Ağırlıklı CNN Topluluğu Tabanlı Kolorektal Kanser Tespiti

Mustafa Yurdakul*[1] ,Şakir Taşdemir[2],

*[1] Kırıkkale Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi Bilgisayar Mühendisliği, KIRIKKALE
[2] Selçuk Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği, KONYA

(Alınış / Received: 16.04.2024, Kabul / Accepted: 16.07.20**, Online Yayınlanma / Published Online: 30.08.2024)

Anahtar Kelimeler
Ağırlıklı Topluluk Öğrenmesi
Derin Öğrenme
Kolorektal Kanser
Konvolüsyonel Sinir Ağı

Öz: Kolorektal Kanser(KKR), dünya çapında yaygın ve potansiyel olarak ölümcül bir hastalıktır. Erken ve doğru teşhis, teşhisin zaman alması, insan hatalarının olasılığı ve uzman doktor eksikliği nedeniyle zor bir süreçtir. Bu çalışmada, tıbbi görüntülerden kolorektal kanserin teşhisini basitleştirmek ve hızlandırmak için derin öğrenme algoritmaları kullanılmıştır. Çeşitli KKR evrelerini içeren Enteroskop Biyopsi Histopatolojik H&E(EBHI) görüntü veri seti kullanıldı. Çeşitli önceden eğitilmiş Evrişimli Sinir Ağı modelleri, görüntüleri iyi huylu veya kötü huylu olarak sınıflandırmak için kullanıldı. Ayrıca, sınıflandırma doğruluğunu artırmak için üç en iyi model ağırlıklı topluluk yöntemiyle birleştirildi. Deneysel sonuçlar, ağırlıklı topluluk yönteminin sınıflandırma performansını önemli ölçüde iyileştirdiğini göstermektedir.

Colorectal Cancer Detection Based on Weighted Ensemble of CNNs

Keywords
Colorectal Cancer
Convolutional Neuralit Network
Deep Learning
Weighted Ensemble,

Abstract: Colorectal Cancer(CRC) is a common and potentially deadly disease around the world. Early and correct diagnosis can be a challenge due to the time takes, the possibility of human mistakes, and the lack of specialized doctors. In this study, deep learning algorithms, a type of machine learning algorithms, were used to simplify and speed up the diagnosis of CRC from medical images. In this study, we used Enteroscope Biopsy Histopathological H&E Image dataset, which contains various stages of CRC. Various pre-trained Convolutional Neural Network models were used to classify images into two groups: malignant or benign. Moreover three best models combined using weighted ensemble method to improve the accuracy of classification. The experimental results show that weighted ensemble method significantly improves classification performance.

*İlgili Yazar: myyurdakul1@gmail.com

1. Introduction

Colorectal Cancer (CRC) occurs as a result of the development of malignant tumors in the large intestine. CRC ranks third in global cancer prevalence and is the second major cause of cancer-related deaths[1]. The vast majority of cases are detected in western countries, and the number of cases is increasing every year[2]. In these cases, one out of every three patients loses their life[3]. All cancer types related to rectal, colorectal and colon cancer are accepted as colorectal cancer[4]. Histopathological imaging of the intestine is one of the methods used in the diagnosis of CRC[3]. In diagnosing of CRC, pathologists analyze digital images stained with Hematoxylin and Eosin (H&E)[5]. Pathologists locate the area of the lesion and then use microscopes to magnify and examine it in detail. However, diagnosing based on histopathological images can be challenging. Pathologists have to examine a large amount of data, which causes a time-consuming process. Pathologists can make mistakes due to fatigue, lack of attention or knowledge. Also, different pathologists may have different opinions and experiences, which can lead to inconsistent diagnoses. Furthermore, the number of pathologists globally is insufficient [6] the training period for an expert pathologist exceeds a decade[7]. For all these reasons, there is a need for fast, accurate and reliable solutions for the diagnosis of CRC. It is possible to make detailed analyzes on images with computer aided systems. Deep learning algorithms has a successful outputs in the detection of breast[8], skin[9], brain[10] and gastrointestinal[11] cancers. Based on these studies, using deep learning for CRC detection can be an effective approach. In this study, the publicly available Enteroscope Biopsy Histopathological H&E Image dataset (EBHI) is used. EBHI contains images obtained with an electron microscope, pertaining to various stages of CRC.

Using microscopic images, a classification process has been performed to categorize them into two groups: malignant and benign. Various pretrained CNN models was used to classify. The performance of the models was optimized using transfer learning and fine-tuning techniques. In addition to all these, the best three models selected according to the performances of the models were used with the weighted ensemble method to increase the performance. Proposed approach combines the strengths of the models and balances their weaknesses, thus providing a more accurate and reliable classification for CRC diagnosis.

The remaining sections of the article are structured as follows: Section 2 provides a discussion on related studies. Section 3 describes the materials and methods used, including the dataset and the proposed classification system. Section 4 presents the results and analysis of the study. Section 5 discusses the findings and implications of the research. Section 6 concludes the article, summarizing the key findings.

2. Related Works

In recent years, deep learning methods have been extensively used in the medical field[9, 12, 13]. In particular, Convolutional Neural Networks (CNN), a deep learning algorithm, contributes to the disease diagnosis process by automating medical image analysis. Successful results have been achieved in numerous studies, such as breast[8], skin[9], brain[10], gastrointestinal[11] cancers and pneumonia detection [14]. Additionally, numerous studies have been achieved at the detection of Colorectal Cancer.

Akbari et al. (2018)[15] proposed a novel CNN model for the detection of polyps, a leading cause of colon cancer. In the proposed model, the size of model was reduced by using binary weights and kernels, thereby enhancing the model's compatibility with medical devices. An accuracy value of 90.28% was achieved on the Asu Mayo Test clinical dataset.

Ponzio et al. (2018)[16] proposed the VGG16 model to distinguish Adenocarcinomas from healthy tissue and benign lesions. An accuracy rate of 90% was achieved with the VGG16 model. Additionally, transfer learning was implemented using the VGG16 model trained on ImageNet data and accuracy rate of 96% was achieved.

Poudel et al. (2020)[17] developed a new robust model for the detection of colorectal diseases with an efficient dilation method in CNNs. In the study that used the KVASIR dataset and classified five different categories (Adenocarcinoma, Adenoma, Crohn's, Ulcerative colitis, Normal), an accuracy of 95.7% was achieved.

Sarwinda et al. (2021) [18] proposed ResNet architectures to classify colon gland images into benign and malignant categories. The images were converted to grayscale and then applied the CLAHE algorithm. Experimental studies showed that ResNet-18 and ResNet-50 models resulted in an accuracy of 88% and 85%, respectively.

