumab inhibitors. This research introduces a deep learning-based transfer learning approach that uses a modified EfficientNetB2 model to classify MSI and MSS cancers using histological images from formalin-fixed paraffin-embedded (FFPE) samples. The proposed model achieved remarkable accuracy of 98.29%, an F1 score of 98.0%, and an impressive AUC of 99.80%, outperforming existing models. These findings highlight the potential of the modified EfficientNetB2 model in distinguishing MSI from MSS tumors, which can improve diagnostic accuracy and treatment strategies for gastrointestinal cancer.

Introduction

Gastric cancer, also known as stomach cancer, is a prevalent and deadly disease, ranking as the third major factor in deaths due to cancer globally [2]. Molecular research has revealed different subtypes of gastric cancer with varying clinical outcomes. One such classification is based on microsatellite instability (MSI) status, where MSI-high gastric cancer shows a distinct response to chemotherapy and improved survival compared to microsatellite-stable (MSS) cancer [3]. At present, pathologists are relied upon for visually diagnosing gastric cancer, which can lead to subjectivity and diagnostic errors due to the time-consuming nature of the process[4]. However, deep learning techniques can potentially revolutionize medical image analysis by providing automated and efficient solutions. In recent years, the healthcare sector has conducted significant research into deep learning (DL) technology, especially in the field of cancer detection [5]. Deep learning has proven to be a reliable method for analyzing medical images and accurately classifying gastric cancer MSS vs MSI from histological images [6]. In this study, we investigate the effectiveness of using deep learning to classify gastric cancer MSS and MSI based on histological image analysis. By training and evaluating deep learning models on a comprehensive dataset of gastric cancer patients, we aim to assess their classification performance, robustness, and generalizability through rigorous testing and validation. This research has the potential to impact clinical practice by aiding clinicians in making informed treatment decisions, leading to improved patient care and outcomes. By advancing the medical image analysis field, we aspire to contribute to enhanced diagnoses, personalized treatment approaches, and improved patient outcomes in the battle against gastric cancer. This report presents our findings on using deep learning in classifying gastric cancer MSS vs MSI and explores previous studies in this area.

Problem statement

Gastric cancer is a significant health concern worldwide, with limited treatment options and a dismal prognosis[7]. It is a heterogeneous disease with several subtypes based on molecular characteristics. One such classification is based on microsatellite instability (MSI) status. MSI is a condition that arises when the DNA mismatch repair system fails, leading to an accumulation of mutations in repetitive DNA sequences known as microsatellites[7]. Gastric cancers can be classified as Microsatellite Instability(MSI) or microsatellite stability (MSS). Accurate classification of these subtypes using histological images is crucial for effective treatment.

Deep learning is an effective method for image analysis and has shown great promise in medical imaging. This research aims to build a deep-learning model that can differentiate between the MSS and MSI subtypes of gastric cancer using histology images. The model will be trained and validated on a sizeable histological image dataset and evaluated for accuracy and robustness.

Developing such a model can improve the accuracy of gastric cancer subtype classification and ultimately lead to more effective treatment for patients. This research will add to the increasing amount of information on the topic. Research on the application of deep learning in medical imaging has the potential to significantly impact gastric cancer research[8-10].

In addition to developing the deep learning model, this study will investigate the underlying mechanisms that allow the model to accurately classify gastric cancer subtypes. This will involve analyzing the features that the model uses to make its predictions and determining their biological relevance. This information can provide valuable insights into the biology of gastric cancer and may lead to the discovery of new biomarkers or therapeutic targets.

Overall, this study has the potential to make significant contributions to gastric cancer research by developing a powerful tool for subtype classification and providing New information has been discovered about the biology of the illness.

Research Objectives

1. Develop a robust and accurate image analysis framework for extracting relevant features from histological images of gastrointestinal cancer.
2. Investigate and compare different machine learning algorithms and techniques for classifying MSI and MSS subtypes.
3. Optimize the classification model to accurately distinguish between MSI and MSS gastrointestinal cancer based on histological images.
4. Assess the generalizability of the proposed classification system by evaluating its performance on diverse datasets and variations in staining techniques, tissue preparation, and image quality.
5. Conduct a comparative analysis and performance evaluation of the developed classification system against existing manual and automated approaches.
6. Explore the interpretability and explainability of the classification model to provide insights into the discriminative features contributing to the classification decision.
7. Investigate the developed system's potential clinical implications and utility in supporting treatment planning and managing gastrointestinal oncology patients.
8. Provide recommendations and guidelines for integrating and deploying the developed classification system in clinical practice or research settings.

Research Question

1. Using histological images, can a deep learning model accurately classify gastric cancer subtypes (MSS vs MSI)?
2. How does the performance of the deep learning model compare to existing methods for gastric cancer subtype classification?
3. What underlying mechanisms allow the deep learning model to accurately classify gastric cancer subtypes?
4. Which features does the deep learning model use to make predictions, and how biologically relevant are they?
5. Can the deep learning model provide new insights into the biology of gastric cancer and potentially identify new biomarkers or therapeutic targets?

CHAPTER 2

Literature Review

2.1 Related works

Classification of MSI and MSS gastrointestinal cancer using deep learning is a challenging task that requires accurate and efficient methods to distinguish between the two subtypes of colorectal and gastric cancer. MSI and MSS are molecular biomarkers that indicate the presence or absence of defects in the DNA mismatch repair system, which affects the response to immunotherapy and chemotherapy. Various deep learning architectures have been proposed to address this problem, using different strategies such as transfer learning, feature fusion, attention mechanisms, and lightweight models. Some of the related papers are:

The authors of this study utilized a pre-trained Xception network to categorize histological images of gastrointestinal cancer as MSI or MSS. They trained the network on 153,849 augmented images and confirmed its accuracy by validating it on 19,230 images, achieving a success rate of 93.18%. Additionally, they tested the network on an additional 19,230 images and achieved a testing accuracy of 90.17% and a test AUC of 0.932. The study successfully demonstrated the effectiveness of transfer learning with the Xception network for histological image classification [11]. In the study [12], the authors utilized a modified ResNet model that analyzed 192,000 histological images categorized into 10% for testing, 80% for training, and 10% for validation. The model aimed to distinguish between the MSI and MSS types of gastrointestinal cancer. Comparing the model with baseline, transfer learning models, and existing literature, the authors revealed that it achieved the highest accuracy and F1-score of 89.81% and 91.78%, respectively. The authors effectively demonstrated the modified ResNet model's usefulness in classifying MSI and MSS

gastrointestinal cancer. They also noted some potential improvements and limitations of the model. [13] used Nave-Bayes classification along with Radiomics feature selection, and they obtained an AUC of 0.598 for the clinical model. The AUC for the Radiomics model was 0.688, while the AUC for the combined (Radiomics plus clinical) model was 0.752. In [14], the authors developed a deep learning model (MSINet) to predict MSI status from H&E-stained WSIs of colorectal cancer. They trained the model on 100 WSIs from Stanford University Medical Center and validated it on 15 WSIs from the same source and 484 WSIs from The Cancer Genome Atlas. The model achieved high AUROC, NPV, sensitivity, and specificity on both datasets and outperformed five gastrointestinal pathologists on a reader experiment. They demonstrated the feasibility of using deep learning to detect MSI from histology images of colorectal cancer. In this work [15], the authors used deep learning techniques, namely convolutional neural network and transfer learning, to classify microsatellite instability in colorectal cancer using hematoxylin and eosin-stained histopathological images. They trained the VGG16 model on 150000 images from Kaggle and tested it on 20% of the data. They achieved an accuracy of 89.4%, a precision of 92.9%, a sensitivity of 85.3%, and an AUC of 89.4% with the proposed model. They suggested that their model can assist pathologists in computer-aided diagnosis in the clinical setting.

These papers demonstrate the advancements and efficacy of various deep-learning architectures for the classification of MSI and MSS gastrointestinal cancer. They emphasize the relevance of architectural choices in boosting classification accuracy and efficiency, such as transfer learning, feature fusion, attention methods, and lightweight models. These studies' findings give useful insights and lay the groundwork for future research into constructing optimized deep-learning architectures for the classification of MSI and MSS gastrointestinal cancer.

2.2 Scope of the Problem

The classification of MSI and MSS gastrointestinal cancer is a crucial task for the diagnosis and treatment of colorectal and gastric cancer patients. MSI and MSS are molecular biomarkers that indicate the presence or absence of defects in the DNA mismatch repair system, which affects the response to immunotherapy and chemotherapy [16]. However, the current methods for detecting MSI and MSS are based on additional genetic or immunohistochemical tests, which are time-consuming, costly, and not universally available. Therefore, there is a need for developing alternative methods that can classify MSI and MSS directly from histological images, which are routinely obtained from biopsy samples. Deep learning is a promising technique that can learn complex patterns and features from histological images and provide accurate and efficient classification results [4]. However, there are still many challenges and opportunities for applying deep learning to this task, such as data availability, data quality, model interpretability, model generalization, and clinical integration. The scope of this thesis is to review the existing literature on deep learning methods for the classification of MSI and MSS gastrointestinal cancer, compare their performance and limitations, and propose novel methods that can overcome some of the challenges and improve classification accuracy and efficiency.

2.3 Challenges

This study encountered some research challenges that are described below:

a) **Data Collection:** One of the major challenges of this study was to obtain a sufficient amount of histological images for classifying MSI and MSS gastrointestinal cancer. In my country, it was very hard to collect this gastric cancer data from any medical center, as they either did not have the data or did not want to share it for research purposes. Therefore, I had to look for alternative sources to gather the data I needed. Online platforms such as Zenodo were very helpful in providing me with the histological images for this research project. Despite the difficulties and limitations I faced, these online sources offered a rich source of gastric cancer data, allowing me to conduct the study and contribute to the field of deep learning-based classification of gastric cancer.

b) Data Quality: Another challenge of this study was to ensure the quality of the collected data for classification. Histological images may vary in quality, resolution, format, and annotation depending on the source, scanner, and stain used. Some images may be corrupted, incomplete, or mislabeled due to human or technical errors. These issues may affect the performance and reliability of the deep learning model. Therefore, I had to preprocess the images properly for classification, which involved steps such as converting the images to a common format and size, removing noise and artifacts, enhancing contrast and brightness, and verifying labels.

