ViT-DeiT: An Ensemble Model for Breast Cancer Histopathological Images Classification

Amira Alotaibi
Computer Science Department
Umm Al-Qura University
Jamoum, Saudi Arabia
s44280250@st.uqu.edu.sa

Tarik Alafif
Computer Science Department
Umm Al-Qura University
Jamoum, Saudi Arabia
tkafif@uqu.edu.sa

Faris Alkhilaiwi
Natural Products and Alternative Medicine Department
King Abdulaziz University
Jeddah, Saudi Arabia
Faalkhilaiwi@kau.edu.sa

Yasser Alatawi
Pharmacy Practice Department
University of Tabuk
Tabuk, Saudi Arabia
Yasser@ut.edu.sa

Hassan Althobaiti
Computer Science Department
Umm Al-Qura University
Jamoum, Saudi Arabia
hmthobaiti@uqu.edu.sa

Abdulmajeed Alrefaei
Biology Department
Umm Al-Qura University
Jamoum, Saudi Arabia
afrefaei@uqu.edu.sa

Yousef Hawsawi
Research Center
King Faisal Specialist Hospital and Research Center
Jeddah, Saudi Arabia
hyousef@kfshrc.edu.sa

Tin Nguyen
Computer Science and Engineering Department
University of Nevada
Nevada, United State of America
tinn@unr.edu

2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC) | 978-1-6654-7275-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICAISC56366.2023.10085467

***Abstract*—Breast cancer is the most common cancer in the world and the second most common type of cancer that causes death in women. The timely and accurate diagnosis of breast cancer using histopathological images is crucial for patient care and treatment. Pathologists can make more accurate diagnoses with the help of a novel approach based on computer vision techniques. This approach is an ensemble model of two pre-trained vision transformer models, namely, Vision Transformer (ViT) and Data-Efficient Image Transformer (DeiT). The ViT-DeiT ensemble model is a soft voting model that combines the ViT model and the DeiT model. The proposed ViT–DeiT model classifies breast cancer histopathology images into eight classes, four of which are categorized as benign, whereas the others are categorized as malignant. The BreakHis public dataset is used to evaluate the proposed model. The experimental results show 98.17% accuracy, 98.18% precision, 98.08% recall, and a 98.12% F1 score, which outperform existing classification models.**

***Index Terms*—breast cancer, vision transformer, histopathological images, image classification, computer-aided system.**

I. Introduction

According to the World Health Organization, 2.3 million women worldwide were diagnosed with breast cancer in 2020, with 685,000 deaths, making it the most prevalent cancer globally. As many as 7.8 million women were diagnosed with breast cancer between 2015 and 2020 [1].

The timely diagnosis of breast cancer can increase the survival rate; hence, several techniques, such as mammography, magnetic resonance imaging, ultrasound, computed tomography, positron emission tomography, biopsy, and microwave imaging, are used in clinics to diagnose this disease [2]. Although years of experience are usually required for radiologists to correctly diagnose malignant tumors from histopathological images, experts occasionally differ in their conclusions. The usage of computer-aided diagnosis (CAD) for image diagnosis can help medical experts make precise decisions [3].

The Food and Drug Administration approved the first viable CAD system for second-opinion screening mammography in 1998 [4]. Histopathological images of breast cancer can be employed for clinical applications to automatically and accurately detect malignant tumors. Moreover, deep learning algorithms have been extensively employed to improve detection performance. The deep learning algorithms have successfully escalated the performance of the classification of histopathological images [5].

Convolutional Neural Networks (CNNs) are extensively implemented in computer vision applications, including the detection of breast cancer using histopathological images [6]–[8]. Many studies have been conducted to improve the performance of CNNs for breast cancer image classification [7], [9], [10]. Training a Vision Transformer (ViT) [11] with sufficiently large data has been shown to achieve remarkable results. The ViT outperforms comparable state-of-the-art CNNs, with four times less computational effort. Nevertheless, transformers were originally innovated for natural language processing [12]. In a transformer, a sequence of tokens is passed as input, but in ViT, image patches are passed as inputs. Similar to ViT, Data-Efficient Image Transformer (DeiT) [13] is proposed, which has made a great success in computer vision applications.