Su et al. (2022) [19] performed feature extraction on histopathology images for the classification of colon and colorectal diseases. In the study, various machine learning algorithms (XGBoost, SVM, RF, LDA, MLP, and LightGBM) were used to classify the extracted features from histopathology images. XGBoost achieved the highest accuracy rate of 99%.

Naga et al. (2023) [20] proposed a study for the classification of colon and lung adenocarcinomas along with squamous cell carcinomas using digital histopathology images.

In the study, feature extraction was performed on histopathology images using the PCANet model. The extracted features were optimized with the Rider Optimization algorithm and classified using the Extreme Learning Machine. They achieved an accuracy rate of 99.72% on Kvasir dataset.

Kumar et al. (2023) [21] proposed a new CNN model called CRCCN-Net for the automatic diagnosis of colorectal cancer. The model is designed to be lightweight and less complex. The CRCCN-Net model has significantly fewer parameters (3.76M) compared to Xception, InceptionResNetV2, DenseNet121, and VGG16 models. It achieved an accuracy of 93.50% on the CRCCN-Net dataset and 96.26% on the NCT-CRC-HE-100K dataset for colorectal cancer. The model outperformed Xception, InceptionResNetV2, DenseNet121, and VGG16 models, demonstrating better results.

Hu et al. (2023) [22] used various machine learning algorithms and deep learning techniques for the classification of histopathological images into benign and malignant categories using the publicly available dataset called EBHI. The classical machine learning algorithms achieved a maximum accuracy rate of 76.02%, deep learning technique was achieved a significantly higher accuracy rate of 95.37%.

Yengeç et al.[23] proposed a Clinical Decision Support System (CDSS) for the detection of adenomatous polyps in colon histopathology images. They employed the ConvNeXt architecture for the classification of histopathological images. Evaluating the proposed method on a custom dataset with over 10,000 colon histopathology images, they achieved a classification accuracy of 95%. Furthermore, on the EBHI and UniToPatho datasets, they achieved classification accuracies of 91.1% and 90%, respectively.

Although there are many studies in the existing literature, computer-aided colourectal cancer detection system is an active field of study. Although EBHI dataset to be used in the study is quite comprehensive and useful, it has been observed that a limited number of studies have been carried out. In this context, in the light of the literature, two different approaches are tested on the EBHI dataset in this study. In the first approach, various well-known CNN models are tested on the EBHI dataset. In the second approach, the three most successful CNN models are used with the weighted voting ensemble method weighted by the grid search algorithm.

3. Material And Methods

The images of Normal, Polyp, Low-Grade IN, High-Grade IN, and Adenocarcinoma in the EBHI dataset were categorized into benign and malignant classes. Subsequently, the data was divided into three parts with an 80% training, 10% testing, and 10% validation split, and online data augmentation techniques were applied. The data was trained using various architectures of different sizes and transfer learning with models pre-trained on ImageNet. The best three models were combined using weighted ensemble method. The details of the dataset, architectures used, and data augmentation processes are described following sections of the study. Fig. 1 shows the schematic diagram of the proposed study.

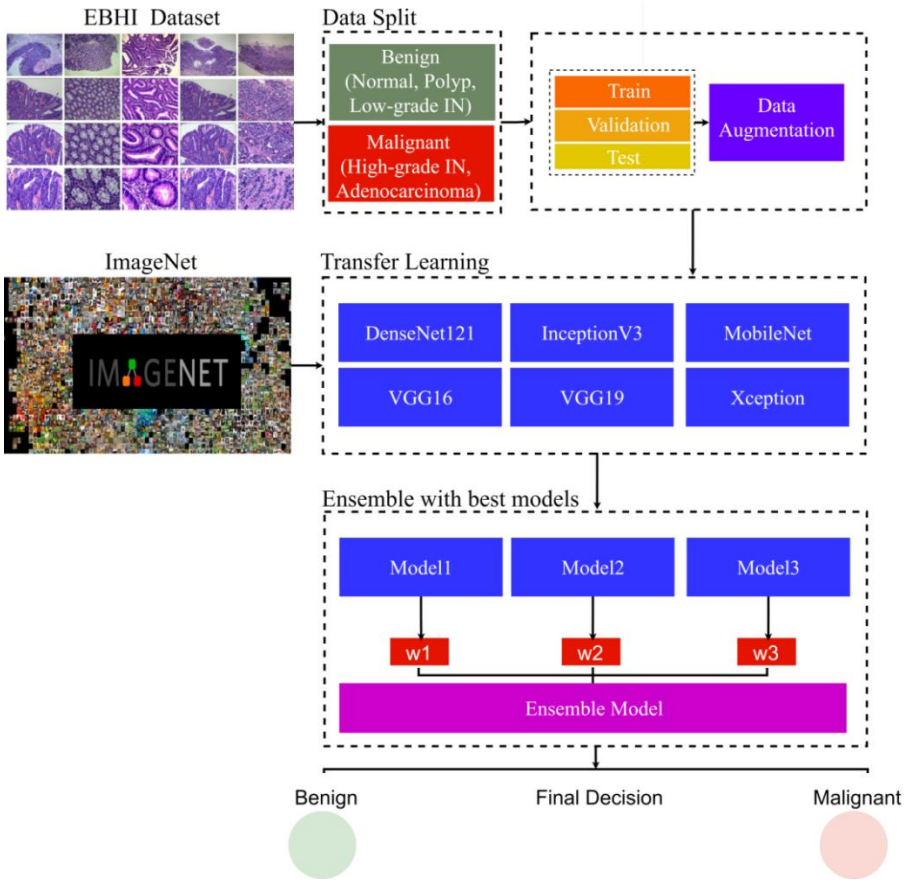

Fig. 1. Schematic Diagram of Proposed Study

4. Dataset

EBHI [22]dataset is a public dataset comprising a total of 5532 electron microscopy images, which includes five stages of colorectal cancer (CRC): Normal, Polyp, Low-Grade Intraepithelial Neoplasia (Low-Grade IN), High-Grade Intraepithelial Neoplasia (High-Grade IN), and Adenocarcinoma. Normal cells in the dataset are characterized by regular morphology and a low mitotic rate. Polyp exhibits elevated and low nuclear division rates, forming raised lesions in the intestinal mucosa. Low-Grade IN is characterized by significant changes in cell morphology and arrangement, representing precursor lesions. High-Grade IN is distinguished by severe nuclear atypia and a high mitotic rate. Adenocarcinoma represents malignant tumors with irregular lumen distribution, infiltrative growth, and a high nucleoplasmic ratio. The dataset includes four magnification levels: 40x, 100x, 200x, and 400x. The distribution of the dataset according to CRC stages and magnification levels is shown in Table 1.

