SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
http://saucis.sakarya.edu.tr/
e-ISSN: 2636-8129
Publisher: Sakarya University
Vol. 8, No. 1, 136-151, 2025
DOI: 10.35377/saucis...1638424
RESEARCH ARTICLE

A Comparative Analysis of Deep Learning Models for Prediction of Microsatellite Instability in Colorectal Cancer

Ziynet Pamuk[1*], Hüseyin Erikçi[1]

[1] Zonguldak Bulent Ecevit University, Department of Biomedical Engineering, Zonguldak, Türkiye, ror.org/01dvabv26

Corresponding author:
Ziynet Pamuk, Zonguldak Bulent Ecevit University, Department of Biomedical Engineering, Zonguldak, Türkiye
ziynet.pamuk@beun.edu.tr

Article History:
Received:12.02.2025
Revised: 06.03.2025
Accepted: 23.03.2025
Published Online: 27.03.2025

ABSTRACT
Colorectal cancer remains one of the most prevalent and fatal malignancies worldwide, underscoring the necessity for early and precise diagnostic approaches to enhance patient prognoses. This study proposes a deep learning-based model for predicting microsatellite instability (MSI) in colorectal cancer using hematoxylin and eosin (H&E)-stained histopathological tissue slides. A classification framework was constructed using convolutional neural networks (CNN) and optimized through transfer learning techniques. The dataset, comprising 150,000 unique H&E-stained histologic image patches, was sourced from an open-access Kaggle repository, with 80% allocated to training and 20% to testing. A comparative evaluation of nine pre-trained models demonstrated that the VGG19 architecture yielded the highest classification performance, achieving an accuracy of 90.60%, a precision of 88.60%, a sensitivity of 93.10%, and an AUC score of 90.60%. Considering its high performance, the proposed model is expected to assist pathologists in clinical decision-making, potentially enhancing diagnostic accuracy in real-world medical applications.

Keywords: Microsatellite instability, Deep learning, Colorectal cancer, Histopathologic image

1. Introduction

Colorectal cancer (CRC), the most common type of cancer worldwide, accounts for a significant proportion of cancer-related deaths [1]. According to the World Health Organization's International Agency for Research on Cancer (IARC), approximately 19.3 million individuals are diagnosed with cancer, while around ten million cancer-related deaths have been reported. With these figures, CRC represents 10% of cancer cases, and its death rate accounts for 9.4% of cancer-related mortality [2, 3]. Consequently, CRC is considered the third most diagnosed malignancy and is among the second leading causes of cancer-related deaths globally [4].

For the diagnosis of colorectal cancer, a tissue sample is typically obtained endoscopically, embedded in paraffin (FFPE), and fixed in formalin [5]. While FFPE is a standard method that aids in developing medical treatments and disease diagnosis, research on the preservation and preparation of biopsy samples has shown that FFPE tissue slides, when properly prepared, are highly durable and can be stored at ambient temperature for extended periods [6]. Subsequently, the tissue sample is stained with hematoxylin and eosin (H&E) and examined microscopically for diagnosis [5]. CRC patients are recommended to undergo advanced molecular testing, especially for microsatellite instability (MSI), which is caused by genetic alterations and affects 10-15% of these patients [7, 8]. MSI is a common tumor phenotype characterized by abnormal repetitions of short DNA motifs resulting from an insufficient mismatch repair (MMR) system, and it is associated with Lynch Syndrome (LS), an inherited cancer syndrome [9].

Systematic MSI screening is generally recommended for evaluating the response to colorectal immunotherapy [8] and has been controversially proposed for chemotherapy response. MSI diagnosis is traditionally performed using polymerase chain reaction (PCR) or next-generation sequencing (NGS) by analyzing H&E-stained CRC tissue and manually detecting abnormalities in microscopic images [10]. These methods are time-consuming and costly and require specialized expertise that may not be available at all centers, which can sometimes lead to errors [11]. Therefore, there is a growing need for tools that are widely accessible, cost-effective, and provide high-accuracy results in a shorter time.

Cite as: Z. Pamuk, H. Erikçi. "A comparative analysis of deep learning models for prediction of microsatellite instability in colorectal cancer". *Sakarya University Journal of Computer and Information Sciences*, vol. 8, no. 1, pp. 136-151, 2025, doi: 10.35377/saucis...1638424

Deep learning technologies, a subfield of artificial intelligence (AI), have become increasingly prevalent in healthcare [12]. Significant contributions to medicine have been made by developing segmentation and classification models that can diagnose and detect tumor mutations and molecular changes using convolutional neural networks (CNNs). These methods have been applied in detecting skin cancer using dermoscopic images [13] and in detecting lung [14], liver [15], and breast [16] cancers using histopathological images. Building upon these studies, this research aims to estimate the MSI status in colorectal cancer by conducting comparative analysis using deep learning models previously trained on H&E-stained histopathological slides.

The paper's organization is as follows: The studies on deep learning techniques applied to histopathological images for MSI prediction are presented in Section 2. The dataset is described in detail in Section 3.1. The deep transfer learning architecture, pre-trained models, and experimental setup parameters are explained in Sections 3.2 and 3.3, respectively. Performance metrics are provided in Section 3.4. The experimental results obtained from the proposed models and discussions are presented in Section 4. Finally, conclusions and future work are summarized in Section 5.

2. Related Works

Several studies have diagnosed MSI using H&E-stained histopathological images in cancer research, employing dual or multi-class classification based on deep learning methods. While some studies used raw data, others applied segmentation techniques through feature extraction. Additionally, the number of datasets used in these studies varies. Among these studies, convolutional neural networks and machine learning are the most commonly evaluated methods, particularly in studies incorporating transfer learning approaches. However, there is still no consensus on the ideal neural network architecture. Furthermore, variations are observed in the success rates of different models with varying parameters.

In contribution to the dataset used in our study, Kather et al. employed the ResNet18 model architecture over ImageNet to train and validate H&E-stained FFPE and frozen images from various independent cohorts, such as gastric and colorectal cancer from The Cancer Genome Atlas (TCGA). The German colorectal cancer cohort (DACHS) was used as an external confirmation set. It was observed that colorectal cancer exhibited better performance results than gastric cancer at the patient level. They demonstrated that MSI could be estimated with an area under the curve (AUC) value of 0.84 for the colorectal cancer set [17].

Cao et al. proposed a deep learning model based on multi-example learning to estimate MSI from histopathological images. They developed the Ensembled Patch Likelihood Aggregation (EPLA) model using two different colorectal cancer cohorts from TCGA-COAD and Asia and the ResNet18 architecture. Initially, it was trained and validated with TCGA-COAD, achieving an AUC of 0.8848. However, a lower AUC of 0.6497 was obtained using the EPLA model on the Asia-CRC external confirmation dataset. By utilizing previously trained networks within the model, the AUC was improved to 0.8504, and MSI estimation was successfully performed. They evaluated model performance by generating pathological signatures from the model [18].

Echle et al. developed a deep learning system based on MSI or dMMR estimation of colorectal cancer patients' H&E-stained slide images. This system was trained using the ESA-based ShuffleNet model architecture and validated externally with the Yorkshire Cancer Research Bowel Cancer Improvement Program (YCR-BCIP) cohort. The system demonstrated high performance in estimating MSI/dMMR tissue characteristics [19, 20].

Lee et al. introduced a two-phase classification method to estimate MSI in CRC. Initially, they segmented the tumor regions into MSI-L and MSI-H areas using a feature pyramid network. Then, tumor classification was performed using Inception-ResNetV2. The pathology images were magnified at 10x and 20x, and the suggested method was reported to outperform traditional methods in terms of performance [21].

Krause et al. developed synthetic images using conditional generative adversarial networks (CGAN) to estimate MSI status in CRC. Training with these synthetic images achieved an AUC of 0.742 on actual image data. When both actual and synthetic images were used for training, the AUC improved to 0.777, demonstrating an approach to augment small datasets with synthetic data [22].

Zhang et al. employed a different visualization method for MSI estimation in rectal cancer using high-resolution T2-weighted magnetic resonance imaging (MRI). Their model, developed with the MobileNetV2 neural network architecture, achieved AUC values of 0.573, 0.82, and 0.868 for the clinical, purely visual, and combined models. It was noted that the clinical model exhibited lower performance compared to the other two models [23].

Yamashita et al. developed the MSINet-100 deep learning model by scanning H&E-stained whole slide images magnified at 40x. The model was externally validated with images from TCGA and using 20x magnification. Additionally, they compared the performance of MSINet with five gastrointestinal pathologists and concluded that the deep learning model outperformed the pathologists in estimating MSI [24].