In this study, an ensemble model based on the ViT and DeiT models is proposed for breast cancer histopathological

image classification. The main contributions of our work are summarized as follows:

- A ViT– DeiT ensemble model is proposed for histopathological image classification. The ensemble model is fine-tuned using transfer learning for multi-class classification.
- We investigate image magnification dependent and independent approaches on a BreakHis dataset [14] for multi-class classification.
- The classification performance is compared and evaluated with existing models. In terms of classification performance, the ViT–DeiT model surpasses other models.

The remaining parts of the paper are organized as follows: We present related works in Section II, and review the dataset in Section III. We provide ViT and DeiT details in the preliminaries Section IV and introduce our proposed model in Section V. Section VI provides an experimental setup, results of the proposed model, and comparison with similar studies. Then, Section VII concludes the paper and presents our future work.

II. Related Work

Recently, studies in the field of breast cancer classification have focused on ultrasound image classification [15]–[17], biopsy data classification [18]–[20], and histopathological image classification [7], [9], [21]–[25]. In this study, we only focus on the latter, considering that previous studies have not reached sufficient performance.

Several studies have used CNNs to classify breast cancer using public datasets, with satisfactory results [7], [9], [21]. Parvin et al. [22] compared the performance of five CNN architectures. The models were evaluated via magnification-dependent classification using a public dataset named BreakHis. The best results were achieved using Inception-v1, showing accuracies of 89% to 94% for binary classification. These results are considered good, as there were few images to train from scratch. Likewise, Agarwal et al. [23] proposed and analyzed the performance of four CNN-based architectures, including VGG-16, VGG-19, MobileNet, and ResNet-50 using the BreakHis dataset. The VGG-16 architecture achieved the highest accuracy of 94.67% for binary classification.

Another deep learning approach proposed by Zhou et al. [24] was based on a resolution adaptive network (RANet) model and anomaly detection with an SVM (ADSVM) for binary and multiclass classification using the BreakHis dataset. The RANet–ADSVM model was trained and compared with and without a balanced dataset. In the experiments, the RANet–ADSVM approach achieved the highest accuracy of 93.35% to 99.14% for multi-class classification with balanced data. Although the RANet–ADSVM method achieved better performance with a balanced dataset than with an imbalanced dataset, there were marked improvements in classification performance.

Seo et al. [25] proposed a method based on the Primal-Dual Multi-Instance SVM model. The method was evaluated for the binary classification of images with magnifications dependent using the BreakHis dataset. The accuracy ranged from 85.3% to 89.8%. Although many studies have used the BreakHis dataset, a large number of them worked on binary classification, and those that worked on multiple classifications such as RANet-ADSVM [24] did not reach significant results. In addition, very few studies were conducted on multi-class magnification-independent classification.

III. BreakHis Dataset

The BreakHis dataset [14] contains microscopic breast tumor biopsy images. The tumors are either benign or malignant. The dataset covers 7,909 images. The images are gathered from 82 patients using four magnification factors (40×, 100×, 200×, and 400×), according to the objective lenses (4×, 10×, 20×, and 40×). Each image is captured using an automatic exposure setting and performed manually using a computer screen to view the image. The dataset contained 2,480 benign tumor images and 5,429 malignant tumor images. The benign tumor images are divided into adenosis (A), tubular adenoma (TA), fibroadenoma (F), and phyllodes tumor (PT). The malignant tumor images are divided into ductal carcinoma (DC), lobular carcinoma (LC), papillary carcinoma (PC), and mucinous carcinoma (MC). The statistics of the BreakHis dataset are shown in Table I.

TABLE I
NUMBER OF IMAGES FOR BENIGN AND MALIGNANT CATEGORIES IN DETAIL.

Main category		Benign				Malignant				Total
Sub category		A	F	TA	PT	DC	LC	MC	PC	
Magnification factor	40X	114	253	109	149	864	156	205	145	1,995
	100X	113	260	121	150	903	170	222	142	2,081
	200X	111	264	108	140	896	163	196	135	2,013
	400X	106	237	115	130	788	137	169	138	1,820
Total		444	1,014	453	569	3,451	626	792	560	7,909

IV. Preliminaries

A. ViT Model

The transformer is widely used for NLP [12]. The structure of a transformer model includes an encoder and a decoder. The decoder is not required in the ViT structure [11]. Therefore, the ViT structure consists of only the encoder for image processing, as shown in Fig. 1.The encoder component consists of normalization layers, a multi-head attention layer, and a feed-forward layer. The multi-head attention is a type of self-attention that functions to provide attention to specific information from various aspects.

