https://africanjournalofbiomedicalresearch.com/index.php/AJBR
Afr. J. Biomed. Res. Vol. 28(1) (January 2025); 29- 39
Research Article

Transformer-Based Neural Architectures For Automated Cancer Classification In Histopathology Images

Lalitha Bhavani Konkyana[1], J Rajanikanth[2], K Chandra Bhushana Rao[3], B Ramesh Naidu[4]

[1]Department of ECE, AITAM, Tekkali, AP, INDIA
[2]Department of CSE, SRKR Engineering College, Bhimavaram, AP, INDIA
[3]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, INDIA
[4]Department of Information Technology, AITAM, Tekkali, AP, INDIA

***Corresponding Author:** K V L Bhavani
*Department of ECE, AITAM, Tekkali, AP, INDIA

Abstract– Timely identification of metastatic cancer via accurate image classification is essential for enhancing patient outcomes. This research introduces a deep learning method for automated tumor identification through Transformer-Based Neural Architectures applied to histopathological images. Our model underwent training using a dataset composed of 96x96 pixel microscopic images and demonstrated remarkable performance, attaining a training accuracy of 93.9% and a validation accuracy of 93.1%. The model showed excellent effectiveness in differentiating "no tumor tissue" from "tumor tissue," reaching an ROC-AUC score of 0.9799. These findings indicate that our method is very proficient at correctly identifying tumor areas, paving the path for better diagnostic instruments in medical image analysis.

Keywords: Transformer-Based Neural Architectures, tumor, image classification, metastatic cancer, Histopathology

Received: 28 November 2024 *Accepted:07 December 2024*

DOI: https://doi.org/10.53555/AJBR.v28i1.4973

INTRODUCTION

Cancer continues to be a significant global health issue, where early detection is essential for effective treatment and patient survival. Histopathological image analysis serves as an essential diagnostic method, enabling pathologists to observe cellular alterations that may suggest possible malignancies. Conventional manual analysis of microscopic tissue specimens is labor-intensive, subjective, and susceptible to human mistakes, leading to a pressing demand for automated and accurate diagnostic technologies[1-2].

Recent progress in artificial intelligence and machine learning has created extraordinary possibilities for analyzing medical images. Deep learning frameworks, especially convolutional neural networks (CNNs), have shown impressive abilities in recognizing patterns and handling intricate image classification challenges. These technological advancements provide hopeful solutions for improving diagnostic precision, minimizing human mistakes, and possibly speeding up cancer detection processes[3-4].

The main research challenge entails creating a strong computational model that can effectively detect metastatic cancer areas in histopathological images. Particular challenges involve dealing with images that contain little tumor tissue, addressing slight cellular differences, and attaining high sensitivity and specificity in classification. Our research tackles these issues by

utilizing advanced neural network designs and refined image preprocessing methods[5].

This study makes a substantial contribution to computational pathology by offering a data-oriented method for cancer identification. Through the automation of microscopic tissue image analysis, we seek to assist pathologists in achieving more precise and prompt diagnoses. The suggested approach may decrease diagnostic turnaround times, decrease human error, and enhance overall patient care by facilitating early and accurate identification of metastatic areas[6].

The main aims of this research are: (a) creating a deep learning model with high performance for the binary classification of histopathological images, (b) assessing the model's effectiveness through standard metrics like accuracy, precision, recall, and ROC-AUC, (c) showcasing the model's feasibility for practical clinical uses, and (d) investigating the applicability of the suggested method to various medical imaging fields.

LITERATURE SURVEY

This section examines notable progress and associated studies in the utilization of deep learning, particularly transformer architectures and convolutional neural networks (CNNs), for the automated classification of cancer in histopathological images. We explore research related to model effectiveness, dataset issues, preprocessing methods, and clinical significance.

Dosovitskiy et al. [7] presented Vision Transformers (ViTs), showcasing the efficacy of transformer architectures in image classification, particularly with the availability of large datasets. Their capability to grasp global context represented a notable improvement over conventional convolutional neural networks (CNNs). This feature has now been broadened to medical imaging, allowing enhanced efficacy in tasks needing thorough contextual comprehension.

Chen et al. [8] introduced TransUNet, a hybrid architecture that integrates transformers and CNNs, aimed specifically at medical image segmentation. Their research underscored TransUNet's capacity to attain high precision and sensitivity, rendering it an ideal option for tasks requiring precise boundary delineation, like organ or lesion segmentation.