Table 1. Dataset Distribution by CRC Stages and Magnification Levels[22]

Magnification	40x	100x	200x	400x	Total
Normal	17	29	61	79	186
Polyp	119	165	254	304	842
Low-grade IN	204	341	603	660	1808
High-grade IN	47	80	130	161	418

Adenocarcinoma	205	471	790	812	2278
Total	592	1086	1838	2016	5532

Fig. 2 shows sample images from different magnification levels of CRC stages in the EBHI dataset.

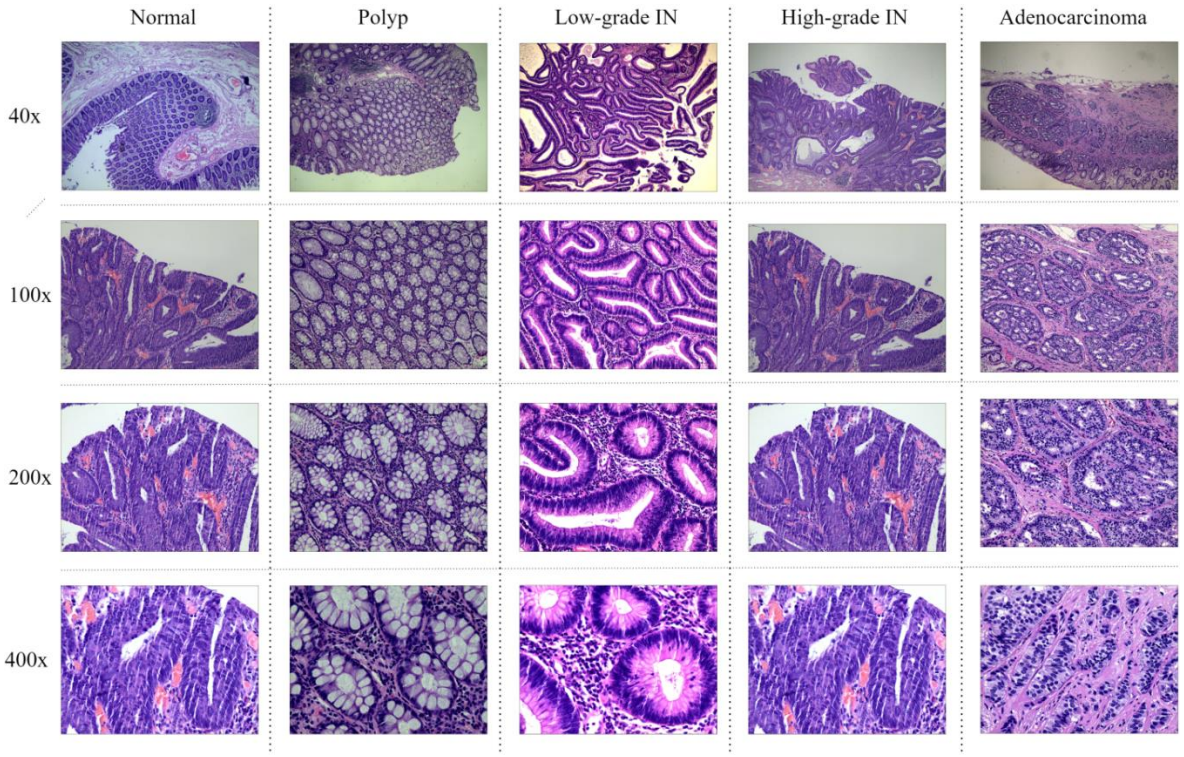

Fig. 2. Sample Images of CRC Stages at Different Magnification Levels and categories in the EBHI Dataset

The five classes corresponding to CRC stages are divided into two main categories according to the Medical Classification method: Benign (Normal, Polyp, and Low-Grade IN) and Malignant (High-Grade IN and Adenocarcinoma). The numerical distribution of the dataset based on the main categories and magnification levels is presented in Table 2.

Table 2. Dataset Distribution by Type And Magnification Levels

Type	**40x**	**100x**	**200x**	**400x**	**Total**
Benign	340	535	918	1043	2836
Malignant	252	551	920	973	2696
Total	592	1086	1838	2016	5532

In this study, classification was performed on 200x and 400x magnification levels, taking into account the distribution of the dataset. Figure 3 presents the t-SNE plot of malignant and benign images at 200x and 400x magnification levels. t-SNE [24] is a visualization method used to represent high-dimensional data by mapping similar data points closer together, making it easier to interpret the data visually.

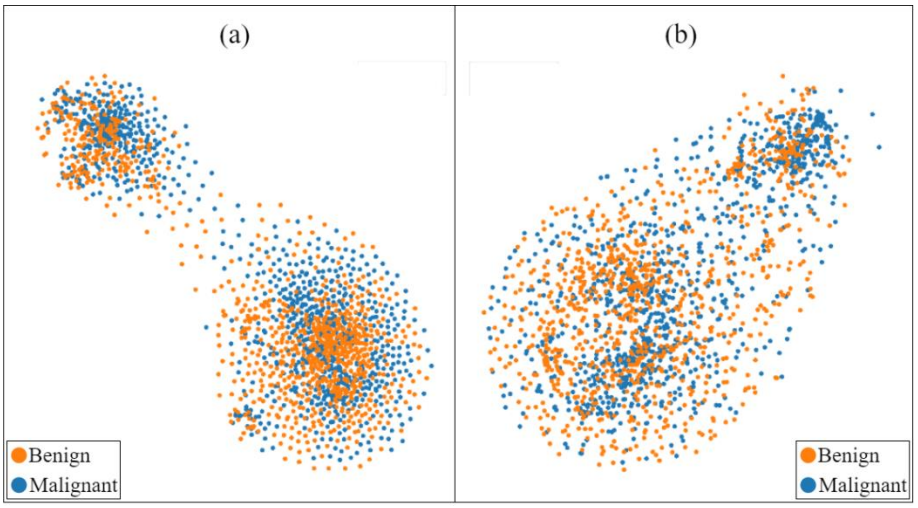

Fig. 3. Dataset visualization in a two dimensional space of benign and malignant images: (a) 200x and (b) 400x

5. Data Augmentation

The performance of deep learning algorithms is directly related to the size of the training data [25]. In cases where the data is insufficient, data augmentation techniques can be applied to increase the number of images in the dataset. In this study, online data augmentation techniques were employed. The parameters of data augmentation methods used in the study are shown in Table 3.

Table 3. Data Augmentation Parameters

Rotation	Width Shift	Height Shift	Shear	Zoom
0.8	0.3	0.6	0.5	0.2

6. Convolutional Neural Networks

CNNs are a special deep learning algorithms that used for image analysis tasks. CNNs achieved succesful outputs in various fields such as farm[25, 26], medical[27, 28] and education[29]. CNNs consist of three main components: convolutional, pooling, and fully connected layers [30-32]. The convolutional layer applies filters to the input image in order to extract relevant features[33]. The pooling layer is used for dimensionality reduction, to make the model run faster and reduce computational complexity. In the fully connected layer, the classification process is performed based on the neural network structure. Various architectures have been proposed using the three main components.