In the ViT model, each image is passed through a linear embedding layer before being fed to the encoder. The embedding layer divides the image into equal-sized patches that are flattened into a one-dimensional vector. The position of the embedding is added to the flattened patches, and the class of the embedded image is added. After the encoder processes these inputs, it produces the output. Then, the output is passed through the MLP head structure, which performs the

classification task. The class is the output of the MLP head structure. The MLP head structure consists of two connected layers with a GELU activation function [11]. Recently, Google AI released more than one model, which were trained on various datasets of different sizes.

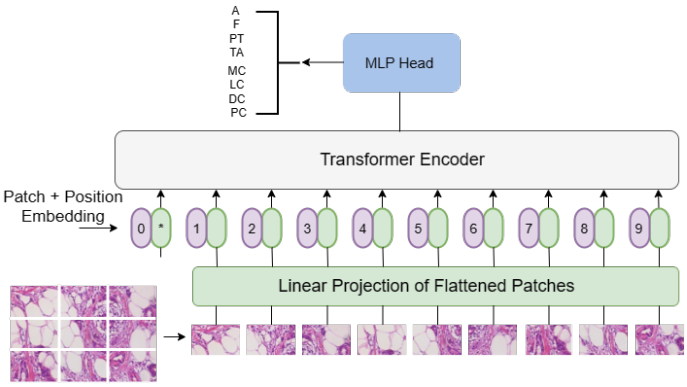

Fig. 1. Vision Transformer.

B. DeiT Model

Facebook AI has recently introduced the DeiT model [13]. The DeiT model is trained in fewer days and on one machine compared to the ViT model [11]. Similar to the ViT model, the DeiT model is also trained on the ImagNet dataset. The model shows improvements over previous ViT models. The structure of the DeiT model is built based on the ViT model [11]. They are added with a feed-forward network (FFN) above the multi-head self-attention (MSA) layer, which comprises two linear layers separated by GELU activation. As shown in Fig. 2, there is an extra input called a distillation token. This token allows the model to learn from the teacher's output. The authors attempted soft distillation and hard distillation. However, the latter achieves the best results.

V. Proposed Method

The proposed method aims to classify histopathological images of breast cancer into eight categories based on magnification-dependent and magnification-independent approaches. As shown in Fig.2, the proposed method has three steps as follows:

1) Preprocessing the dataset.
2) Fine-tuning the ViT and DeiT models.
3) Developing the ViT–DeiT ensemble model.

A. Preprocessing the Dataset

In the BreakHis dataset, the number of images in the subcategories is uneven. The DC subcategory has the largest number of images at different magnifications, as shown in Table I. The F subcategory has the second-largest number of images, and the other subcategories have a similar number of images. Thus, this causes an imbalance in the data, which leads to overfitting [26]. To avoid the overfitting issue, an undersampling technique is employed. The undersampling technique reduces the number of samples in each subcategory with a large