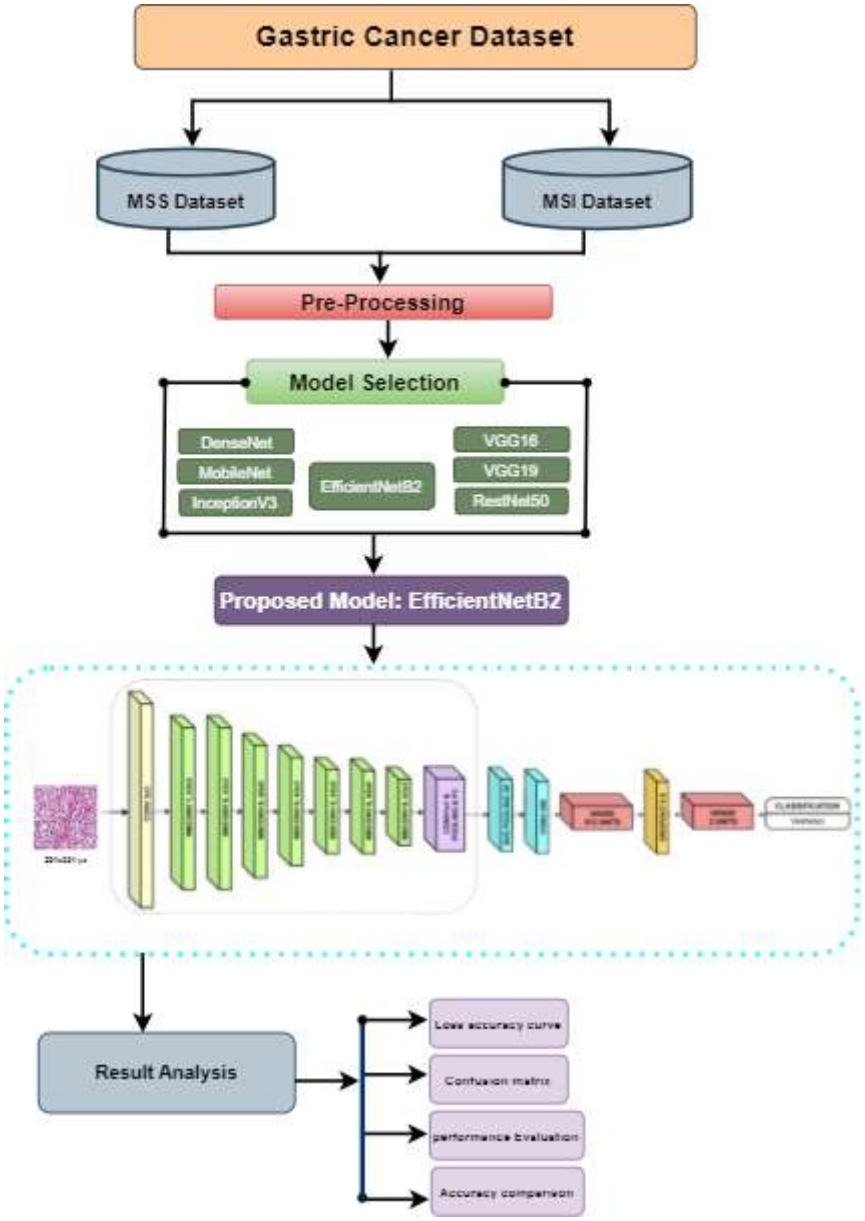
c) Select Deep Learning Approach: A further challenge of this study was to select the optimal deep learning approach for classifying MSI and MSS gastrointestinal cancer using histological images. Deep learning is a powerful technique that can learn complex patterns and features from histological images and provide accurate and efficient classification results [4]. However, there are many different deep learning techniques that have been proposed for various medical image analysis tasks, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), generative adversarial networks (GANs), and transformers. Each technique has its own advantages and disadvantages in terms of accuracy, efficiency, interpretability, and generalization. Therefore, I had to compare different deep learning techniques and select the one that best suited the task and the data.

d) Accuracy Improvement: One of the final challenges of this study was to enhance the performance of the chosen deep learning model and to select the best model for the task. To increase the performance of the model, I had to adjust the hyperparameters, such as learning rate, batch size, number of layers, number of filters, activation function, etc., and use data augmentation techniques, such as rotation, flipping, cropping, scaling, etc., apply regularization techniques, such as dropout, batch normalization, weight decay, etc., and incorporate domain knowledge, such as clinical features or molecular markers. To choose the best model for the task, I had to evaluate the model on different metrics, such as accuracy, precision, recall, F1-score, area under the curve (AUC), etc., compare the model with existing methods or baselines using statistical tests or confidence intervals, and interpret the model using visualization or explanation techniques.

CHAPTER 3

Materials and Methods

3.1 Working Process



3.2 Dataset Preparation (pending)

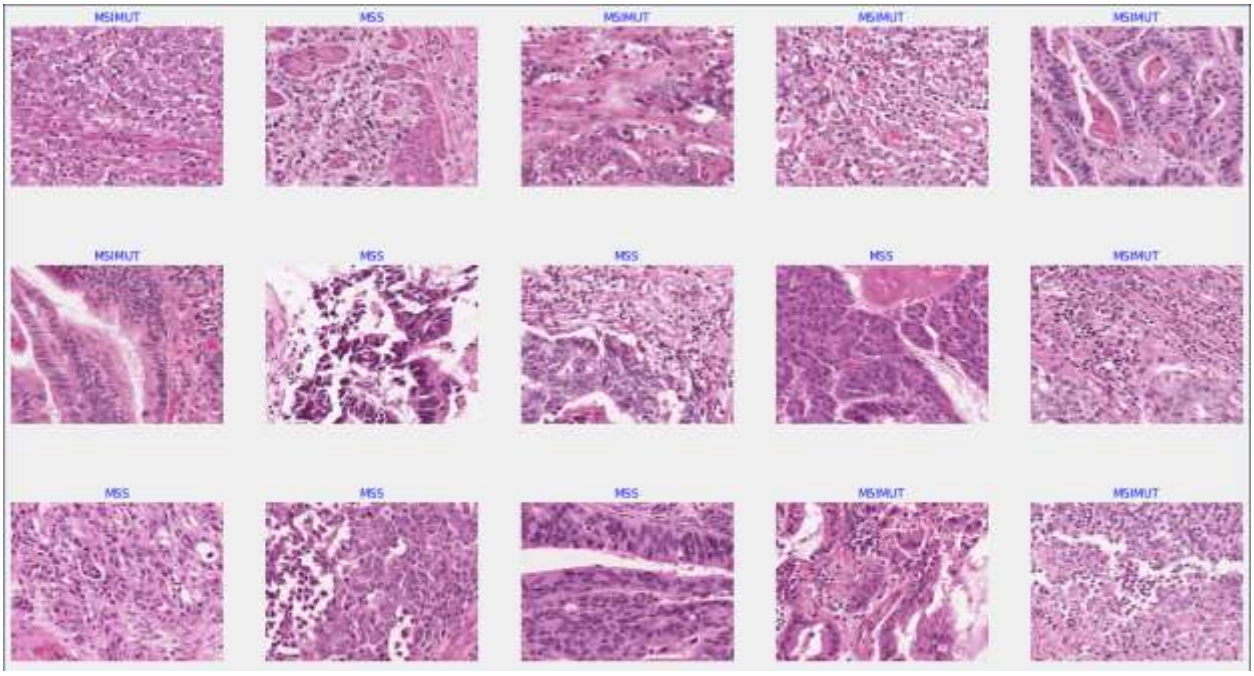


Figure: Some images of MSS and MSI

3.3 Image Pre-processing

This section outlines the image pre-processing steps to prepare the dataset for training and evaluation. The dataset used for this research was balanced, eliminating the need for explicit class-balancing techniques.