Within the study, several well-known and high-performing CNN models were used; DenseNet121 [34], InceptionV3 [35], MobileNet[36], VGG16[37], VGG19 [37], and Xception [38].

DenseNet121

DenseNet [34]was proposed by Huang et al. in 2016. The DenseNet model utilizes features from previous layers by establishing direct connections between each layer and all preceding layers. The DenseNet-121 architecture, consists of 4 dense blocks, 3 transition layers, and a total of 121 layers.

InceptionV3

InceptionV3 [35], is a network with multiple Inception modules. These modules enhance the network's capacity for generalization and adaptability to different scales.

MobileNet

MobileNet [36], is a 2017 model designed for mobile devices. It reduces network parameters and complexity by using depthwise separable convolutions.

Visual Geometry Group(VGG)

VGG [37], was developed at Oxford University in 2015. It has a deeper and more homogeneous structure than other popular CNN architectures. The VGG16 and VGG19 models consist of 16 and 19 convolutional layers, respectively, with filter sizes typically set as 3x3.

Xception

Xception[38] proposed by Chollet in 2017, is inspired by the Inception V3 architecture. The key innovation introduced in Xception is the inverse usage of depthwise separable convolution modules that modification resulted in improved performance compared to Inception V3 on the ImageNet dataset. The Xception model consists of a total of 14 modules and 36 convolutional layers.

In this study, various CNN models of different sizes, including DenseNet121, InceptionV3, MobileNet, VGG16, VGG19, and Xception, were used for image classification tasks. These models are pre-trained on the ImageNet dataset, enabling them to extract more general features. The classifier block at the last layer of models was removed, and a new classifier layer was added, facilitating transfer learning. The CNN models were trained with a batch size of 32 and an SGD optimizer for 100 epochs.

7. Ensemble Learning

The ensemble method is an approach that aims to create a more powerful and robust model by strategically combining base models[39]. In this study, we have employed the weighted ensemble method. The weighted ensemble method combines different models, with each model given a weight according to its contribution[27]. This method enhances the prominence of better models and reduces the influence of those with poor performance. Pretrained models (DenseNet121, InceptionV3, MobileNet, VGG16, VGG19, and Xception) were trained on the training dataset. Each model's performance was evaluated with the test dataset. Regarding their performance, the best three models were combined using the weighted ensemble

method. The weights of the models, ensuring a total sum of one, were determined through the grid search technique.

8. Evalation metrics

Classification performance consists of a series of metrics used to measure how accurately a model performs the classification task. In binary classification, TP (True Positive) denotes the count of correctly classified positive samples, TN (True Negative) represents the count of correctly classified negative samples, FP (False Positive) signifies the count of falsely classified positive samples, and FN (False Negative) indicates the count of falsely classified negative samples[40].

Based on TP, TN, FP, and FN values, several metrics such as accuracy, precision, recall (sensitivity), F1 score, and AUC can be calculated to assess the classification performance[40].

The mathematical formulas for these performance evaluation metrics are provided in Equations 1-4.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

9. Results

In this study, binary classification(benign and malignant) was performed on images from the EBHI dataset with 200x and 400x magnification ratios. . Pretrained models (DenseNet121, InceptionV3, MobileNet, VGG16, VGG19, and Xception) were trained on the training dataset. The graphs illustrating the accuracy and loss values during the training process, based on the training and validation data, are shown in Figure 4 and Figure 5 for the 200x magnification ratio, and in Figure 6 and Figure 7 for the 400x magnification ratio. When Figure 4,5,6 and 7 are analyzed, there is a significant difference between the train accuracy and validation accuracy at 200x and 400x magnification, which shows that the models are not overfitting. However, VGG19 experienced fluctuations in validation accuracy and loss value, indicating that the model's prediction on unseen data is worse than the other models.