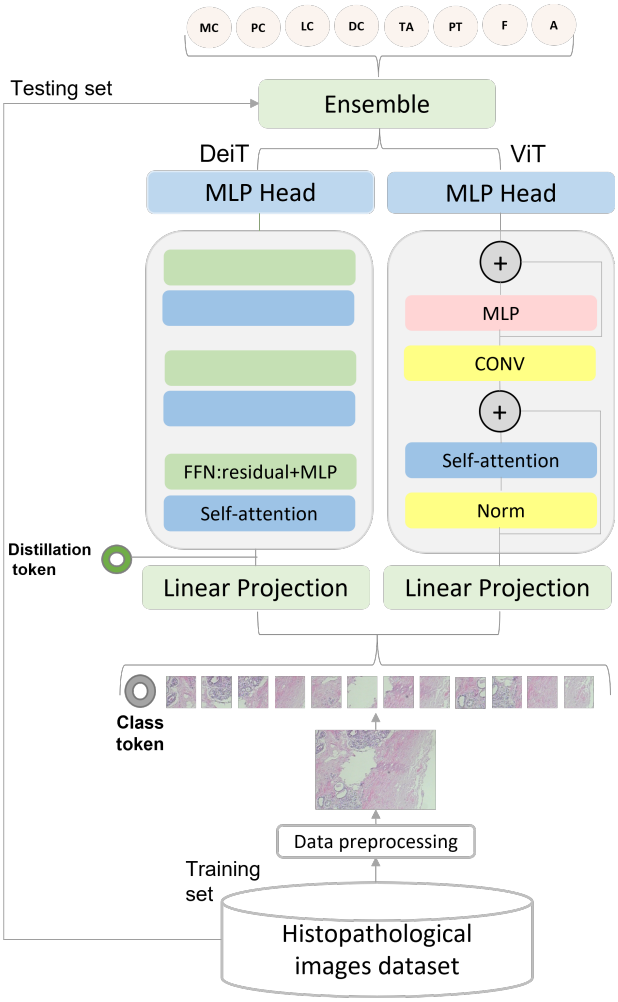

Fig. 2. Structure of the proposed ViT–DeiT ensemble model.

sample size. Moreover, the undersampling technique is also used to balance the dataset before training the ensemble model using magnification-dependent and magnification-independent approaches.

B. Fine-tuning the ViT and DeiT Models

To train the models, deep learning requires a large number of samples. To address this problem, a transfer learning technique is employed. With a few images, a pre-trained model derived from training on large datasets can be fine-tuned, thereby greatly reducing the training time. In addition to improving the convergence speed and generalization ability of the model, the transfer learning also reduces the risk of overfitting. In this work, both the ViT and the DeiT models are fine-tuned. The models are fine-tuned by placing a prediction head on top of the final hidden state of the class token to classify the images from the eight subcategories. Although the models consist of twelve self-attention heads, the outputs of the heads are combined to produce a final attention score. The final attention score is used to provide attention to a region of interest inside each histopathological image. The

region of interest represents the cancer cells, which is used for detecting and determining the type of breast cancer. However, the acquired attention scores from the ViT and the DeiT models differ since they have different structures.

C. Developing the ViT-DeiT Ensemble Model

The proposed ensemble model uses multi-learning to achieve better performance than that achieved by any single model. Fig. 2 shows the structure of the ensemble model, which combines two different models, the ViT model and the DeiT model. Compared to the ViT model, the DeiT model uses the distillation token to learn effectively from a teacher. The distillation token is learned through backpropagation by interacting with classes and patch tokens through the self-attention layers.

A soft voting technique is proposed to attain the highest probability value from the ensemble model. The soft voting technique works by assigning the high average probability as a predicted label for each sample. When an image is taken as an input, the two models provide probability values for each class. Then, the probabilities are summed for each class and divided by the number of classifiers. Then, the highest probability value is assigned for the predicted label, as computed in (1).

$$\hat{y} = \underset{i}{\operatorname{argmax}} \left\{ \frac{1}{N} \sum_{j=1}^{n} p_{ij} \right\} \tag{1}$$

where N is the number of classifiers, and p_{ij} is the probability value of j^{th} classifier for the i^{th} category.

VI. Experimental Results and Evaluations

A. Experimental Setup

In the training stage, the assigned value to the learning rate is 1e-4, the weight decay is set to 0.001. The batch size is set to 16 while the number of epochs is set to 15. The histopathological images were then divided into 80% for training and 20% for testing. The training images are balanced to avoid bias and overfitting in the classification. The code is written in Python. It is publicly available in the GitHub repository.

B. Results

The proposed model is evaluated using a magnification-independent approach. The same parameters are used across the magnification groups.

As shown in Table I, the BreakHis dataset consists of images grouped into eight subcategories within two main categories, benign and malignant. Accordingly, the performance of the breast cancer classification approach is evaluated for both binary and multiple classes. To evaluate the effect of the balanced data scenario, we conducted an analysis with and without magnification factors for the classification system based on the ViT–DeiT ensemble model.