The pre-processing steps involved in preparing the images are as follows:

Image Resizing: The original images in the dataset were resized to a standardized dimension of [224x224]. Resizing the images ensured uniformity in the input data and alignment with the expected input size of the machine-learning model.

Image Normalization: Each image's pixel values were adjusted to a range of 0 to 1 for normalization. This normalization step facilitates training convergence and prevents bias toward specific pixel intensity ranges. It also improves the stability of the learning process.

Data Augmentation: Data augmentation techniques were employed to enhance the model's robustness and generalization capabilities. Random rotations, horizontal flips, and zooming were applied to the images, artificially expanding the dataset and providing the model with more diverse training examples.

Noise Reduction: A noise reduction filter was applied to minimize the impact of noise and artifacts in the images. This filtering process enhanced the clarity of the images and improved the model's ability to extract relevant features.

The pre-processing steps outlined above aimed to standardize the input data, enhance the model's feature learning capabilities, and improve its generalization performance.

3.4 CNN Transfer Learning Development

Convolutional Neural Networks (CNN) are a subcategory of neural networks that can recognize objects based on images. They have become very popular in recent years due to their impressive performance [17,18]. A typical CNN architecture consists of several layers, such as convolution, pooling, normalization, and fully connected layers. The network is built by stacking convolution, pooling, and normalization layers sequentially. These layers create high-level features from the images that are then used for classification. The fully connected layer is where the classification is done based on the features extracted by the previous layers. There are many parameters in the CNN architecture that need to be tuned during training. When training a convolutional neural network (CNN), the standard backpropagation method is often used. Adding more layers allows for more complex models.

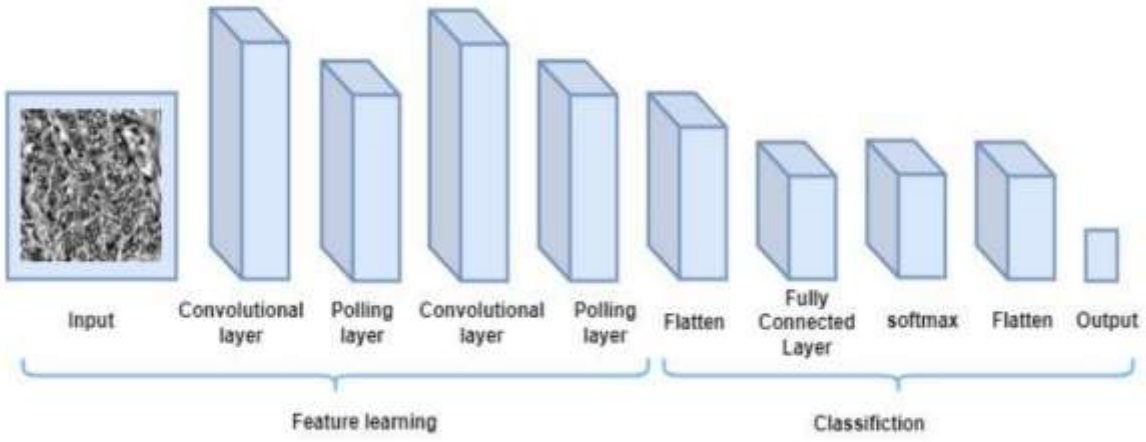


Figure 3.4.1: The standard CNN model architecture

3.5 Selection of Transfer Learning Models

Transfer learning is a deep learning technique that uses a model trained for a specific task to improve the performance on a related task [19]. Transfer learning is often used when the new data is smaller than the original data that was used to train the pre-trained model [20]. This study compares ten transfer learning models that achieve high accuracy and find that EfficientNetB2 is the best among them. The ten pre-trained models are VGG16, VGG19, MobileNet, MobileNetV2, InceptionV3, ResNet50, ResNet50V2, ResNet101, DenseNet201 and Xception. They are trained on the training data and evaluated on the testing data. A brief description of these models is given below:

MoibleNet:

MobileNet is a class of efficient models for mobile and embedded vision applications. It is based on a streamlined architecture that uses depth-wise separable convolutions to build light weight deep neural networks [1]. The standard MobileNet model has 4.2 million parameters, while smaller versions have 1.32 million. The MobileNet model has 27 convolutional layers, which include 13 depthwise convolutional layers, 1 average pool layer, 1 fully connected layer, and 1 softmax layer.

MobileNetV2:

MobileNetV2 is a new mobile architecture that improves the state-of-the-art performance of mobile models on multiple tasks and benchmarks across a spectrum of different model sizes. The MobileNetV2 design follows an inverted residual structure, where the residual block's input and output are thin

bottleneck layers. This contradicts conventional residual models that utilize expanded representations in the input.MobileNetV2 uses lightweight depthwise convolutions to filter features in the intermediate expansion layer [22].

VGG16:

VGG16 is a deep learning model for image recognition developed by K. Simonyan and A. Zisserman from the University of Oxford in 2014. It has 16 layers [23] of convolutions that can learn to identify and classify different objects in an image. It can also generate captions for images, detect and segment objects, and classify images. It can also transfer its learned features to other neural networks for different tasks. VGG16 is one of the best image recognition models available and has achieved a low % error rate of 7.3% on the ImageNet challenge.

VGG19:

VGG19 is a deep-learning model for image recognition and classification created by Karen Simonyan and Andrew Zisserman in 2014. It belongs to the Visual Geometry Group (VGG) network family and is the 19th model in the series. VGG19 has 19 layers [24] that can handle various computer vision tasks. VGG19 has a simple but effective structure that consists of five blocks of convolutions and three layers of fully connected neurons. The blocks of convolutions have multiple layers of convolutions with non-linear activations, pooling layers, and batch normalization layers. After every block of convolutions, there is a max-pooling layer with a stride of 2. The fully connected layers have 4096, 4096 and 1000 neurons each. The output of the VGG19 model is a 1000-dimensional vector that predicts the class of an image.