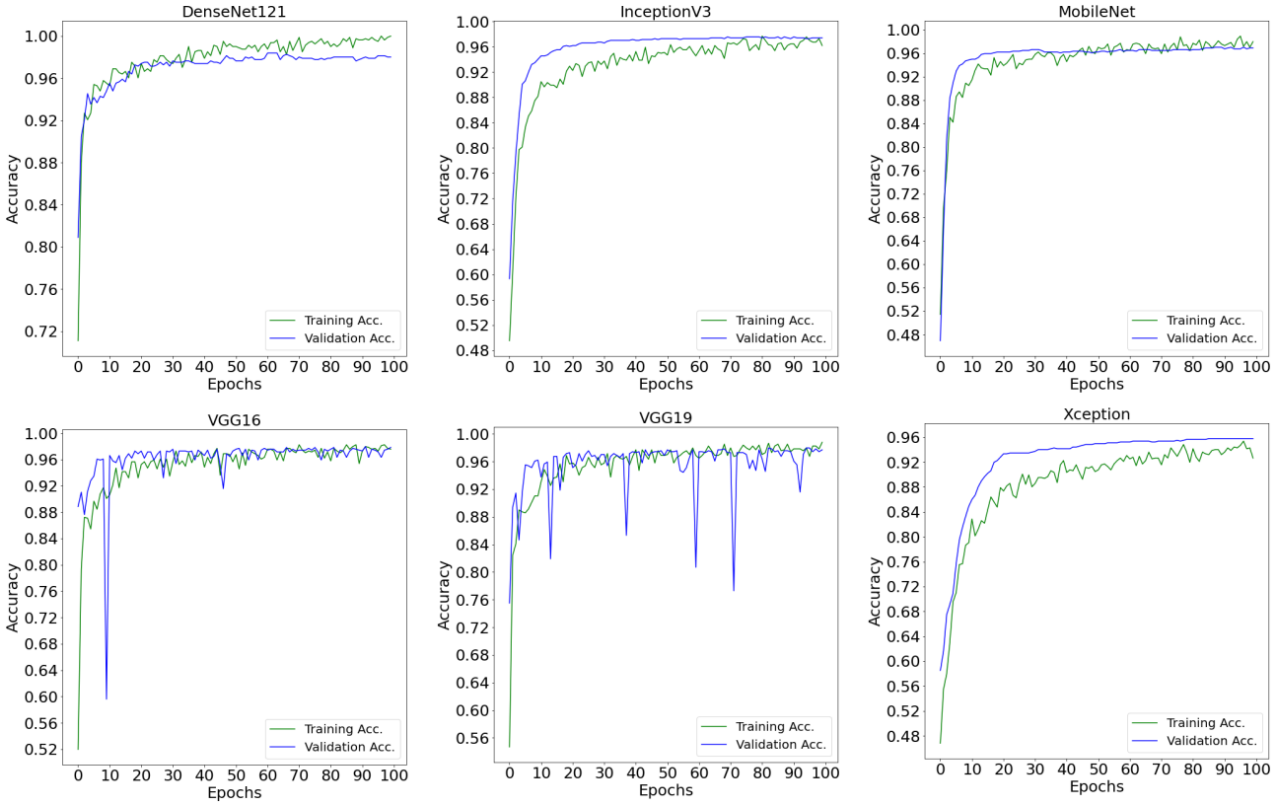

Fig. 4. Mini-batch accuracy plots for models at 200x magnification

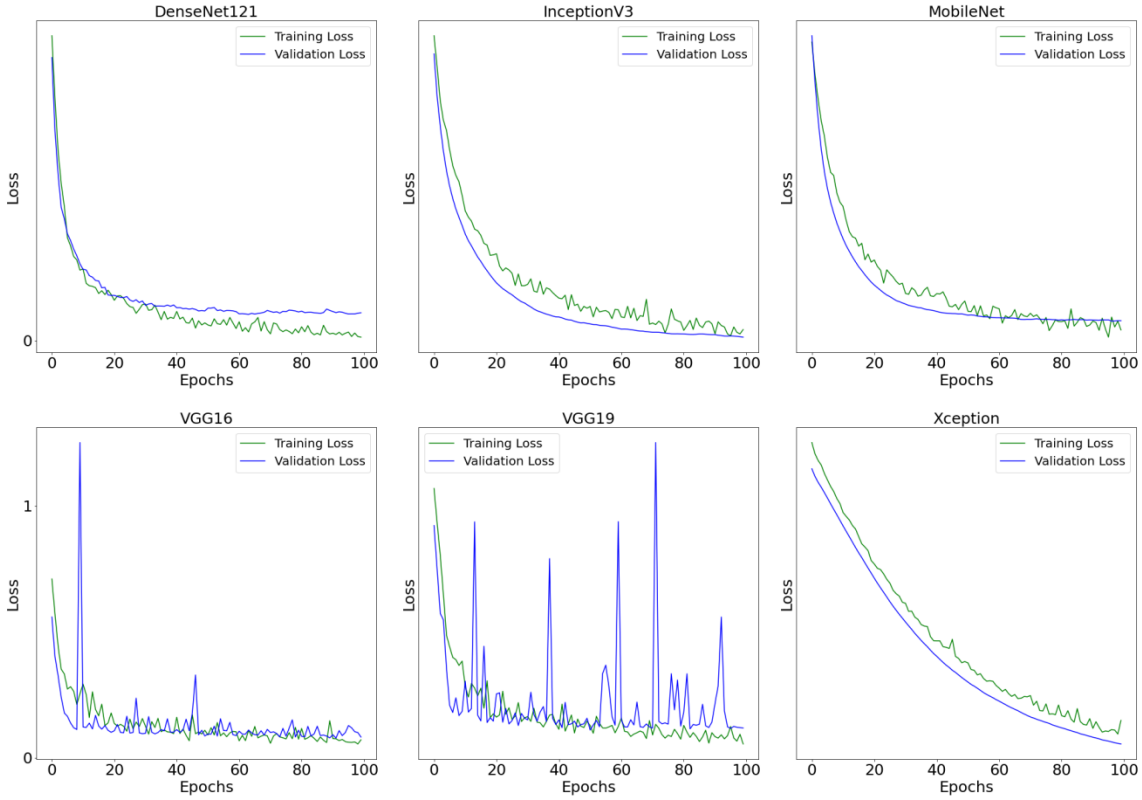

Fig. 5. Mini-batch loss plots for models at 200x magnification

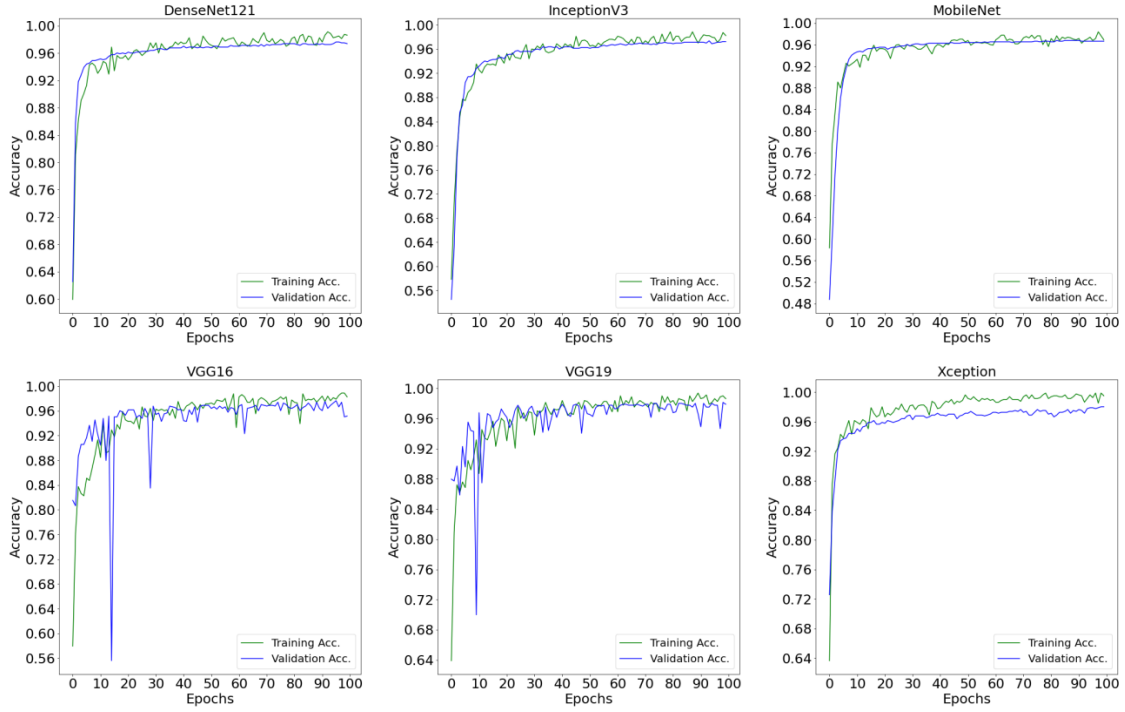

Fig. 6. Mini-batch accuracy plots for models at 400x magnification

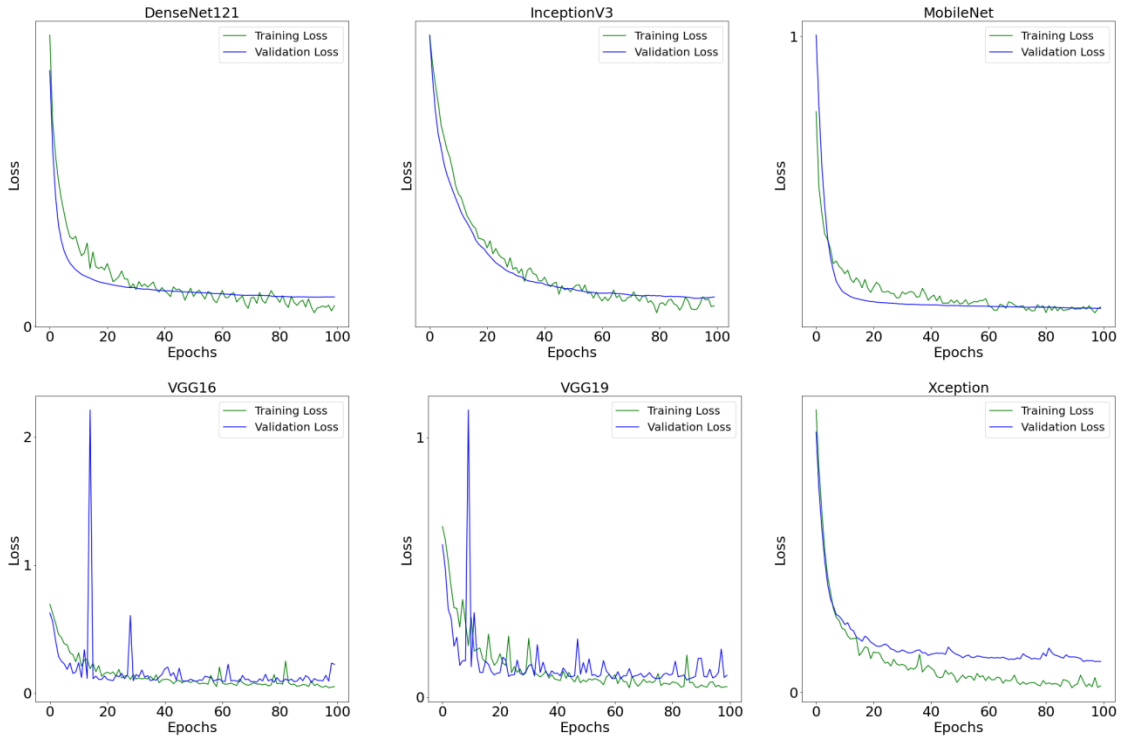

Fig. 7. Mini-batch loss plots for models at 400x magnification