For the binary classification, the ViT and DeiT cover all images for magnification-independent classification. For the multi-class classification, Table II shows the performance of the ViT and the DeiT models on the testing set. The overall performance of the ViT and the DeiT individual models is close, in which the accuracy of the ViT model is 0.28% higher than that of the DeiT model.

TABLE II
Performance of ViT and DeiT models before combination.

Model	Accuracy %	Precision %	Recall %	F1 score %
ViT	97.75	97.78	97.67	97.71
DeiT	97.47	97.47	97.41	97.43

The accuracy can be used to judge a model's classification ability, but it cannot reflect specific details. When the classification model makes predictions, the confusion matrix indicates the prediction details for each category by comparing the predicted result with the actual value. As shown in Fig. 3 the confusion matrix is used to further evaluate the classification ability and details of the ViT-DeiT model.

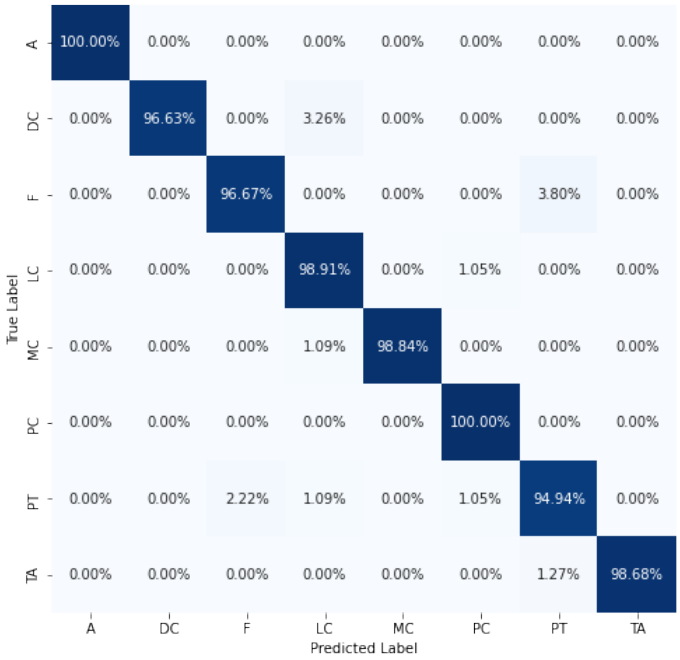

Fig. 3. Confusion matrix for the results from the ensemble model.

In addition, Table III shows the performance of the ViT-DeiT model with the balancing and imbalancing issues using the BreakHis dataset. In the magnification-independent evaluation, using the balanced dataset scenario, the multi-class classification accuracy attains 3.99% higher than without balancing the dataset. In addition, other evaluation metrics also demonstrate that using the balanced dataset scenario improves the overall classification performance.

For magnification-independent evaluation, an accuracy of 98.17% is achieved for the multi-class classification using the balanced dataset as shown in Table III. However, for magnification-dependent evaluation, the best accuracy (99.43%) for multi-class classification is achieved at 40× mag-

TABLE III
PERFORMANCE OF THE VIT–DEIT MODEL (%) FOR MAGNIFICATION-INDEPENDENT MULTI-CLASS CLASSIFICATION USING BALANCED AND IMBALANCED DATASET SCENARIOS.

Scenario	Accuracy	Precision	Recall	F1 score
Imbalanced dataset	94.18	94.62	93.08	93.80
Balanced dataset	**98.17**	**98.18**	**98.08**	**98.12**

nification. Also, the model achieves better precision, recall, and F1 score at the same magnification as shown in Table IV.

TABLE IV
PERFORMANCE OF THE PROPOSED VIT-DEIT MODEL (%) USING ALL MAGNIFICATION FACTORS FOR MULTI-CLASS CLASSIFICATION TASK.

Magnification	Accuracy	Precision	Recall	F1 score
40X	**99.43**	**99.38**	**99.46**	**99.40**
100X	98.34	98.31	98.51	98.35
200X	98.27	98.32	98.27	98.23
400X	98.82	98.57	98.78	98.65

The ViT–DeiT model avoids the worst case in cancer diagnosis, that is, the diagnosis of a malignant sample as benign [27]. The most misclassified images are the malignant samples that are predicted to be other types of malignant tumors, as shown in Fig. 4.