Bustos et al. developed a deep learning system that reduced multi-biases using adversarial networks from tissue microarrays to estimate MSI in colorectal cancer. This system was validated with data from 1,788 patients from EPICOLON and HGUA. Tissue types and model performance effects were evaluated using different magnifications at both

tissue and patient levels. The study marked the first attempt to estimate MSI using a multi-bias ablation technique with tissue microarrays [25].

Further studies on MSI estimation in CRC include those by Lee et al., Qui et al., Su et al., Muti et al., and Saillard et al., who applied deep learning-based classification methods using neural networks for MSI estimation in multiple cancer types, including gastric cancer and Epstein-Barr virus-related cancers [26-30].

3. Material and Methods

3.1. Dataset

In this study, histopathological tissue images of colorectal cancer (COAD) from The Cancer Genome Atlas (TCGA), specifically focusing on MSI and MSS cases, were obtained from the open-access Kaggle website and provided by Joan Gibert. This dataset comprises 192,312 unique image patches extracted from the histopathological images of cancer patients. These images were reproduced from diagnostic slides that had been fixated with formalin and embedded in paraffin (FFPE) [31]. FFPE diagnostic slides play a crucial role in medical diagnoses, as they are produced by fixating a tissue sample in formaldehyde and embedding it in a paraffin wax block for sectioning. A clear, well-defined image is created to prepare the sample for computational analysis [32].

The dataset was categorized into two groups: MSS and MSI. It contains 117,276 images in the MSS category and 75,039 in the MSI category. During the creation of the dataset, Kather performed several preprocessing steps. These steps included automatic tumor detection, resizing the images to a 0.5 µm/px resolution with dimensions of 224x224 pixels, followed by color normalization using the Macenco method [33]. The images were then classified as either "MSS" (microsatellite stable) or "MSIMUT" (microsatellite unstable or highly mutated) based on the patients' conditions [34]. As a result of these preprocessing steps, the images were of high quality and were used in their final form as JPG files in the study. Sample images are shown in Figure 1.

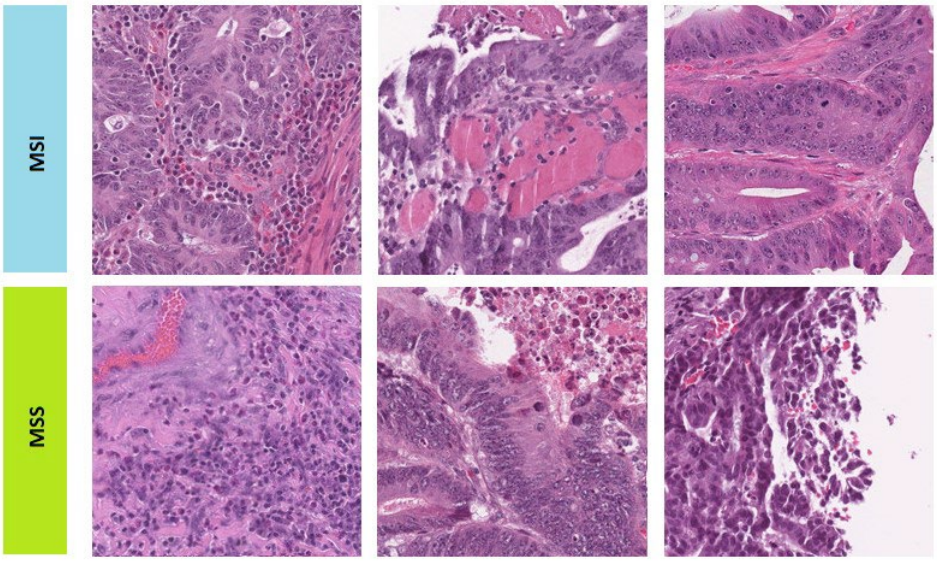

Figure 1. Histopathological Image Samples of MSI and MSS Classes

3.2. Architecture of Deep Transfer Learning

Deep learning is a subfield of machine learning that operates by mimicking the working style of the human brain, processing information through multiple layers. In recent years, as deep learning technology has been applied in areas such as autonomous systems, computer vision, and speech recognition, it has also made significant advancements in medical data analysis and medical image processing [35, 36]. In our study, we achieve substantial results by applying deep learning methods to medical imaging data.

Studies in classification, segmentation, and lesion detection are conducted using images obtained from medical imaging modalities, such as positron emission tomography-computed tomography (PET-CT), magnetic resonance imaging (MRI), computed tomography (CT), and x-rays. As a result of these analyses, the detection and classification of certain diseases have become more rapid and cost-effective [37-39].

Deep learning models rely on several algorithms, one of the most widely used being the Deep ESA algorithm in image classification. ESA is a multi-layer, feed-forward neural network that utilizes sensors to analyze data through controlled learning processes [40]. This algorithm captures the features of the image data used in the study and classifies them through various stages. As illustrated in Figure 2, feature extraction is initially performed in the convolutional layers. A filter is

applied in the feature extraction process, and a comparison is made between the newly generated image and the image label by sliding the filter over the image. The network's inconsistencies are addressed by reducing parameters and performing calculations in the pooling layer. In short, the fully connected layer completes the classification process [41]. Therefore, data passes through three distinct layers—convolutional, pooling, and fully connected—during the application of ESA. These layers are discussed in further detail below.

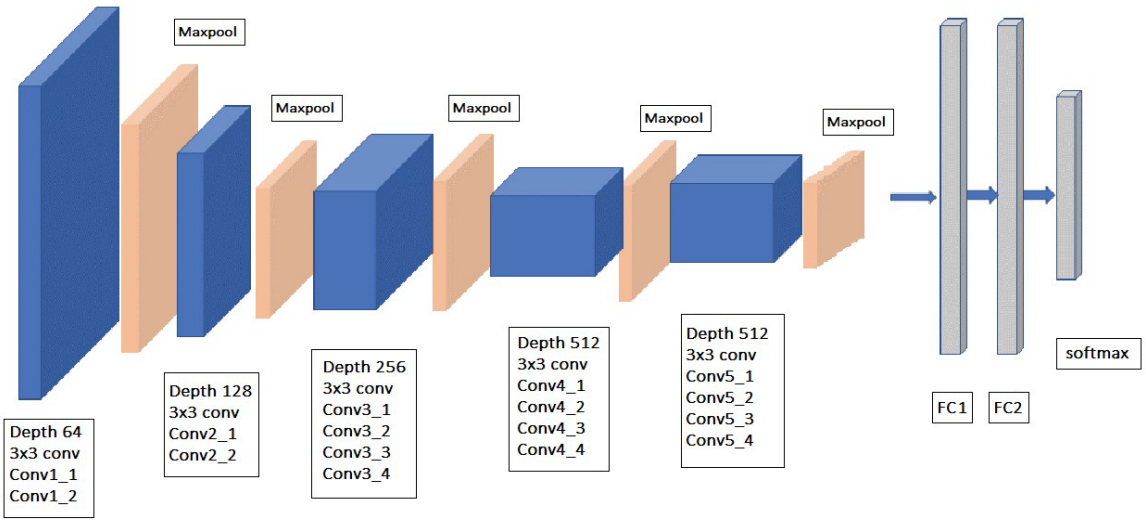

Figure 2. Architecture of the VGG19 Model

This layer is the first stage of the ESA algorithm. It is responsible for detecting the features of image data. The convolutional layer extracts both low-level and high-level features from the image by applying a filtering process. These filters assist in extracting the edge information and fine details of the image. As shown in Figure 3, the applied filters are 3x3x3 multidimensional matrices composed of pixel values. The first value of the filter represents the matrix's height, the second value represents the matrix's width, and the third value indicates the matrix's depth [42].