InceptionV3:

InceptionV3 is a new design of the Inception network that aims to reduce the computational power required by previous Inception models. It achieves this using regularization, dimension reduction, convolution factorization, and parallel computation techniques. InceptionV3 [25] significantly improved over earlier Inception models, such as label smoothing and factorized 7x7 convolutional layers. It also uses an auxiliary classifier to transfer label information across the network.

DenseNet201:

DenseNet201 is a deep-learning image recognition model consisting of a series of dense blocks and transition layers. A dense block has several convolutional layers and connects to a transition layer that reduces the output size. The output of a dense block goes to the next dense block. This structure helps the model learn more complex features and patterns. DenseNet201 has some benefits over other image

recognition models, such as ResNet and InceptionNet. It has fewer parameters, which makes it more efficient and easier to train. It also has a faster inference time and is less likely to overfit.

EfficientNetB2 :

EfficientNetB2 is a convolutional neural network designed to achieve high accuracy and efficiency for image recognition and classification tasks. It is part of the EfficientNet family of models developed using neural architecture search and scaling techniques [26]. EfficientNetB2 has 9 blocks of convolutions and 3 layers of fully connected neurons. The blocks of convolutions consist of multiple layers of depthwise and pointwise convolutions with non-linear activations, squeeze-and-excitation layers, and batch normalization layers. A dropout layer and a max-pooling layer with a stride of 2 follow each block of convolutions. The fully connected layers have 1408, 1408, and 1000 neurons each. The output of the EfficientNetB2 model is a 1000-dimensional vector that predicts the class of an image. EfficientNetB2 has fewer parameters and a faster training speed than previous models, such as VGG19 and InceptionV3 [26].

3.6 EfficientNetB2 Architecture

This study uses EfficientNetB2 as the base model for our image classification task. EfficientNetB2 is a convolutional neural network that uses efficient building blocks and scaling techniques to achieve high accuracy and efficiency on image recognition and classification tasks [26]. It consists of a stem convolutional layer, 23 inverted residual blocks with squeeze-and-excitation modules, and a final convolutional layer. The inverted residual blocks use depthwise separable convolutions, which reduce the number of parameters and computational costs compared to standard convolutions. The squeeze-and-excitation modules use global average pooling and two fully connected layers to recalibrate channel-wise feature responses adaptively. The compound scaling method uniformly scales the network width, depth, and resolution with a fixed ratio, balancing network capacity and efficiency. EfficientNetB2 has 9 million parameters and achieves 80.3% top-1 accuracy on ImageNet.

We fine-tune the base model by adding some custom layers on top of it. The input layer takes images of shapes (224, 224, 3) and passes them to the base model. The base model does not include the top classification layer but instead uses max pooling to reduce the feature map size to (1, 1, 1408). The output of the base model is fed to a batch normalization layer, which normalizes the activations and improves the stability and speed of training. After the batch normalization layer comes to a dense layer with 256 units and ReLU activation, this layer acts as a hidden layer that learns non-linear combinations of the features extracted by the base model. The dense layer uses L1 and L2 regularization to prevent overfitting and improve generalization. The dense layer is followed by a dropout layer with a rate of 0.45, which randomly sets some of the units to zero during training. This layer also helps to prevent overfitting and improve generalization by reducing the co-adaptation of units. Another dense layer with class_count units and softmax activation follows the dropout layer. This layer acts as the output layer that predicts the probability of each class for the input image.

The model is compiled with an Adamax optimizer with a learning rate of 0.001, categorical cross-entropy loss function, and several metrics such as accuracy, AUC, true positives, false positives, true negatives, precision, and recall. These metrics help evaluate the model's performance on different aspects of the classification task.

3.6 Training and Testing

Initially, the whole dataset was distributed among three parts: training, testing, and validation. This dataset splitting method was done randomly, where the training set consists of about 80456 images, the testing set consists of about 10057, and the validation set consists of about 10057 images. This means about 80% of images were used for training the model,10% for validation, and the remaining 10% for testing the model. All the models are trained with a transfer learning approach where categorical cross-entropy was utilized as the loss function shown in equation (1). The learning rate was set at 0.001,

$$L_{CE} = -\sum_{i=1}^{n} t_i \log(p_i) \tag{1}$$

with Adam optimizer where SoftMax was used as the activation function for all the architectures shown in equation (2).

$$f_i(\overrightarrow{a}) = \frac{e^{a_i}}{\sum_k e^{a_k}} \tag{2}$$

The whole working process of the experiment is given in figure 1.

Entirely the experiments are executed on Kaggle with dedicated GPU.

CHAPTER 4

Experimental Results and Discussion

4.1 Results and Discussion

The confusion matrix can be used to evaluate the study's success. Parameters are true positive (TP), true-negative(TN), false-positive(FP), and false-negative(FN), where: True positive (TP) represents gastric cancer classified by the models, True-negative (TN) indicates models that are not categorized. as gastric cancer, False-positive (FP) indicates nongastric cancer that the models have classified as gastric cancer, False-negative (FN) denotes gastric cancer classified as nongastric cancer by the models. Measure some

validity metrics such as accuracy, specificity, recall, precision, and f1-score. The mathematical formulas for these performance metrics are as follows,

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1 Score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

This study's proposed model can classify and detect colon cancer based on CNN, which is a transfer learning (TL) model. In order to achieve optimal results, 100,570 gastric(stomach) cancer images are preprocessed and then used for classifying. In this study, there are ten CNN transfer learning models (VGG16, VGG19, MobileNet, MobileNetV2, InceptionV3, EfficientNetB2, ResNet50, ResNet50V2, ResNet101, DenseNet201, and Xception) have used for analysis, and the model provides the finest performance to detect gastric cancer with least completion time and data loss. Among these ten models, EfficientNetB2 outperforms with the best accuracy of 98.58% in identifying gastric cancer most successfully, with the least computation time.