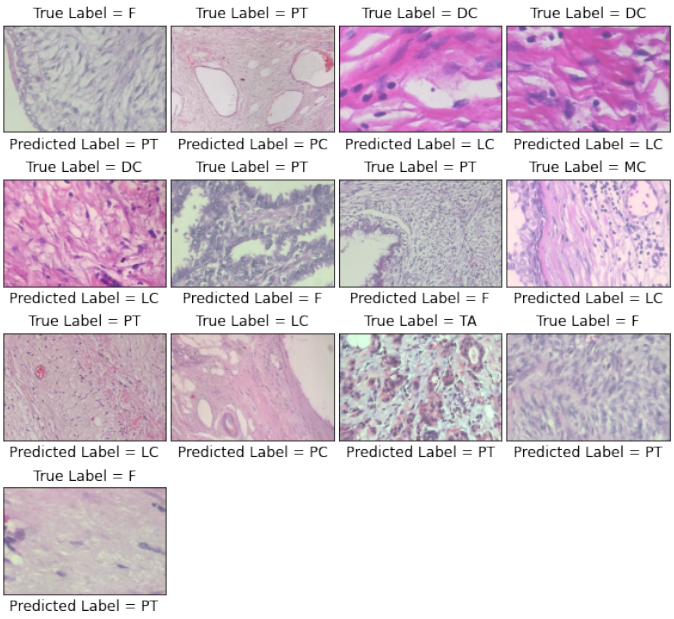

Fig. 4. Misclassified samples by the ViT-DeiT ensemble model with true and predicted labels, showing that no malignant tumor was classified as a benign tumor.

Fig.5 shows the attention maps of the ViT and DeiT models after the training using the BreakHis dataset. The attention maps are focused on cancerous cells while they provide little attention to wrong regions as they are also diagnosed by domain of experts. The effectiveness of soft voting appears to minimize the error rate in diagnosing breast cancer.

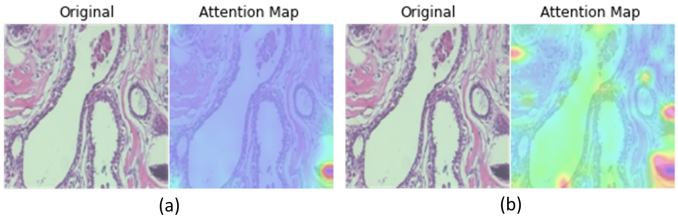

Fig. 5. Attention maps from the (a) DeiT and (b) ViT models on a sample image from the BreakHis dataset.

C. Comparison with Existing and Recent Works

The performance of the methods in recent studies is compared with that of the proposed ViT–DeiT ensemble model. The performance of the multi-class classification with magnification-dependent evaluation is shown in Table V. Our model achieves the best results on various magnification factors with large margins among state-of-the-art models [3], [24], [28], [29].

TABLE V
PERFORMANCE COMPARISON OF MAGNIFICATION-DEPENDENT MULTI-CLASS CLASSIFICATION AGAINST EXISTING MODELS.

Model	Magnification	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Deep-Net [3]	40X	94.43	95.25	95.55	95.39
	100X	94.45	94.51	94.64	94.42
	200X	92.27	90.71	92.24	91.42
	400X	91.15	90.74	91.09	90.75
ResNet-18 [28]	40X	94.49	93.81	94.78	94.15
	100X	93.27	92.94	91.59	92.23
	200X	91.29	91.18	88.28	89.47
	400X	89.56	87.97	87.97	87.77
SE-ResNet [29]	40X	86.89	-	-	-
	100X	88.69	-	-	-
	200X	86.53	-	-	-
	400X	86.37	-	-	-
RANet-ADSVM [24]	40X	91.14	-	-	-
	100X	96.83	-	-	-
	200X	98.05	-	-	-
	400X	90.30	-	-	-
ViT-DeiT (Ours)	40X	**99.43**	**99.38**	**99.46**	**99.40**
	100X	**98.34**	**98.31**	**98.51**	**98.35**
	200X	**98.27**	**98.32**	**98.27**	**98.23**
	400X	**98.82**	**98.57**	**98.78**	**98.65**

Few studies have focused on magnification-independent fields. Table VI shows the performance of studies for multi-class classification for the BreakHis dataset with magnification-independent evaluation. Our model achieves the highest results for magnification-independent classification. However, the classification results showed 98.17% accuracy, 98.18% precision, 98.08% recall and 98.12% F1 score, which were better than the results (93.32% accuracy, 92.98% precision, 92.36% recall, and 92.44% F1 score) obtained with the Xception approach [30].