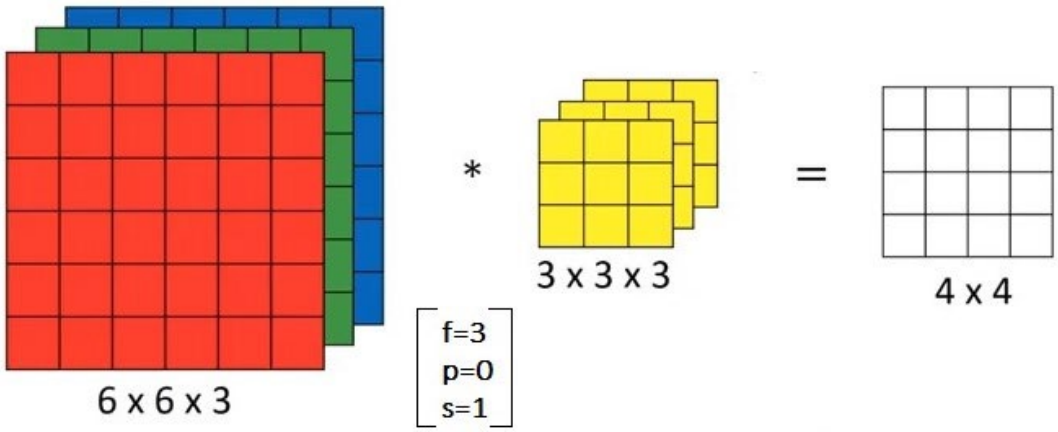

Figure 3. Calculation Process of Layers Using the Convolutional Formula (N+2p-f) +1

N: the input image size, p: padding (pixel adding), f: filter size, s: stride (scroll step).

The pooling layer is designed to reduce the number of computations and parameters within the network. Our study utilized both maximum pooling and global average pooling variants. Maximum pooling involves the application of a 2x2 filter, which selects the largest value within the area covered by the filter on the feature map, as illustrated in Figure 4. This process reduces the size of the image, thereby decreasing the computational load while preserving the essential features of the image data. On the other hand, global average pooling connects to the fully connected layer by aggregating all the neurons in the feature map into a single value, thereby reducing the dimensionality of the data [43].

This layer is the final layer of ESA and connects each neuron to every neuron in the subsequent layer. It is responsible for learning through artificial neural networks. In our study, the Rectified Linear Unit (ReLU) activation function was employed in this layer, while the sigmoid activation function was used to estimate the output data in the neurons. The mathematical formulations for these activation functions are provided in equations (1) and (2).

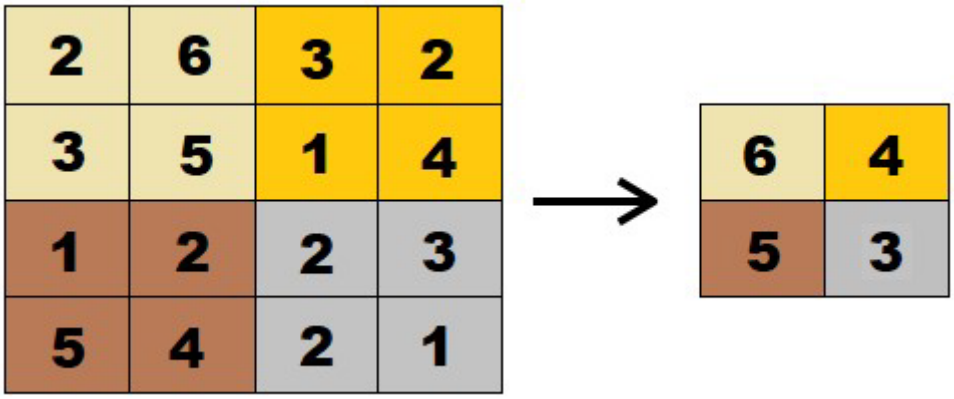

Figure 4. Illustration of the Pooling Process

$$ReLU(x) = \begin{cases} 0, if\ x < 0 \\ x, if\ x \geq 0 \end{cases} \quad (1)$$

$$Sigmoid(x) = \sigma(x) = \frac{1}{1 + \exp(-x)} \quad (2)$$

In these equations, x represents the input data. Transfer learning is a machine learning technique in which a model developed for one study is reused as the starting point for a subsequent model in another study. In other words, instead of training convolutional neural networks (CNNs) from scratch, transfer learning accelerates the learning process by leveraging information from a previously trained model, thus facilitating faster transactions and improving training efficiency. Pre-trained networks are models that inherit weights from a model that has been previously applied to a similar or even a different problem. While these networks may not always offer the optimal solution for every new problem, they help avoid the need for redundant training, enabling faster results with reduced computational effort. Pre-trained networks typically leverage the ImageNet visual object recognition database [44, 45].

In this study, deep ESA-based pre-trained networks, including VGG19, VGG16, MobileNet, MobileNetV2, ResNet50, InceptionV3, ResNet18, GoogleNet, and AlexNet, were employed to classify colorectal cancer histopathological tissue images into two categories: MSI and MSS. The performance of these networks was compared. The applied processes are illustrated in Figure 5. The study's resources are also publicly available in the open-access repository: https://github.com/hsynrkc/MSI-Prediction.

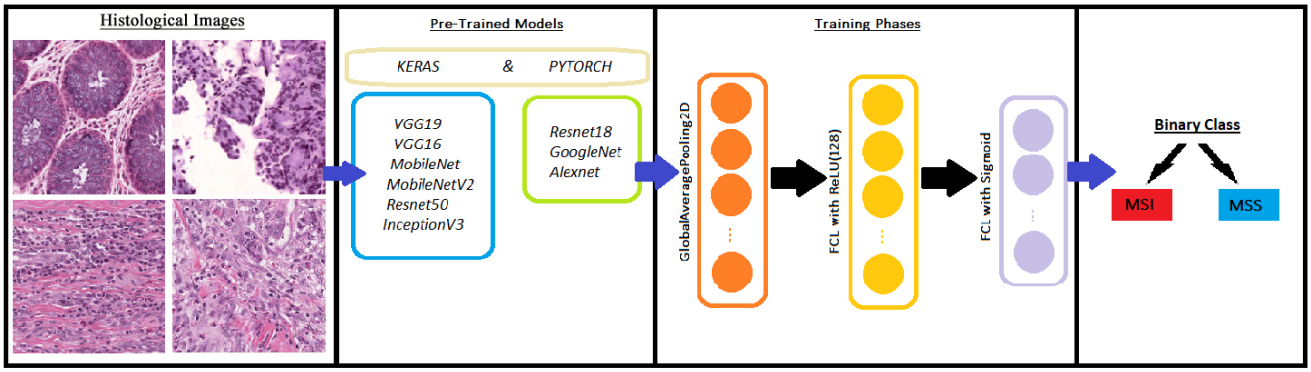

Figure 5. Schematic Flowchart Illustrating the Estimation Process of Pre-Trained Models for MSI and MSS Classification

Keras, an open-access deep learning library written in Python, is an application programming interface (API) for neural networks that operate with frameworks such as TensorFlow, Theano, and CNTK (Microsoft Cognitive Toolkit). Keras applications provide deep learning models with pre-trained weights, which can be utilized for tasks such as estimation, feature extraction, and fine-tuning. The models employed in this study are detailed below with their respective descriptions [46].

The VGG19 model is a convolutional neural network architecture with five pooling layers developed by the Visual Geometry Group (VGG) at the University of Oxford. It consists of 19 layers, 16 of which are convolutional, and three are fully connected. The network was trained using over one million images from the ImageNet database and performs

classification tasks on various images with an input size of 224x224 pixels [47]. Among the models used in this study, VGG19 achieved the best performance regarding computational cost, accessibility, high accuracy, and low loss. After the convolutional layers in the VGG19 architecture, a fully connected layer with 128 neurons is added, utilizing the sigmoid activation function.

The VGG16 model is another convolutional neural network architecture consisting of 16 layers, including 13 convolutional layers and three fully connected layers. VGG16 and VGG19 are similar, with the key difference being the number of layers. A notable feature of this network is that it includes joint layers after 2x2 or 3x3 convolutional layers, distinguishing it from other models [47].

The MobileNet model, developed by Google researchers, is a low-dimensional convolutional neural network designed to reduce computational cost and the number of parameters by employing depthwise separable convolution layers [48, 49]. MobileNetV2, an evolution of MobileNet, incorporates bottleneck layers and inverted blocks after 32 filtered convolutional layers. This variant is more efficient in dimensionality, with fewer parameters and greater depth than the original MobileNet [50].

ResNet (Residual Network) is an artificial neural network that incorporates residual connections, which involve stacking bottleneck blocks. The ResNet50 model is a convolutional neural network with a depth of 50 layers, comprising 48 convolutional layers, one max pooling layer, and one average pooling layer. It is a widely used model in deep learning applications [51].

InceptionV3, developed by Google, is another convolutional neural network trained on the ImageNet dataset. It includes multiple convolutional layers, max-pooling layers, and a fully connected layer at the final stage [52].