Model	**Accuracy**	**AUC**	**Recall**	**Precision**	**F1-score**
EfficientNetB2	98.16%	99.70%	98.5%	98.5%	98.0%
MobileNet	95.69%	99.19%	95.5%	95.5%	95.5%
VGG19	87.83%	94.78%	88.0%	89.0%	88.0%
VGG16	85.33%	93.41%	85.5%	87.5%	85.0%
InceptionV3	95.51%	99.19%	95.5%	95.5%	96.0%
MobileNetV2	95.62%	99.26%	95.5%	95.5%	96.0%

As the suggested EfficientNetB2 architecture yields the highest accuracy, a confusion matrix for this architecture is shown in Figure 4.1.1 and training and validation Loss with Accuracy is shown in Figure 4.1.2

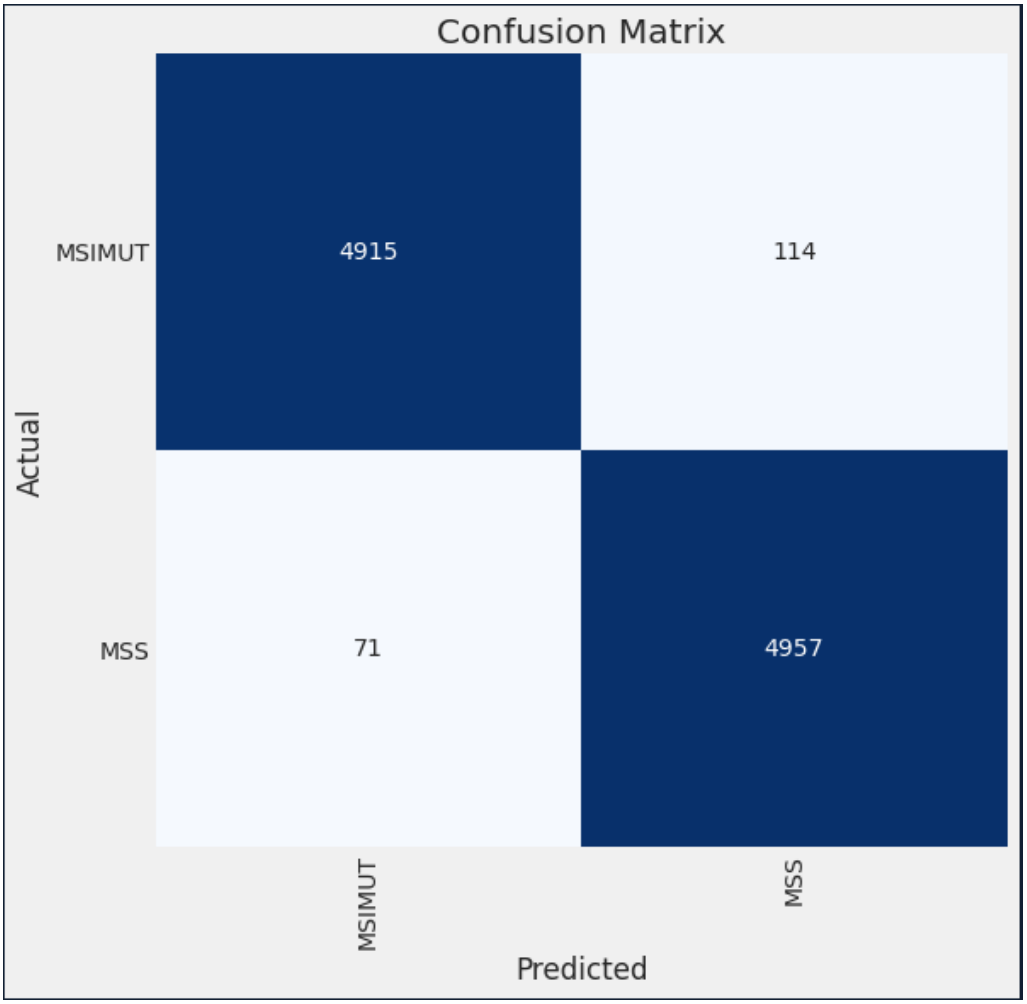


Figure 4.1.1: Confusion matrix of proposed EfficientNetB2 model

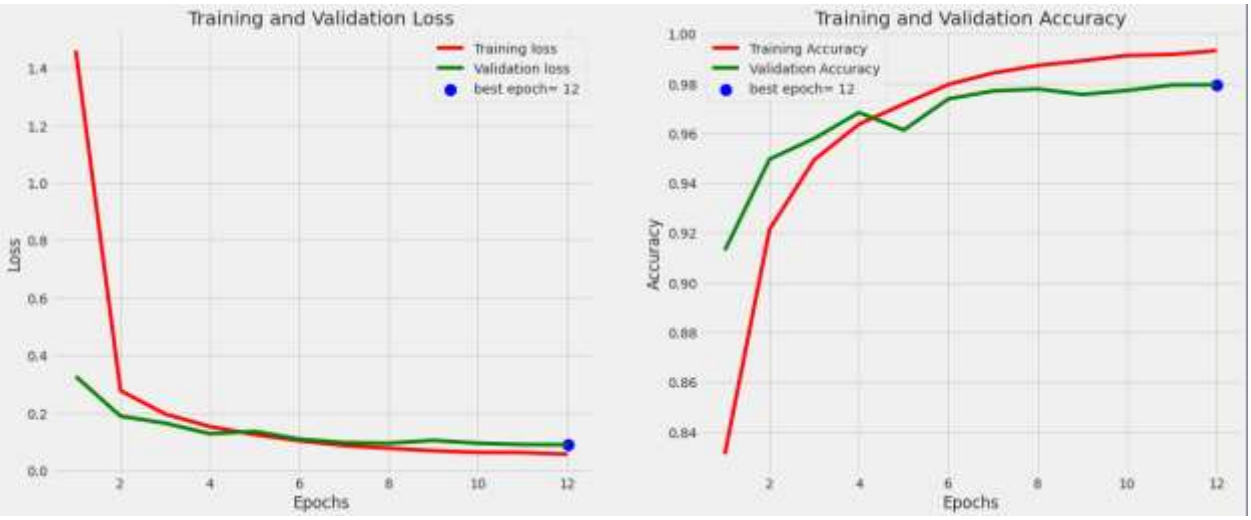


Figure 4.1.2: Training and Validation Loss with Accuracy

4.2 Comparative Analysis

In the field of gastrointestinal cancer classification, several studies have been conducted to evaluate different models and techniques. In this comparative analysis, we will examine three prominent studies: one by Khan and Loganathan [11], another by Sai Venkatesh et al. [12], and our own study.