VII. CONCLUSION

An ensemble classification model is proposed to accurately classify eight types of breast cancer from histopathological images and assist pathologists in diagnosis. The BreakHis

TABLE VI
PERFORMANCE COMPARISON OF MAGNIFICATION-INDEPENDENT MULTI-CLASS CLASSIFICATION AGAINST EXISTING MODELS.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
ResNet-18 [28]	92.03	91.39	90.28	90.77
6B-Net [31]	90.10	-	-	-
Xception [30]	93.32	92.98	92.36	92.44
ViT-DeiT (Ours)	**98.17**	**98.18**	**98.08**	**98.12**

dataset is used to evaluate the proposed model that integrates the pre-trained ViT and DeiT models. A soft voting technique is also introduced to work on top of the ensemble model to attain the highest probability value for a cancerous class. The performance of the ensemble model in magnification-dependent and magnification-independent multi-class classification achieves better results with larger margins compared to the state-of-the-art classification models. Moreover, the ensemble model avoids misclassifying a malignant tumor as benign. These promising results demonstrate that CAD systems can be trusted to classify breast cancer, which is another step toward automating the diagnosis of breast cancer. In the future, we will investigate the possibility of combining multiple pre-trained models to attain higher performance in this complex medical domain.

REFERENCES

[1] "Breast cancer." https://www.who.int/news-room/fact-sheets/detail/breast-cancer, March 2021.

[2] L. Wang, "Early diagnosis of breast cancer," *Sensors*, vol. 17, no. 7, p. 1572, 2017.

[3] Y. Jiang, L. Chen, H. Zhang, and X. Xiao, "Breast cancer histopathological image classification using convolutional neural networks with small se-resnet module," *PloS one*, vol. 14, no. 3, p. e0214587, 2019.

[4] H.-P. Chan, R. K. Samala, and L. M. Hadjiiski, "Cad and ai for breast cancer—recent development and challenges," *The British journal of radiology*, vol. 93, no. 1108, p. 20190580, 2019.

[5] H. A. Khikani, N. Elazab, A. Elgarayhi, M. Elmogy, and M. Sallah, "Breast cancer classification based on histopathological images using a deep learning capsule network," *arXiv preprint arXiv:2208.00594*, 2022.

[6] Y. Celik, M. Talo, O. Yildirim, M. Karabatak, and U. R. Acharya, "Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images," *Pattern Recognition Letters*, vol. 133, pp. 232–239, 2020.

[7] D. Albashish, R. Al-Sayyed, A. Abdullah, M. H. Ryalat, and N. A. Almansour, "Deep cnn model based on vgg16 for breast cancer classification," in *2021 International Conference on Information Technology (ICIT)*, pp. 805–810, IEEE, 2021.

[8] H. Alghodhaifi, A. Alghodhaifi, and M. Alghodhaifi, "Predicting invasive ductal carcinoma in breast histology images using convolutional neural network," in *2019 IEEE National Aerospace and Electronics Conference (NAECON)*, pp. 374–378, IEEE, 2019.

[9] K. Gupta and N. Chawla, "Analysis of histopathological images for prediction of breast cancer using traditional classifiers with pre-trained cnn," *Procedia Computer Science*, vol. 167, pp. 878–889, 2020.

[10] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Classification of histopathological biopsy images using ensemble of deep learning networks," *arXiv preprint arXiv:1909.11870*, 2019.

[11] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[12] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in neural information processing systems*, pp. 5998–6008, 2017.

[13] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, "Training data-efficient image transformers & distillation through attention," in *International Conference on Machine Learning*, pp. 10347–10357, PMLR, 2021.