According to the graphs, it can be seen that the difference in accuracy and loss values between the training and validation data is minimal. So the models did not overfit the training dataset. Performance of each model was calculated using a confusion matrix. The confusion matrices for the models, categorized by magnification ratios, are shown in Figure 8 and Figure 9.

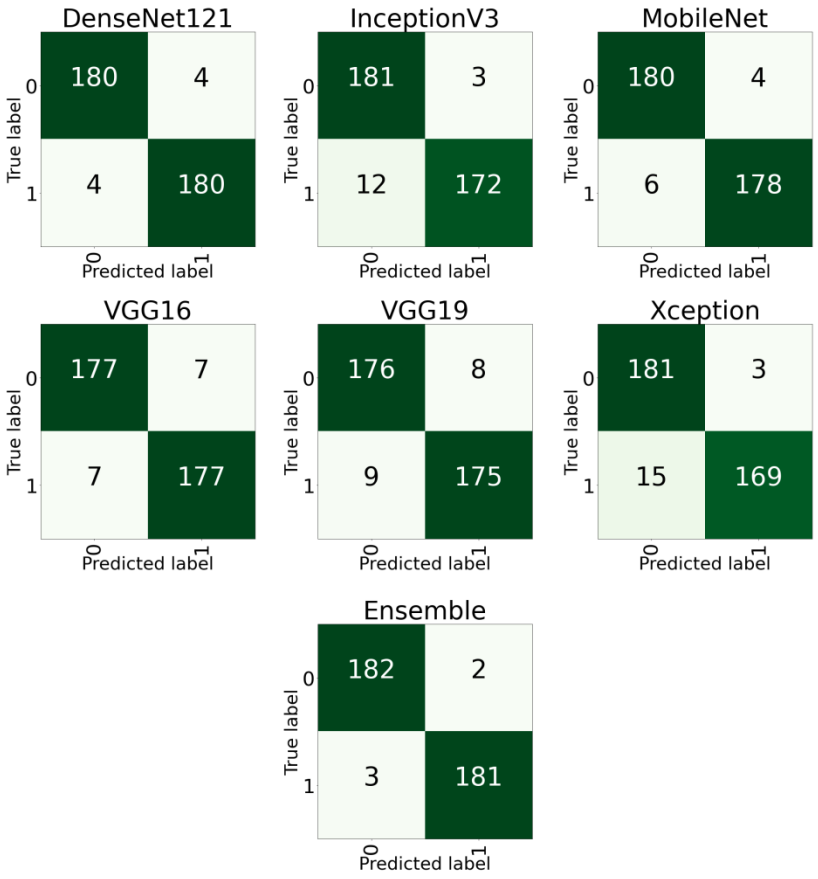

Fig. 8. Model Confusion Matrices at 200x Magnification(0:Malignant,1:Benign)