Table 1. Parameters Used in Deep Learning Models for MSI and MSS Classification

Parameter	Description
Model Architecture	The specific architecture of the model (e.g., VGG19, VGG16, MobileNet, etc.).
Activation Function	The activation function used in the model (e.g., ReLU, Sigmoid).
Optimizer	The optimization algorithm used (e.g., Adam, SGD).
Learning Rate	The learning rate used in the training process (e.g., 0.001, 0.0001).
Epochs	The number of times the entire dataset is passed through the network during training.
Batch Size	The number of training samples used in one forward/backward pass.
Input Image Size	The size of the input images fed into the model (e.g., 224x224).
Loss Function	The loss function used to compute the error during training (e.g., categorical crossentropy).
Training Data Size	The number of training images used in the model.
Testing Data Size	The number of images used to evaluate the model's performance during testing.
Pre-Trained Model	The pre-trained model used for transfer learning (e.g., VGG19, ResNet50, MobileNetV2).
GPU Utilization	Whether GPU is used for model training (e.g., Yes, No).
Accuracy	The accuracy rate of the model on the test set.
Precision	The precision value for the model's predictions.
Recall	The recall rate for the model's predictions.
F1-Score	The harmonic mean of precision and recall for evaluating model performance.
AUC	The Area Under the ROC Curve value, indicating model's success rate.

PyTorch is an open-access Python library used to develop deep learning models. It provides a practical environment for creating neural networks with speed and flexibility, leveraging the power of graphics processing units (GPUs). PyTorch's popularity stems from its Pythonic nature and ease of use in constructing neural network models [53]. Below are descriptions of some models employed within the PyTorch environment:

The ResNet18 model is a convolutional neural network with 18 layers, trained using the ImageNet dataset. It is a residual network similar to ResNet50 and is implemented using PyTorch [51, 54].

GoogleNet, developed by Christian Szegedy and colleagues in 2014, is a convolutional neural network architecture with 22 layers. This PyTorch implementation model includes two auxiliary classifier layers, which combine multiple inception model layers [55].

AlexNet, developed by Alex Krizhevsky and colleagues in 2012, is a neural network architecture with a depth of 8 layers, consisting of five convolutional layers and three fully connected layers. It is implemented using PyTorch, and the ImageNet dataset was trained on it [56].

As shown in Table 1, the parameters used in the deep learning models for MSI and MSS classification are summarized, highlighting each model's key components and configurations.

3.3. Experimental Setup

The Python programming language was utilized to train the proposed deep transfer learning models. All colorectal cancer histopathological image experiments were conducted at no cost using Google Colaboratory (Colab) servers, leveraging online cloud services and Tesla T4 Graphics Processing Unit (GPU) hardware. In addition, personal computers equipped with an Intel Core i3-5005U processor, 2 GB RAM, an NVIDIA GeForce 920M graphics card, and 8 GB RAM were used for supplementary tasks. Training for ESA models (VGG19, VGG16, MobileNet, MobileNetV2, ResNet50, InceptionV3, ResNet18, GoogleNet, AlexNet) was performed using the root-mean-square propagation (RMSprop) optimization technique with the cross-entropy loss function.

In the proposed model, the batch size, learning rate, and number of epochs were set to 64, 1e-5, and 10, respectively. The datasets were randomly divided into two independent subsets for training and testing, with 80% of the data used for training and 20% for testing. During model training, 80% of the data was employed, while the validation process was conducted by randomly generating a 20% validation dataset from the training data. The model's performance was evaluated using the remaining 20% of the dataset, which served as the test set.

3.4. Performance Metrics

Several performance metrics are utilized to evaluate the performance of deep learning transfer learning models. These metrics include:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = TPR = \frac{TP}{TP + FN} \tag{5}$$

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{6}$$

In these equations, TP, FP, TN, and FN represent the counts of true positive, false positive, true negative, and false negative cases, respectively. Specifically, for the test dataset and model, TP refers to the number of MSI cases correctly predicted as positive by the model, FP refers to MSS cases incorrectly predicted as positive (i.e., false positives), TN represents the number of MSS cases correctly predicted as negative, and FN refers to MSI cases incorrectly predicted as negative.

The proposed model for MSI diagnosis in colorectal cancer and other models used in this study are evaluated using classification performance metrics such as accuracy, precision, recall, and F1-score. Accuracy is the proportion of the total number of correct predictions made by the model to the total number of samples. Precision refers to the proportion of correctly predicted positive samples (MSI) out of all samples predicted as positive. Recall indicates the proportion of actual positive samples (MSI) the model correctly identified. The F1-score represents the harmonic meaning of precision and recall, balancing the two.

The ROC curve is used to visualize the model's ability to discriminate between the two classes (MSI and MSS) at various threshold values. It plots the false positive rate (FPR) on the x-axis, which is defined as 1−Specificity, and the true positive rate (TPR) on the y-axis. These are defined as:

$$Specificity = TNR = \frac{TN}{TN + FP} \quad (7)$$

$$FPR = \frac{FP}{FP + TN} \quad (8)$$

4. Experimental Results and Discussion

In this study, we conducted dual-class classification between MSI (Microsatellite Instability) and MSS (Microsatellite Stable). To evaluate the performance of various models, we compared nine different pre-trained models: ResNet18, GoogleNet, and AlexNet (from the PyTorch library) and VGG19, VGG16, MobileNet, MobileNetV2, ResNet50, and InceptionV3 (from the Keras library).

For model training, 80% of the dataset was randomly selected for the training/validation set, while the remaining 20% was reserved for the test set. This process was repeated until each model's test set was evaluated.

The detailed results of these experiments, including the performance metrics for each model, are summarized in Table 2.

Table 2. Training Accuracy and Loss Values of Pre-Trained Models Used in the Study

Models	Values						
Keras	Train			Validation		Test	
	Acc	Loss	Time	Acc	Loss	Acc	Loss
VGG19	**91**	**0.19**	**17666s**	**91**	**0.21**	**90**	**0.22**
VGG16	91	0.19	13989s	90	0.22	89	0.25
MobileNet	93	0.17	6040s	86	0.31	85	0.33
MobileNetV2	91	0.21	6653s	86	0.31	85	0.34
ResNet50	87	0.27	11355s	87	0.29	86	0.30
InceptionV3	81	0.40	7736s	81	0.39	80	0.41
Pytorch							
ResNet18	86	0.30	7166s	85	0.33	86	~
GoogleNet	82	0.38	4887s	83	0.37	83	~
AlexNet	75	0.53	3595s	77	0.52	76	~

According to the best loss value during the training period for the models applied to the dataset, the accuracy and loss values are presented in Table 1. In evaluating the model's performance, accuracy and loss rates are considered during training. The accuracy graphs for the Keras library models are shown in Figure 6.

The loss function is critical to evaluate the discrepancy between the true and predicted values. A decrease in the loss function signifies an increase in the model's robustness [58]. The VGG19 model outperforms the others, achieving lower loss values and higher detection rates for the MSI-MSS classes. While MobileNet models show better accuracy and loss values during the training phase, they perform worse than VGG19 and VGG16 in the accuracy and test phases. Additionally, although ResNet50 and InceptionV3 models yield good training results, their accuracy and test values are relatively poor. Therefore, VGG19 demonstrates the highest overall performance with an accuracy of 91%. The results summarized in Table 3 reflect the training outcomes on the histopathological images of all the pre-trained models used in this study. The accuracy values represent the general success rates of the models.