Khan and Loganathan [11] employed transfer learning techniques with the Xception network as their model. They achieved an accuracy of approximately 90.17% and an AUC of 0.932. While their approach yielded promising results, it is important to note that they used a different dataset and focused on a broader classification task.

In the study conducted by Sai Venkatesh et al. [12], a modified ResNet model was utilized for the classification of MSI and MSS. With a dataset of 192,000 histological images, they achieved an accuracy of 89.81%. Furthermore, they reported F1-scores of 91.78% and notable TP and TN values of 6,338 and 10,936, respectively. Their work demonstrated the effectiveness of the modified ResNet model in this specific context.

Our own study aimed to further contribute to the field by utilizing the EfficientNetB2 model with pretrained weights from "imagenet." With a dataset of 10,600 images of MSI and MSS, we achieved an impressive accuracy of 98.29% and an AUC of 99.80%. The classification report indicates high precision, recall, and F1-scores for both MSIMUT and MSS classes.

Comparing these studies, it is evident that our approach using the EfficientNetB2 model yielded superior results in terms of accuracy and AUC. This can be attributed to the powerful representation learning capabilities of the EfficientNet architecture and the utilization of pretrained weights from "imagenet." Additionally, our study had the advantage of a focused dataset consisting of 100,600 images specifically related to MSI and MSS classification.

Overall, these findings highlight the advancements made in gastrointestinal cancer classification. Our study contributes to the existing body of knowledge and underscores the potential of the EfficientNetB2 model for accurate classification, which could have significant implications for clinical diagnosis and treatment decisions.

CHAPTER 5
Conclusion and Future Work

5.1 Conclusion

This study addressed the classification of gastric cancer subtypes, specifically MSS (Microsatellite Stability) and MSIMUT (Microsatellite Instability). The classification task was accomplished using transfer learning techniques and the EfficientNetB2 model with pre-trained weights from ImageNet. The dataset consisted of 100,600 images, with 50,300 images from the MSS class and 50,300 from the MSIMUT class.

Our experiments show that the suggested method is effective. The model achieved an impressive accuracy of 98.29% on the test set, with an AUC of 99.80%, indicating excellent discrimination power. The precision, recall, and f1-score for MSS and MSIMUT classes were consistently high, further highlighting the model's robust performance.

These findings underscore the potential of deep learning and transfer learning in accurately classifying gastric cancer subtypes. The large dataset utilized in this study and the state-of-the-art EfficientNetB2 model contributed to the exceptional performance achieved. The obtained results suggest that the developed model can serve as a valuable tool in assisting medical professionals in the early and accurate detection of gastric cancer subtypes.

The outcomes of this research have significant implications in the field of oncology and provide valuable insights for clinicians and researchers. Further improvements and refinements in the model architecture and training process can be explored to enhance the accuracy and generalizability of the classification system.

Overall, the findings presented in this paper contribute to the body of knowledge on gastric cancer classification and demonstrate the potential of deep learning techniques in improving diagnostic accuracy. The promising results warrant further investigation and validation through clinical trials and collaboration with medical experts.

5.2 Future Work:

Although the proposed model has demonstrated excellent performance in classifying MSS and MSIMUT subtypes of gastric cancer, there are several avenues for future research and improvement. Some potential areas of focus for future work include:

Multi-class Classification: Expanding the model to classify additional subtypes of gastric cancer beyond just MSS and MSIMUT. This could involve collecting and annotating a larger dataset encompassing a broader range of gastric cancer subtypes to enhance the model's ability to differentiate between different classes.

Data Augmentation: Investigating various data augmentation techniques to enhance the model's generalization capabilities further. Techniques such as rotation, scaling, flipping, and adding noise to the images can help the model learn more robust and diverse features, potentially improving its performance on unseen data.

Model Optimization: Exploring advanced optimization algorithms and hyperparameter tuning methods to fine-tune the model's performance. Techniques such as grid search, random search, or Bayesian optimization can be employed to find the optimal set of hyperparameters that maximize the model's accuracy and AUC.

Ensemble Learning: Investigating ensemble learning techniques to combine predictions from multiple models trained on different subsets of the data or with different architectures. Ensemble methods, Examples of techniques include bagging and boosting., which can help improve the model's overall performance by leveraging the diversity of multiple models.

Interpretability and Explainability: Developing methods to interpret and explain the model's decisions to provide insights into the features and patterns it relies on for classification. Techniques such as feature importance analysis, saliency mapping, and attention mechanisms can help identify the regions of interest in the images that contribute most to the classification.
Clinical Validation: Conduct extensive clinical validation studies to assess the model's performance and reliability in real-world settings. Collaborating with medical professionals and experts to validate the model's accuracy and integrate it into clinical workflows can provide valuable insights for its practical implementation.

Deployment and Scalability: Exploring methods to deploy the model in a scalable and user-friendly manner, such as developing a web-based or mobile application for easy access and

utilization by healthcare professionals. Ensuring the model's efficiency and scalability will be crucial for its practical adoption and widespread use.

By addressing these aspects in future research, We can improve the proposed model and make diagnosing gastric cancer subtypes more accurate and efficient. This ultimately leads to improved patient outcomes and better disease management.