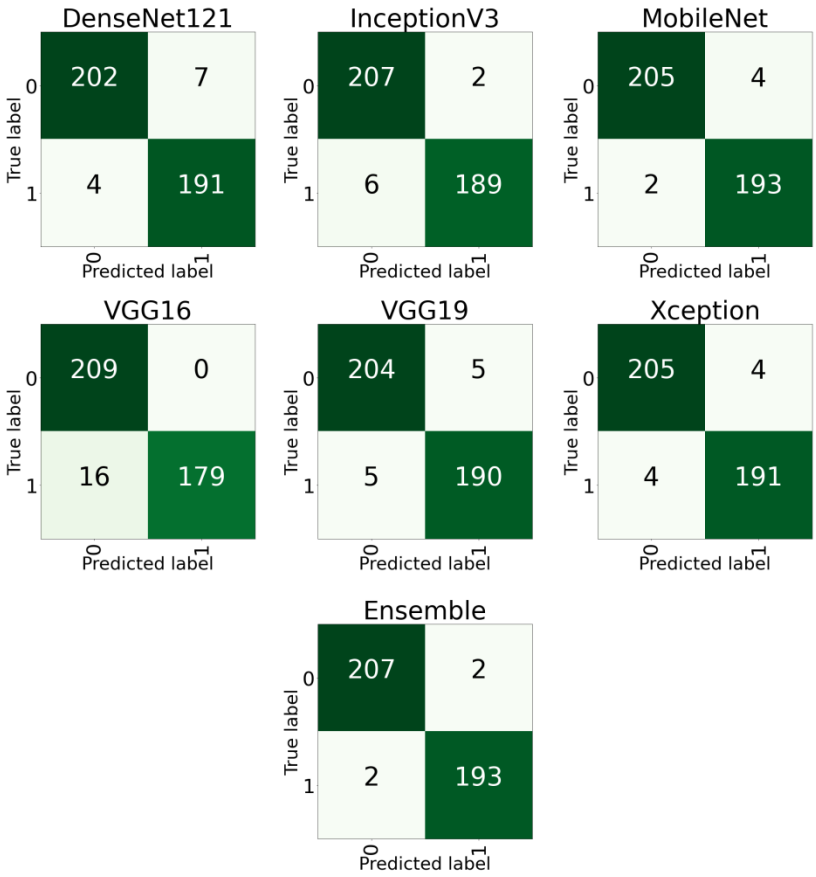

Fig. 9. Model Confusion Matrices at 400x Magnification

The performance of CNN models and the ensemble method is presented in Table 4 and Table 5, respectively, for the 200x and 400x magnification ratios. Among the models at the 200x magnification ratio, DenseNet121, MobileNet, and VGG16 achieved higher accuracy values compared to the other models. The models obtained accuracy rates of 0.9782, 0.9728, and 0.9619, respectively. Hence, these three models were combined using the weighted ensemble method to perform the final classification. On the other hand, for the 400x magnification ratio, MobileNet, Inception-V3, and Xception models achieved the highest accuracy rates. MobileNet performed the best with an accuracy rate of 0.9851, followed by Inception-V3 and Xception with accuracy rates of 0.9801. These three models were also combined using the weighted ensemble method for the final classification. With the weighted ensemble method, the accuracy value increased to 0.9864 for the 200x magnification ratio and 0.9901 for the 400x magnification ratio.

Table 4. Classification Results of Models at 200x Magnification on EBHI Dataset

Model	Acc.	Category	Precision	Recall	F1 Score	AUC
DenseNet121	0.9782	Benign	0.9783	0.9783	0.9783	0.9782
		Malignant	0.9783	0.9783	0.9783	
Inception-V3	0.9592	Benign	0.9378	0.9837	0.9602	0.9592
		Malignant	0.9829	0.9348	0.9582	
MobileNet	0.9728	Benign	0.9677	0.9783	0.9730	0.9728
		Malignant	0.9780	0.9674	0.9727	
VGG16	0.9619	Benign	0.9620	0.9620	0.9620	0.9619
		Malignant	0.9620	0.9620	0.9620	
VGG19	0.9538	Benign	0.9514	0.9565	0.9539	0.9538
		Malignant	0.9563	0.9511	0.9537	
Xception	0.9510	Benign	0.9235	0.9837	0.9526	0.9510
		Malignant	0.9826	0.9185	0.9494	
Ensemble	**0.9864**	Benign	**0.9838**	**0.9891**	**0.9864**	**0.9864**
		Malignant	**0.9891**	**0.9837**	**0.9864**	

Table 5. Classification Results of Models at 400x Magnification on EBHI Dataset

Model	Acc.	Category	Precision	Recall	F1 Score	AUC
DenseNet121	0.9727	Benign	0.9806	0.9665	0.9735	0.9729
		Malignant	0.9646	0.9795	0.9720	
Inception-V3	0.9801	Benign	0.9718	0.9665	0.9810	0.9798
		Malignant	0.9895	0.9692	0.9793	
MobileNet	0.9851	Benign	0.9903	0.9809	0.9856	0.9853
		Malignant	0.9797	0.9897	0.9847	
VGG16	0.9603	Benign	0.9289	1.0000	0.9631	0.9589
		Malignant	1.0000	0.9179	0.9572	
VGG19	0.9752	Benign	0.9761	0.9761	0.9761	0.9752
		Malignant	0.9744	0.9744	0.9744	
Xception	0.9801	Benign	0.9809	0.9809	0.9809	0.9801
		Malignant	0.9795	0.9795	0.9795	
Ensemble	**0.9901**	Benign	**0.9904**	**0.9904**	**0.9904**	**0.9900**
		Malignant	0.9897	**0.9897**	**0.9897**	

In conclusion, different deep learning models and the weighted ensemble method were employed in this study to classify tumors in histopathology images with varying magnification ratios. The highest performance was achieved with the DenseNet121, MobileNet, and VGG16 models for the 200x magnification ratio, and the MobileNet, Inception-V3, and Xception models for the 400x magnification ratio. The combination of these models using the weighted ensemble method resulted in higher accuracy rates.

10. Discussion

The experimental results are compared with other successful studies in the literature. Table 6 illustrates the comparison with previous studies on the EBHI dataset. Hu et al. [21] achieved an accuracy of 0.9537 for the 200x magnification ratio. In the study by Yengec et al. [32], a binary classification was performed without specifying the magnification level, and an accuracy rate of 0.911 was achieved. In comparison with these studies, our proposed model achieved accuracy rates of 0.9864 for the 200x magnification ratio and 0.9901 for the 400x magnification ratio. Our results indicate a higher performance compared to the outcomes of previous studies.

Table 6. Performance Comparison with previous works on EBHI Dataset

Study	Magnification	Accuracy	Category	Precision	Recall	F1 Score	AUC
Hu et al.[22]	200x	0.9537	Benign	0.944	0.965	0.954	-
			Malignant	0.964	0.943	0.953	
Yengec et al.[23]	-	0.911	Binary	0.8874	0.9436	0.9146	-
Proposed Model	200x	**0.9864**	Benign	**0.9838**	**0.9891**	**0.9864**	**0.9864**
			Malignant	**0.9891**	**0.9837**	**0.9864**	
	400x	**0.9901**	Benign	**0.9904**	**0.9904**	**0.9904**	
			Malignant	0.9897	**0.9897**	**0.9897**	

11. Conclusion

In this study, deep learning models and the weighted ensemble method were used for binary classification of tumors in histopathology images as benign and malignant, on images with different magnification ratios. Binary classification was conducted using various pre-trained CNN models and the transfer learning method. Highest accuracy rates were achieved for 200x and 400x magnification ratios, respectively, in the DenseNet121, MobileNet, and VGG16 models for 200x; and the MobileNet, Inception-V3, and Xception models for 400x. Combining the best models with weighted ensemble, the accuracy value increased to 0.9864 for the 200x magnification ratio and to 0.9901 for the 400x magnification ratio. The results of this study demonstrate

that high performance can be achieved in the classification of tumors in histopathology images through the use of various deep learning models and ensemble methods.