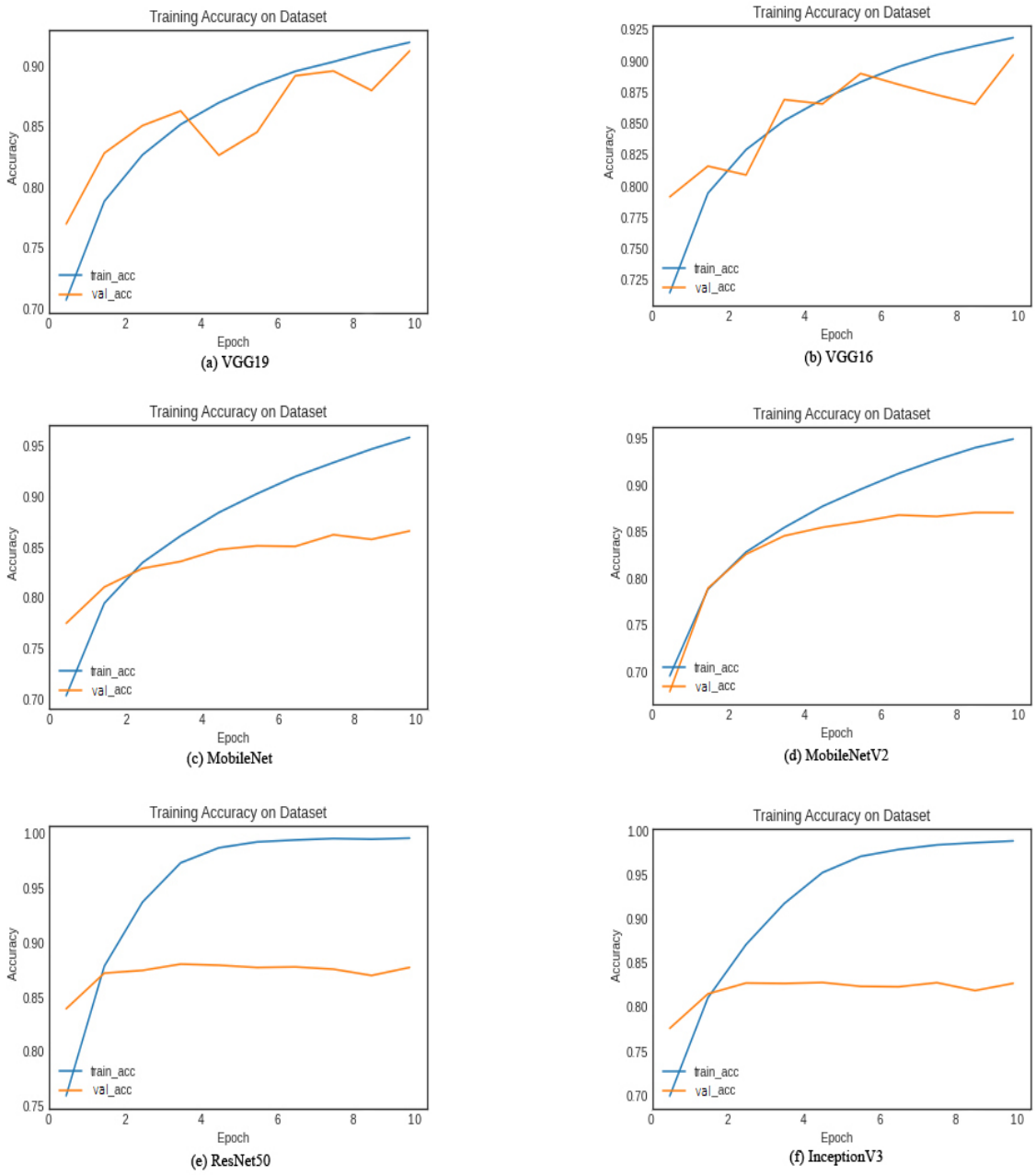

Figure 6. Training and Validation Accuracy Curves of the Pre-trained Models

Table 3. Performance Metrics of Six Different Models

Models	SPE (%)	ACC (%)	PRE (%)	REC (%)	AUC (%)	F1-Score (%)
VGG19	**88.1**	**90.6**	**88.6**	**93.1**	**90.6**	**90.8**
VGG16	93.5	89.4	92.9	85.3	89.4	88.9
ResNet50	86.7	86.6	86.6	86.4	86.6	86.5
MobileNet	83.9	85.1	84.3	86.3	85.1	85.3
MobileNetV2	86.3	85.2	86.0	84.0	85.2	85.0
InceptionV3	81.9	80.9	81.5	79.8	80.9	80.6

The VGG19 model outperforms the other models primarily due to its deeper architecture, which enables superior feature extraction from complex medical images. Its 19 layers effectively capture low-level and high-level features, improving classification accuracy. Furthermore, VGG19 benefits from transfer learning, having been pre-trained on large datasets like ImageNet, which helps it generalize well to new tasks with limited data. The model's optimized activation functions, such as ReLU, enhance its learning efficiency and stability. These factors contribute to VGG19's superior performance metrics, including accuracy, precision, recall, AUC, and F1-score.

The VGG16 model, while similar to VGG19, has a slightly shallower architecture with 16 layers, which may limit its ability to capture more intricate features in complex data. As a result, it achieved lower performance in certain metrics, particularly recall, compared to VGG19. While it still benefits from transfer learning and optimized activation functions, the reduced depth of the network restricts its ability to exploit the available data fully.

ResNet50, which incorporates residual connections to overcome the vanishing gradient problem, performs well but not as highly as VGG models. Its architecture allows for deeper networks without the typical degradation of performance seen in traditional deep networks. Yet, its overall performance is limited by factors such as the model's relatively lower ability to extract fine-grained features and its complexity compared to VGG-based models. The accuracy, precision, and recall performance are slightly lower, which could be attributed to the model's difficulty in generalizing the medical image dataset used in this study.

MobileNet and MobileNetV2 are optimized for mobile and embedded devices and designed to be lightweight and efficient, but they sacrifice some accuracy in favor of computational efficiency. While good for applications requiring lower computational cost, these models show decreased performance compared to the VGG and ResNet models, particularly in accuracy and recall. The reduced model size and fewer parameters reduce the capacity to learn detailed features from the complex data.

Finally, InceptionV3, despite its sophisticated architecture designed to capture a wide variety of features at multiple scales, demonstrates the lowest performance in this study. This could be due to its relatively high complexity, which may hinder its ability to effectively learn from smaller or specialized medical datasets, as well as potential issues with optimization for the specific task of heart disease prediction.

In conclusion, while deeper and more complex models such as VGG19 tend to perform better due to their ability to capture detailed features and benefit from transfer learning, lighter models like MobileNet and InceptionV3 show lower performance because of their design limitations and computational trade-offs.

The confusion matrix table, which is created separately for each model, is shown in detail in Figure 7. In these confusion matrices, the predictions made by the models based on the dataset split and the actual values of the images are presented. Based on the pre-trained VGG19, which provided the best performance among the models, the test dataset consists of 3000 images. Out of 15,000 MSI images, 1,321 were correctly predicted, while 1,789 were misclassified as MSS. For the MSS images, 13,967 out of 15,000 were correctly identified, while 1,033 MSI images were incorrectly classified. The confusion matrices of the other pre-trained models can be interpreted similarly.

The estimation rates presented in Table 3 are calculated separately for each model's performance metrics, with the corresponding calculation formulas provided in equations (3)-(8). The AUC value (Area Under the Curve) refers to the "area under the ROC curve," which indicates the model's success rate based on the area it covers. The ideal value for the area under the curve is 1 [59]. Figure 8 illustrates the pre-trained models' ROC curve and the area under the curve. The AUC, accuracy, F1 score, and other metrics presented in Table 3 are used as criteria for evaluating the performance of the models. According to studies by Ling et al., the AUC value is considered a more reliable metric than accuracy, as it provides a more accurate assessment of the model's performance [60].

A comparison of several studies on the estimation of MSI is presented in Table 4. Previous research has generally focused on a limited number of histopathological image datasets. In contrast, this study compares nine pre-trained models using a dataset comprising 150,000 image samples. Notably, the proposed method is fully automated, with no manual intervention, utilizing images resized to 224x224 pixels. Moreover, this study stands out by leveraging a larger dataset and incorporating more pre-trained models than many existing studies in the field. However, as the dataset size increases, the number of stages involved in the process decreases. This trade-off is mainly due to the limitations of the free cloud environment (Google Colab), which restricts the amount of processing time and capacity. Despite these constraints, the proposed system offers a cost-effective decision support tool for pathologists and doctors, aiding them in rapidly diagnosing and diagnosing MSI in colorectal cancer patients through histopathological images.