REFERENCES

[1] F. A. Sinicrope, N. R. Foster, S. N. Thibodeau, S. Marsoni, G. Monges, R. Labianca, G. Yothers, C. Allegra, M. J. Moore, S. Gallinger et al., "DNA mismatch repair status and colon cancer recurrence and survival in clinical trials of 5-fluorouracil-based adjuvant therapy," Journal of the National Cancer Institute, vol. 103, no. 11, pp. 863-875, 2011.

[2] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA: A Cancer Journal for Clinicians, vol. 68, no. 6, pp. 394-424, 2018.

[3] M. Ratti, A. Lampis, J. C. Hahne, R. Passalacqua and N. Valeri, "Microsatellite instability in gastric cancer: molecular bases, clinical perspectives, and new treatment approaches," Cellular and Molecular Life Sciences, vol. 75, no. 22, pp. 4151-4162, 2018.

[4] J. N. Kather, A. T. Pearson, N. Halama, D. Ja¨ger, J. Krause, S. H. Loosen, A. Marx, P. Boor, F. Tacke, U. P. Neumann et al., "Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer," Nature medicine, vol. 25, no. 7, pp. 1054–1056, 2019.

[5] K. A. Tran, O. Kondrashova, A. Bradley, E. D. Williams, J. V. Pearson and N. Waddell, "Deep learning in cancer diagnosis, prognosis and treatment selection," Genome Medicine, vol. 13, no. 1, p. 152, 2021.

[6] Y. Zhang, W. Chen, L. Zhang, Y. Wang and J.-Z.Wang,"Deep learning for gastric cancer classification based on histopathological images," BMC Bioinformatics, vol. 21,no. 1,p.546 ,2020.

[7] Cristescu R, Lee J, Nebozhyn M, et al., "Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes," Nat Med, vol. 21, no. 5, pp. 449-456, 2015.

[8] O. Iizuka, F. Kanavati, K. Kato, M. Rambeau, K. Arihiro and M. Tsuneki, "Deep learning models for histopathological classification of gastric and colonic epithelial tumours," Scientific Reports, vol. 10, no. 1, p. 1504, 2020.

[9] Y. Qin, Y. Deng, H. Jiang, N. Hu and B. Song, “Artificial intelligence in the imaging of gastric cancer: current applications and future direction,” Frontiers in Oncology, vol. 11, p. 631686, 2021.

[10] J.-H. Lee and J.-H. Kim, “A review of the application of deep learning in medical image analysis,” Journal of Biomedical Engineering Research, vol. 41, no. 2, pp. 77-88, 2020.

[11] S. Venkatesh, C. Meriga, M.G.V.L Geethika, T. Lakshmi Gayatri, and V.B.K.L Aruna, “Transfer Learning Based Classification of MSI and MSS Gastrointestinal Cancer,” EasyChair Preprint, Jan. 2022.

[12] C.H. Sai Venkatesh, Caleb Meriga, M.G.V.L Geethika, T. Lakshmi Gayatri, and V.B.K.L Aruna, “Modified ResNet Model for MSI and MSS Classification of Gastrointestinal Cancer.”

[13] S. Fan, X. Li, X. Cui, L. Zheng, X. Ren, W. Ma, and Z. Ye, “Computed tomography-based radiomic features could potentially predict microsatellite instability status in stage ii colorectal cancer: a preliminary study,”Academic radiology, vol. 26, no. 12, pp. 1633–1640, 2019.

[14] R. Yamashita, J. Long, T. Longacre, L. Peng, G. Berry, B. Martin, J. Higgins, D. L. Rubin and J. Shen, “Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study,” The Lancet Oncology, vol. 22, no. 1, pp. 132-141, 2021.

[15] H. Erikçi and Z. Pamuk, “VGG16 classification of microsatellite instability in colorectal cancer by deep learning,” International Referred Journal of Engineering and Sciences, 2022.

[16] F. Dedeurwaerdere, K. B. M. Claes, J. Van Dorpe, I. Rottiers, J. Van der Meulen, J. Breyne, K. Swaerts and G. Martens, “Comparison of microsatellite instability detection by immunohistochemistry and molecular techniques in colorectal and endometrial cancer,” Scientific Reports, vol. 11, no. 1, p. 12880, 2021.

[17] A. Krizhevsky, I. Sutskever and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” Communications of the ACM, vol. 60, no. 6, pp. 84-90, 2017.

[18] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.

[19] Gao, Y. and Mosalam, K.M., 2018. Deep transfer learning for image-based structural damage recognition. Computer-Aided Civil and Infrastructure Engineering, 33(9), pp.748-768.

[20] LarsenFreeman, D., 2013. Transfer of learning transformed. Language Learning, 63, pp.107-129

[21] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” arXiv preprint arXiv:1704.04861, 2017 .

[22] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” arXiv preprint arXiv:1801.04381, 2018.

[23] Qassim, H., Verma, A. and Feinzimer, D., 2018, January. Compressed residual-VGG16 CNNmodel for big data places image recognition. In 2018 IEEE 8th annual computing and communication workshop and conference (CCWC) (pp. 169-175). IEEE.

[24] Guan, S., Lei, M. and Lu, H., 2020. A steel surface defect recognition algorithm based on improved deep learning network model using feature visualization and quality evaluation. IEEE Access, 8,pp.4988549895.

[25] Wang, C., Chen, D., Hao, L., Liu, X., Zeng, Y., Chen, J. and Zhang, G., 2019. Pulmonary image classification based on inception-v3 transfer learning model. IEEE Access, 7, pp.146533-146541. [44] Yao, N., Ni, F., Wang, Z., Luo, J., Sung, W.K., Luo, C. and Li, G., 2021. L2MXception: an improved Xception network for classification of peach diseases. Plant Methods, 17(1), pp.1-13.

[26] M. Tan and Q. V. Le. EfficientNet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning, pages 6105–6114. PMLR, 2019.

[1] Sai Venkatesh, C.H., Meriga, Caleb, Geethika, M.G.V.L, Lakshmi Gayatri, T., Aruna, V.B.K.L. "Modified ResNet Model for MSI and MSS Classification of Gastrointestinal Cancer." Journal/Conference Name, Volume(Issue), Page Numbers (Year).