Declarations

Data availability

My manuscript has associated data in a data repository

Conflict of Interests

The authors declare that they have no conflict of interest.

REFERENCES

[1] Keum, N. and E. Giovannucci, *Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies.* Nature reviews Gastroenterology & hepatology, 2019. **16**(12): p. 713-732.

[2] Mármol, I., et al., *Colorectal carcinoma: a general overview and future perspectives in colorectal cancer.* International journal of molecular sciences, 2017. **18**(1): p. 197.

[3] Świderska, M., et al., *The diagnostics of colorectal cancer.* Contemporary Oncology/Współczesna Onkologia, 2014. **18**(1): p. 1-6.

[4] Pamudurthy, V., N. Lodhia, and V.J. Konda. *Advances in endoscopy for colorectal polyp detection and classification*. in *Baylor University Medical Center Proceedings*. 2020. Taylor & Francis.

[5] Wang, K.-S., et al., *Accurate diagnosis of colorectal cancer based on histopathology images using artificial intelligence.* BMC medicine, 2021. **19**(1): p. 1-12.

[6] Hitchcock, C.L., *The future of telepathology for the developing world.* Archives of pathology & laboratory medicine, 2011. **135**(2): p. 211-214.

[7] Black-Schaffer, W.S., et al., *Training pathology residents to practice 21st century medicine: a proposal.* Academic Pathology, 2016. **3**: p. 2374289516665393.

[8] Murtaza, G., et al., *Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges.* Artificial Intelligence Review, 2020. **53**: p. 1655-1720.

[9] Esteva, A., et al., *Dermatologist-level classification of skin cancer with deep neural networks.* nature, 2017. **542**(7639): p. 115-118.

[10] Mohsen, H., et al., *Classification using deep learning neural networks for brain tumors.* Future Computing and Informatics Journal, 2018. **3**(1): p. 68-71.

[11] Kuntz, S., et al., *Gastrointestinal cancer classification and prognostication from histology using deep learning: Systematic review.* European Journal of Cancer, 2021. **155**: p. 200-215.

[12] Chen, X., et al., *Recent advances and clinical applications of deep learning in medical image analysis.* Medical Image Analysis, 2022: p. 102444.

[13] Litjens, G., et al., *A survey on deep learning in medical image analysis.* Medical image analysis, 2017. **42**: p. 60-88.

[14] Rajpurkar, P., et al., *Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning.* arXiv preprint arXiv:1711.05225, 2017.

[15] Akbari, M., et al. *Classification of informative frames in colonoscopy videos using convolutional neural networks with binarized weights*. in *2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. 2018. IEEE.

[16] Ponzio, F., et al. *Colorectal cancer classification using deep convolutional networks*. in *Proceedings of the 11th international joint conference on biomedical engineering systems and technologies*. 2018.

[17] Poudel, S., et al., *Colorectal disease classification using efficiently scaled dilation in convolutional neural network.* IEEE Access, 2020. **8**: p. 99227-99238.

[18] Sarwinda, D., et al., *Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer.* Procedia Computer Science, 2021. **179**: p. 423-431.

[19] Su, Y., et al., *Colon cancer diagnosis and staging classification based on machine learning and bioinformatics analysis.* Computers in Biology and Medicine, 2022. **145**: p. 105409.

[20] Naga Raju, M.S. and B. Srinivasa Rao, *Lung and colon cancer classification using hybrid principle component analysis network-extreme learning machine.* Concurrency and Computation: Practice and Experience, 2023. **35**(1): p. e7361.

[21] Kumar, A., A. Vishwakarma, and V. Bajaj, *Crccn-net: Automated framework for classification of colorectal tissue using histopathological images.* Biomedical Signal Processing and Control, 2023. **79**: p. 104172.

[22] Hu, W., et al., *EBHI: A new Enteroscope Biopsy Histopathological H&E Image Dataset for image classification evaluation.* Physica Medica, 2023. **107**: p. 102534.

[23] Yengec-Tasdemir, S.B., et al., *Improved classification of colorectal polyps on histopathological images with ensemble learning and stain normalization.* Computer Methods and Programs in Biomedicine, 2023. **232**: p. 107441.

[24] Van der Maaten, L. and G. Hinton, *Visualizing data using t-SNE.* Journal of machine learning research, 2008. **9**(11).

[25] Yurdakul, M., İ. Atabaş, and Ş. Taşdemir. *Flower Pollination Algorithm-Optimized Deep CNN Features for Almond (Prunus dulcis) Classification*. in *2024 International Conference on Emerging Systems and Intelligent Computing (ESIC)*. 2024. IEEE.

[26] Yurdakul, M., İ. Atabaş, and Ş. Taşdemir, *Almond (Prunus dulcis) varieties classification with genetic designed lightweight CNN architecture.* European Food Research and Technology, 2024: p. 1-14.

[27] Uyar, K., M. Yurdakul, and Ş. Taşdemir, *Abc-based weighted voting deep ensemble learning model for multiple eye disease detection.* Biomedical Signal Processing and Control, 2024. **96**: p. 106617.

[28] Ayhan, B., E. Ayan, and Y. Bayraktar, *A novel deep learning-based perspective for tooth numbering and caries detection.* Clinical Oral Investigations, 2024. **28**(3): p. 178.

[29] Ayan, E., et al., *Dental student application of artificial intelligence technology in detecting proximal caries lesions.* Journal of Dental Education, 2024. **88**(4): p. 490-500.

[30] Ge, S., et al. *Detecting masked faces in the wild with lle-cnns*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

[31] Akbar, S., et al. *Transitioning between convolutional and fully connected layers in neural networks*. in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3*. 2017. Springer.

[32] Zhang, Q. *Convolutional neural networks*. in *Proceedings of the 3rd International Conference on Electromechanical Control Technology and Transportation*. 2018.

[33] Albawi, S., T.A. Mohammed, and S. Al-Zawi. *Understanding of a convolutional neural network*. in *2017 international conference on engineering and technology (ICET)*. 2017. Ieee.

[34] Huang, G., et al. *Densely connected convolutional networks*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

[35] Szegedy, C., et al. *Rethinking the inception architecture for computer vision*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[36] Howard, A.G., et al., *Mobilenets: Efficient convolutional neural networks for mobile vision applications.* arXiv preprint arXiv:1704.04861, 2017.

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

[38] Chollet, F. *Xception: Deep learning with depthwise separable convolutions*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

[39] Polikar, R., *Ensemble learning.* Ensemble machine learning: Methods and applications, 2012: p. 1-34.

[40] Inaba, Y., J.A. Chen, and S.R. Bergmann, *Carotid plaque, compared with carotid intima-media thickness, more accurately predicts coronary artery disease events: a meta-analysis.* Atherosclerosis, 2012. **220**(1): p. 128-133.