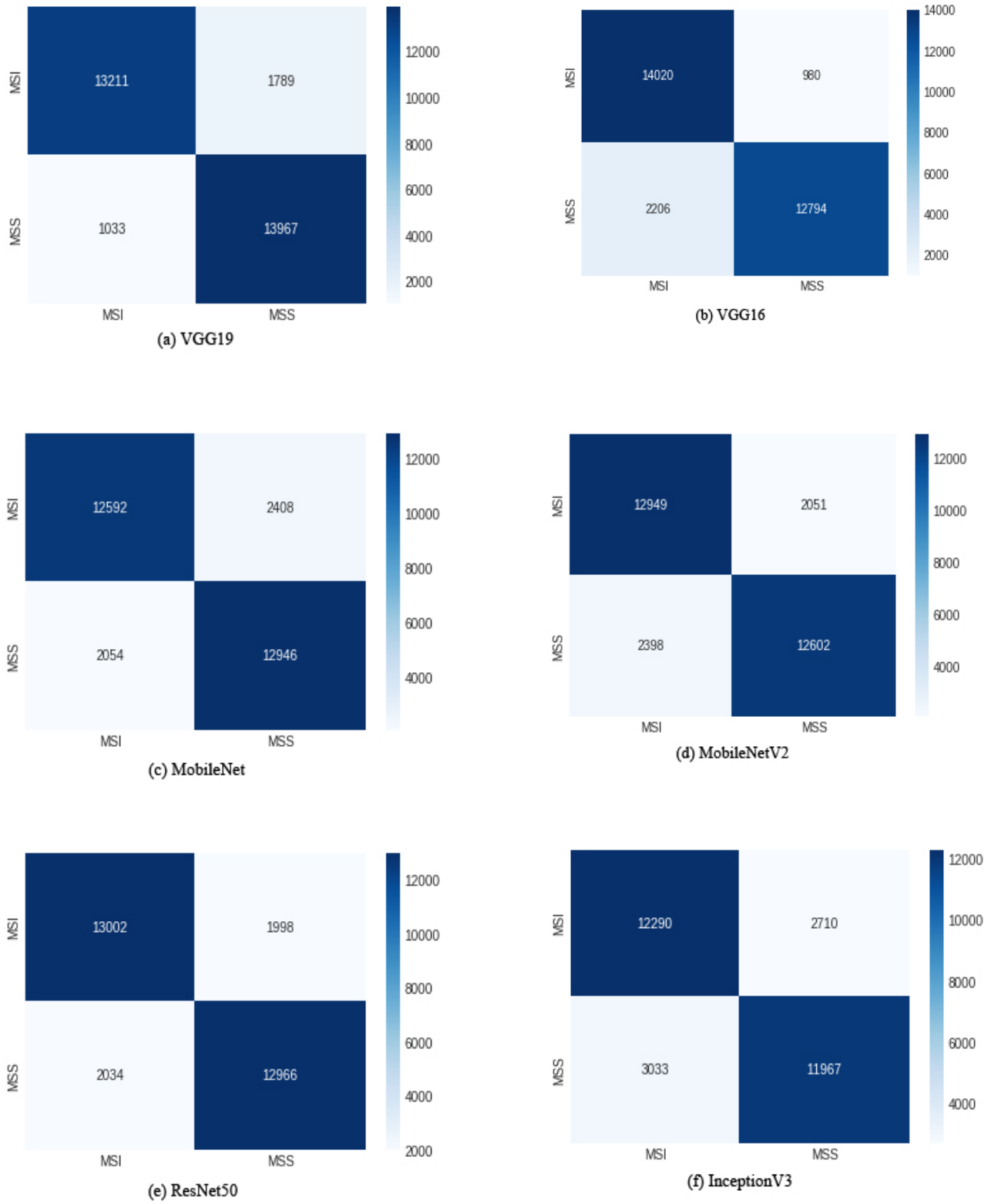

Figure 7. Confusion Matrices for the Evaluated Models

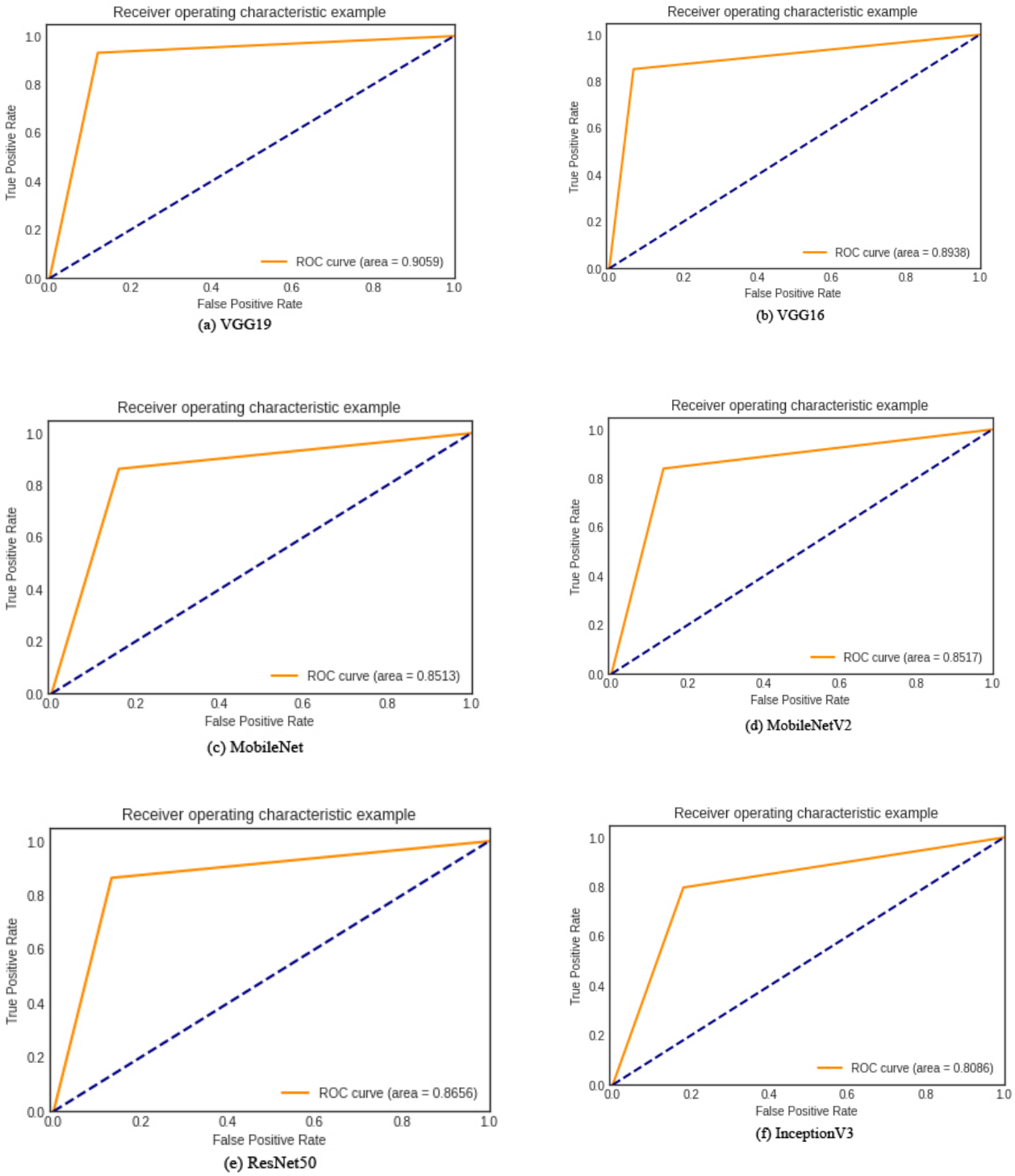

Figure 8. ROC Curves of the Pre-Trained Models

Table 4. Comparison of MSI Estimation Studies

Study	Cancer Type	Dataset	CNN Model	Performance
Kather et al. [17]	Colorectal Cancer	TCGA Cohort (738 patients)	ResNet18	AUC = 0.84
Cao et al. [18]	Colorectal Cancer	TCGA Cohort (429 WSI) Asian Cohort (785 WSI)	ResNet18	TCGA Cohort AUC = 0.885 Asian Cohort AUC = 0.850
Krause et al. [22]	Colorectal Cancer	TCGA Cohort (256 WSI) NLCS Cohort (1457 WSI) 10000 Synthetic Data	ShuffleNet	TCGA Cohort AUROC = 0.742 NLCS Cohort AUROC = 0.757 Synthetic images only AUROC = 0.743 Combination of both synthetic and real data AUROC = 0.777
Zhang et al. [23]	Rectal Cancer	West China Hospital (High-resolution T2-weighted MRI images - 491 patients)	MobileNetV2	AUC = 0.868
Bustos et al. [25]	Colorectal Cancer	Spain Hospitals EPICOLON (1705 patients) HGUA (283 patients)	ResNet34	AUC = 0.87
Qiu et al. [28]	Colorectal Cancer	TCGA Cohort (100000 H&E unique images)	ResNet34	AUC = 0.809
Su et al. [30]	Gastric cancer	Beijing Cancer Hospital (467 patients)	ResNet18	AUC = 0.785
Muti et al. [27]	Gastric cancer	Multicenter patient cohort (South Korea, Switzerland, Japan, Italy, Germany, UK, and the USA – 4128 patients)	ShuffleNet	Cross validation MSI AUROC = 0.836, EBV AUROC = 0.897. External validation MSI AUROC =0.863, EBV AUROC =0.859
This Study	Colorectal Cancer	TCGA Cohort (150000 H&E unique images)	VGG19	AUC = 0.906

5. Conclusion

The early detection of microsatellite instability (MSI) in colorectal cancer patients plays a crucial role in determining appropriate diagnostic and treatment strategies. This can help minimize side effects, reduce the time and costs associated with the disease, and ultimately improve patient outcomes. Colorectal cancer histopathological images are obtained through special staining techniques and scanning devices applied to tissue samples collected from patients, which are then digitized for analysis. This study proposes a deep transfer learning-based approach to estimate MSI from these images automatically.