[14] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *Ieee transactions on biomedical engineering*, vol. 63, no. 7, pp. 1455–1462, 2015.

[15] M. Li, "Research on the detection method of breast cancer deep convolutional neural network based on computer aid," in *2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, pp. 536–540, IEEE, 2021.

[16] B. Gheflati and H. Rivaz, "Vision transformers for classification of breast ultrasound images," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 480–483, IEEE, 2022.

[17] J. F. Lazo, S. Moccia, E. Frontoni, and E. De Momi, "Comparison of different cnns for breast tumor classification from ultrasound images," *arXiv preprint arXiv:2012.14517*, 2020.

[18] S. Punitha, F. Al-Turjman, and T. Stephan, "An automated breast cancer diagnosis using feature selection and parameter optimization in ann," *Computers & Electrical Engineering*, vol. 90, p. 106958, 2021.

[19] O. I. Obaid, M. A. Mohammed, M. Ghani, A. Mostafa, and F. Taha, "Evaluating the performance of machine learning techniques in the classification of wisconsin breast cancer," *International Journal of Engineering & Technology*, vol. 7, no. 4.36, pp. 160–166, 2018.

[20] O. Alagoz, M. A. Ergun, M. Cevik, B. L. Sprague, D. G. Fryback, R. E. Gangnon, J. M. Hampton, N. K. Stout, and A. Trentham-Dietz, "The university of wisconsin breast cancer epidemiology simulation model: an update," *Medical decision making*, vol. 38, no. 1_suppl, pp. 99S–111S, 2018.

[21] Y. Yari, T. V. Nguyen, and H. T. Nguyen, "Deep learning applied for histological diagnosis of breast cancer," *IEEE Access*, vol. 8, pp. 162432–162448, 2020.

[22] F. Parvin and M. A. M. Hasan, "A comparative study of different types of convolutional neural networks for breast cancer histopathological image classification," in *2020 IEEE Region 10 Symposium (TENSYMP)*, pp. 945–948, IEEE, 2020.

[23] P. Agarwal, A. Yadav, and P. Mathur, "Breast cancer prediction on breakhis dataset using deep cnn and transfer learning model," in *Data Engineering for Smart Systems*, pp. 77–88, Springer, 2022.

[24] Y. Zhou, C. Zhang, and S. Gao, "Breast cancer classification from histopathological images using resolution adaptive network," *IEEE Access*, vol. 10, pp. 35977–35991, 2022.

[25] H. Seo, L. Brand, L. S. Barco, and H. Wang, "Scaling multi-instance support vector machine to breast cancer detection on the breakhis dataset," *Bioinformatics*, vol. 38, no. Supplement_1, pp. i92–i100, 2022.

[26] Z. Li, K. Kamnitsas, and B. Glocker, "Analyzing overfitting under class imbalance in neural networks for image segmentation," *IEEE transactions on medical imaging*, vol. 40, no. 3, pp. 1065–1077, 2020.

[27] V. Papageorgiou, Z. Apalla, E. Sotiriou, C. Papageorgiou, E. Lazaridou, S. Vakirlis, D. Ioannides, and A. Lallas, "The limitations of dermoscopy: false-positive and false-negative tumours," *Journal of the European Academy of Dermatology and Venereology*, vol. 32, no. 6, pp. 879–888, 2018.

[28] S. Boumaraf, X. Liu, Z. Zheng, X. Ma, and C. Ferkous, "A new transfer learning based approach to magnification dependent and independent classification of breast cancer in histopathological images," *Biomedical Signal Processing and Control*, vol. 63, p. 102192, 2021.

[29] V. Kate and P. Shukla, "Breast cancer image multi-classification using random patch aggregation and depth-wise convolution based deep-net model," 2021.

[30] A. M. Zaalouk, G. A. Ebrahim, H. K. Mohamed, H. M. Hassan, and M. M. Zaalouk, "A deep learning computer-aided diagnosis approach for breast cancer," *Bioengineering*, vol. 9, no. 8, p. 391, 2022.

[31] T. Jain, V. K. Verma, M. Agarwal, A. Yadav, and A. Jain, "Supervised machine learning approach for the prediction of breast cancer," in *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, pp. 1–6, IEEE, 2020.