Various activation functions, optimization methods, neuron counts, and epochs were explored to identify the optimal classification model for deep learning. Based on these parameters, the results of the models exhibit variations in performance. Classification accuracy is the primary metric used to evaluate the effectiveness of the models.

This study compared nine pre-trained models with identical activation and optimization methods to classify the MSI and MSS histopathological images derived from colorectal cancer patients. VGG19 demonstrated the best performance among the models evaluated, yielding accuracy, precision, recall, and F1-score values of 90.60%, 88.60%, 93.10%, and 90.80%, respectively. Models from the PyTorch library did not meet the expected performance standards. The findings suggest that

the VGG19 model, with its superior performance, holds the potential to assist pathologists in making more informed decisions in clinical settings.

Further improvements can be made in future studies, such as exploring more advanced models, optimizing computational resources, and increasing the dataset size. These enhancements are expected to lead to even more accurate results. The high-performance demonstrated in this study contributes significantly to existing literature, offering a valuable tool for researchers and clinicians working on MSI estimation and colorectal cancer diagnosis. By advancing this methodology, we can provide more efficient decision support systems, ultimately aiding in better diagnosis and treatment planning.

References

[1] K. Bardhan and K. Liu, "Epigenetics and colorectal cancer pathogenesis," *Cancers (Basel)*, vol. 5, no. 2, pp. 676-713, 2013. doi: 10.3390/cancers5020676.

[2] R. L. Siegel, K. D. Miller, A. Goding Sauer, S. A. Fedewa, L. F. Butterly, J. C. Anderson, A. Cercek, R. A. Smith, and A. Jemal, "Colorectal cancer statistics, 2020," *CA Cancer J Clin*, vol. 70, no. 3, pp. 145-164, 2020. doi: 10.3322/caac.21601.

[3] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA Cancer J Clin*, vol. 71, no. 3, pp. 209-249, 2021. doi: 10.3322/caac.21660.

[4] American Cancer Society, *Colorectal Cancer Facts & Figures 2020-2022*, 2020.

[5] T. A. A. Tosta, P. R. de Faria, L. A. Neves, and M. Z. do Nascimento, "Computational normalization of H&E-stained histological images: Progress, challenges and future potential," *Artificial Intelligence in Medicine*, vol. 95, pp. 118-132, 2019. doi: 10.1016/j.artmed.2018.10.004.

[6] J. R. Wiśniewski, "Proteomic sample preparation from formalin-fixed and paraffin-embedded tissue," *JoVE (Journal of Visualized Experiments)*, no. 79, p. e50589, 2013. doi: 10.3791/50589.

[7] R. Bonneville et al., "Landscape of microsatellite instability across 39 cancer types," *JCO Precision Oncology*, vol. 1, pp. 1-15, 2017. doi: 10.1200/PO.17.00073.

[8] J. N. Kather, N. Halama, and D. Jaeger, "Genomics and emerging biomarkers for immunotherapy of colorectal cancer," *Seminars in Cancer Biology*, pp. 189-197, 2018. doi: 10.1016/j.semcancer.2018.02.010.

[9] J. N. Nojadeh, S. B. Sharif, and E. Sakhinia, "Microsatellite instability in colorectal cancer," *EXCLI Journal*, vol. 17, p. 159, 2018. doi: 10.17179/excli2017-948.

[10] A. J. Kacew et al., "Artificial intelligence can cut costs while maintaining accuracy in colorectal cancer genotyping," *Frontiers in Oncology*, vol. 11, p. 2066, 2021. doi: 10.3389/fonc.2021.630953.

[11] K. Li, H. Luo, L. Huang, H. Luo, and X. Zhu, "Microsatellite instability: A review of what the oncologist should know," *Cancer Cell International*, vol. 20, no. 1, pp. 1-13, 2020, doi: 10.1186/s12935-019-1091-8.

[12] J. Xu, K. Xue, and K. Zhang, "Current status and future trends of clinical diagnoses via image-based deep learning," *Theranostics*, vol. 9, no. 25, p. 7556, 2019, doi: 10.7150/thno.38065.

[13] S. Hosseinzadeh Kassani and P. Hosseinzadeh Kassani, "A comparative study of deep learning architectures on melanoma detection," *Tissue and Cell*, vol. 58, pp. 76-83, 2019, doi: 10.1016/j.tice.2019.04.009.

[14] N. Coudray, P. S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, A. L. Moreira, N. Razavian, and A. Tsirigos, "Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning," *Nature Medicine*, vol. 24, no. 10, pp. 1559-1567, 2018, doi: 10.1038/s41591-018-0177-5.

[15] M. Chen, B. Zhang, W. Topatana, J. Cao, H. Zhu, S. Juengpanich, Q. Mao, H. Yu, and X. Cai, "Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning," *npj Precision Oncology*, vol. 4, no. 1, p. 14, 2020, doi: 10.1038/s41698-020-0120-3.

[16] Ş. Öztürk and B. Akdemir, "HIC-net: A deep convolutional neural network model for classification of histopathological breast images," *Computers & Electrical Engineering*, vol. 76, pp. 299-310, 2019, doi: 10.1016/j.compeleceng.2019.04.012.

[17] J. N. Kather, A. T. Pearson, N. Halama, D. Jäger, 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, doi: 10.1038/s41591-019-0462-y.

[18] R. Cao, F. Yang, S. C. Ma, L. Liu, Y. Zhao, Y. Li, D. H. Wu, T. Wang, W. J. Lu, W. J. Cai, et al., "Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in colorectal cancer," *Theranostics*, vol. 10, no. 24, pp. 11080-11091, 2020, doi: 10.7150/thno.49864.

[19] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6848-6856, 2018, doi: 10.48550/arXiv.1707.01083.

[20] A. Echle, H. I. Grabsch, P. Quirke, P. A. van den Brandt, N. P. West, G. G. A. Hutchins, L. R. Heij, X. Tan, S. D. Richman, J. Krause, et al., "Clinical-grade detection of microsatellite instability in colorectal tumors by deep learning," *Gastroenterology*, vol. 159, no. 4, pp. 1406-1416.e11, 2020, doi: 10.1053/j.gastro.2020.06.021.

[21] H. Lee, J. Seo, G. Lee, J. Park, D. Yeo, and A. Hong, "Two-stage classification method for MSI status prediction based on deep learning approach," *Applied Sciences*, 2021, doi: 10.3390/app11010254.
[22] J. Krause, H. I. Grabsch, M. Kloor, M. Jendrusch, A. Echle, R. D. Buelow, P. Boor, T. Luedde, T. J. Brinker, C. Trautwein, et al., "Deep learning detects genetic alterations in cancer histology generated by adversarial networks," *J. Pathol.*, vol. 254, no. 1, pp. 70-79, 2021, doi: 10.1002/path.5638.
[23] W. Zhang, H. Yin, Z. Huang, J. Zhao, H. Zheng, D. He, M. Li, W. Tan, S. Tian, and B. Song, "Development and validation of MRI-based deep learning models for prediction of microsatellite instability in rectal cancer," *Cancer Medicine*, vol. 10, no. 12, pp. 4164-4173, 2021, doi: 10.1002/cam4.3957.
[24] 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, doi: 10.1016/S1470-2045(20)30535-0.
[25] A. Bustos, A. Payá, A. Torrubia, R. Jover, X. Llor, X. Bessa, A. Castells, Á. Carracedo, and C. Alenda, "xDEEP-MSI: Explainable bias-rejecting microsatellite instability deep learning system in colorectal cancer," *Biomolecules*, vol. 11, no. 12, 2021, doi: 10.3390/biom11121786.
[26] S. H. Lee, I. H. Song, and H. J. Jang, "Feasibility of deep learning-based fully automated classification of microsatellite instability in tissue slides of colorectal cancer," *Int. J. Cancer*, vol. 149, no. 3, pp. 728-740, 2021, doi: 10.1002/ijc.33599.
[27] H. S. Muti et al., "Development and validation of deep learning classifiers to detect Epstein-Barr virus and microsatellite instability status in gastric cancer: A retrospective multicentre cohort study," *The Lancet Digital Health*, vol. 3, no. 10, pp. e654-e664, 2021, doi: 10.1016/S2589-7500(21)00133-3.
[28] W. Qiu, J. Yang, B. Wang, M. Yang, G. Tian, P. Wang, and J. Yang, "Evaluating microsatellite instability of colorectal cancer based on multimodal deep learning integrating histopathological and molecular data," *Frontiers in Oncology*, p. 3011, 2022, doi: 10.3389/fonc.2022.925079.
[29] C. Saillard, O. Dehaene, T. Marchand, O. Moindrot, A. Kamoun, B. Schmauch, and S. Jegou, "Self-supervised learning improves dMMR/MSI detection from histology slides across multiple cancers," *arXiv preprint arXiv:2109.05819*, 2021, doi: 10.48550/arXiv.2109.05819.
[30] F. Su, J. Li, X. Zhao, B. Wang, Y. Hu, Y. Sun, and J. Ji, "Interpretable tumor differentiation grade and microsatellite instability recognition in gastric cancer using deep learning," *Laboratory Investigation*, vol. 102, no. 6, pp. 641-649, 2022, doi: 10.1038/s41374-022-00742-6.
[31] J. Gibert, "TCGA COAD MSI vs MSS Prediction (JPG)," Kaggle, 2019. [Online]. Available: https://www.kaggle.com/datasets/joangibert/tcga_coad_msi_mss_jpg. Accessed: Dec. 24, 2024.
[32] A. Janowczyk, "DOWNLOAD TCGA DIGITAL PATHOLOGY IMAGES (FFPE)," 2018. [Online]. Available: http://www.andrewjanowczyk.com/download-tcga-digital-pathology-images-ffpe. Accessed: Dec. 24, 2024.
[33] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. G., C. Schmitt, and N. E. Thomas, "A method for normalizing histology slides for quantitative analysis," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2009, pp. 1107-1110, doi: 10.1109/ISBI.2009.5193250.
[34] J. N. Kather, "Histological images for MSI vs. MSS classification in gastrointestinal cancer, FFPE samples," Zenodo, 2019. [Online]. Available: https://doi.org/10.5281/zenodo.2530835. Accessed: Dec. 25, 2024.
[35] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015, doi: 10.1038/nature14539.
[36] D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.-Z. Yang, "Deep learning for health informatics," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 4-21, 2016, doi: 10.1109/JBHI.2016.2636665.
[37] S. J. Choi, E. S. Kim, and K. Choi, "Prediction of the histology of colorectal neoplasm in white light colonoscopic images using deep learning algorithms," *Scientific Reports*, vol. 11, no. 1, p. 5311, 2021, doi: 10.1038/s41598-021-84299-2.
[38] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Analysis and Applications*, vol. 24, no. 3, pp. 1207-1220, 2021, doi: 10.1007/s10044-021-00984-y.
[39] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240-1251, 2016, doi: 10.1109/TMI.2016.2538465.
[40] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," *arXiv preprint* arXiv:1511.08458, 2015.
[41] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.
[42] İ. Özkan and E. Ülker, "Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri," *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, vol. 6, no. 3, pp. 85-104, 2017.
[43] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual understanding of convolutional neural network-A deep learning approach," *Procedia Computer Science*, vol. 132, pp. 679-688, 2018, doi: 10.1016/j.procs.2018.05.069.

[44] J. Lu, V. Behbood, P. Hao, H. Zuo, S. Xue, and G. Zhang, “Transfer learning using computational intelligence: A survey,” *Knowledge-Based Systems*, vol. 80, pp. 14-23, 2015, doi: 10.1016/j.knosys.2015.01.010.
[45] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, and M. Bernstein, “Imagenet large scale visual recognition challenge,” *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211-252, 2015, doi: 10.1007/s11263-015-0816-y.
[46] Keras. https://tr.wikipedia.org/wiki/Keras. Accessed 10.07.2024
[47] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014. [Online]. Available: https://doi.org/10.48550/arXiv.1409.1556.
[48] S. L. Rabano, M. K. Cabatuan, E. Sybingco, E. P. Dadios, and E. J. Calilung, "Common garbage classification using mobilenet," in *2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, 2018, pp. 1-4. doi: 10.1109/HNICEM.2018.8666300.
[49] W. Wang, Y. Li, T. Zou, X. Wang, J. You, and Y. Luo, "A novel image classification approach via dense-MobileNet models," *Mobile Information Systems*, vol. 2020, 2020. [Online]. Available: https://doi.org/10.1155/2020/7602384.
[50] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510-4520. [Online]. Available: https://doi.org/10.48550/arXiv.1801.04381.
[51] K. Bardhan and K. Liu, "Epigenetics and colorectal cancer pathogenesis," *Cancers (Basel)*, vol. 5, no. 2, pp. 676-713, 2013. Doi: 10.3390/cancers5020676.
[52] R. L. Siegel, K. D. Miller, A. Goding Sauer, S. A. Fedewa, L. F. Butterly, J. C. Anderson, A. Cercek, R. A. Smith, and A. Jemal, "Colorectal cancer statistics, 2020," *CA Cancer J Clin*, vol. 70, no. 3, pp. 145-164, 2020. Doi: 10.3322/caac.21601.
[53] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA Cancer J Clin*, vol. 71, no. 3, pp. 209-249, 2021. Doi: 10.3322/caac.21660.
[54] ResNet-18-Pytorch. https://docs.openvino.ai/2021.1/omz_models_public_resnet_18_pytorch_resnet_1 8_pytorch.html. Accessed 25.08.2024
[55] T. A. A. Tosta, P. R. de Faria, L. A. Neves, and M. Z. do Nascimento, "Computational normalization of H&E-stained histological images: Progress, challenges and future potential," *Artificial Intelligence in Medicine*, vol. 95, pp. 118-132, 2019. Doi: 10.1016/j.artmed.2018.10.004.
[56] J. R. Wiśniewski, "Proteomic sample preparation from formalin fixed and paraffin embedded tissue," *JoVE (Journal of Visualized Experiments)*, no. 79, p. e50589, 2013. Doi: 10.3791/50589.
[57] R. Bonneville, M. A. Krook, E. A. Kautto, J. Miya, M. R. Wing, H-Z. Chen, J. W. Reeser, L. Yu, and S. Roychowdhury, "Landscape of microsatellite instability across 39 cancer types," *JCO Precision Oncology*, vol. 1, pp. 1-15, 2017. Doi: 10.1200/PO.17.00073.
[58] J. N. Kather, N. Halama, and D. Jaeger, "Genomics and emerging biomarkers for immunotherapy of colorectal cancer," *Seminars in Cancer Biology*, Elsevier, pp. 189-197, 2018. Doi: 10.1016/j.semcancer.2018.02.010.
[59] J. N. Nojadeh, S. B. Sharif, and E. Sakhinia, "Microsatellite instability in colorectal cancer," *EXCLI Journal*, vol. 17, p. 159, 2018. Doi: 10.17179/excli2017-948.
[60] A. J. Kacew, G. W. Strohbehn, L. Saulsberry, N. Laiteerapong, N. A. Cipriani, J. N. Kather, and A. T. Pearson, "Artificial intelligence can cut costs while maintaining accuracy in colorectal cancer genotyping," *Frontiers in Oncology*, p. 2066, 2021. Doi: 10.3389/fonc.2021.630953.

Authors Contributions
HE, obtained the study concept and dataset, conducted the experiments, and analysed the results. ZP wrote the draft and critically reviewed the draft. All authors reviewed the manuscript.

Conflict of Interest Notice
Authors declare that they have no conflict of interest.

Ethical Approval and Informed Consent
This article contains no data or other information from studies or experiments involving human or animal subjects.

Availability of data and material
The data used in the study were obtained from the Kaggle data repository.
https://www.kaggle.com/datasets/joangibert/tcga_coad_msi_mss_jpg

Plagiarism Statement
This article has been scanned by iThenticate ™.