Araujo et al. [9] utilized CNNs for analyzing histopathology images of breast cancer, resulting in strong classification results. Their research highlighted the effectiveness of deep learning in identifying complex cellular traits linked to cancer, showcasing its promise for early and accurate diagnosis.

Wang et al. [10] created a deep CNN model to automatically identify metastatic cancer in lymph node specimens. Leveraging the CAMELYON16 challenge dataset, their model demonstrated impressive accuracy and sensitivity, highlighting the effectiveness of CNNs in streamlining labor-intensive diagnostic tasks.

Ciresan et al. [11] were innovators in using CNNs for classifying medical images, implementing a sliding window technique for detecting tumors at the pixel level in histopathological images. Their efforts established the groundwork for contemporary deep learning uses in medical imaging.

Komura and Ishikawa[12] highlighted the significance of preprocessing methods in enhancing the effectiveness of deep learning models for analyzing histopathological images. Methods like normalization and stain color standardization were recognized as essential for reducing data variability and improving model resilience.

Litjens et al. [13] examined publicly accessible datasets, such as CAMELYON16 and TCGA, which have greatly aided progress in cancer detection through the provision of high-quality histopathological images. Their research underscored the importance of open data in fostering innovation within medical imaging studies.

Janowczyk and Madabhushi[14] discussed the difficulties linked to histopathological data analysis, including class imbalance, variability in staining, and artifacts. They promoted strong preprocessing workflows and data enhancement methods to tackle these challenges and enhance model performance.

Russakovsky et al. [15] emphasized the significance of standardized evaluation metrics, including accuracy, precision, recall, and ROC-AUC, for assessing the performance of image classification models. Their suggestions have played a crucial role in creating uniform evaluation methods in medical imaging.

Murtaza et al. [16] introduced a framework for assessing histopathological image analysis, highlighting sensitivity and specificity as critical measures for clinical significance. Their framework offered an organized method to evaluate the effectiveness of AI models in practical diagnostic situations.

Shin et al. [17] investigated the use of transfer learning for radiology images, showing its efficacy in situations where labeled medical data is limited. Utilizing pretrained models, their method attained strong performance with little fine-tuning required.

Yu et al. [18] assessed the ability of deep learning models to transfer across various medical imaging modalities. They discovered that pretrained models could achieve competitive outcomes with little domain-specific fine-tuning, highlighting their adaptability.

Reshma and colleagues [19] created an AI-supported diagnostic instrument for histopathological evaluation, reaching clinical-grade accuracy. Their research emphasized the ability of deep learning to support and improve the skills of pathologists in identifying cancer.

Esteva et al. [20] investigated how AI can aid dermatologists, comparing its effectiveness in cancer detection to its role in pathology. Their results highlighted the transformative capabilities of AI in various medical fields.

Liu et al. [21] explored the use of Swin Transformers for biomedical image segmentation. Their study showed that these models were able to understand context better than CNNs, resulting in enhanced segmentation precision.

Shamshad et al. [22] examined the function of attention mechanisms in transformers and their capability to improve feature extraction in medical imaging applications. Their examination emphasized how attention mechanisms enhance model efficacy by

concentrating on the most pertinent areas within an image.

Mahoro et al. [23] introduced hybrid models that integrated CNNs and transformers for the classification of breast cancer. Their method combined local and global feature extraction, leading to improved classification performance.

Zou et al. [24] investigated the application of multi-scale attention mechanisms in hybrid models for histopathological image classification. Their results showcased the advantages of integrating features at various scales to enhance diagnostic precision.

Xu et al. [25] examined the development of AI in computational pathology, highlighting the promise of transformer-based architectures as advanced diagnostic instruments. Their analysis also addressed the difficulties and prospects linked to implementing these sophisticated models in clinical environments.

Mohammed et al. [26] highlighted the significance of explainable AI models in medical imaging to foster trust among clinicians and encourage broader clinical implementation. Their research emphasized the importance of transparency in AI decision-making systems to guarantee dependability and practicality in real-world uses.

The literature highlights the transformative capabilities of transformer-based and CNN models in cancer detection via histopathological image analysis. Although CNNs have shown considerable success, recent developments in transformer-based models present chances to further improve performance. Issues like data variability, class imbalance, and clinical relevance persist, demanding ongoing advancements in neural architectures and assessment frameworks.

METHODOLOGY

The approach for applying a Transformer-Based Neural Architecture in image classification tasks focuses on harnessing the capabilities of transformer models, initially created for sequential data, to efficiently handle image data. The structure is illustrated in figure 1. This approach consists of six stages.

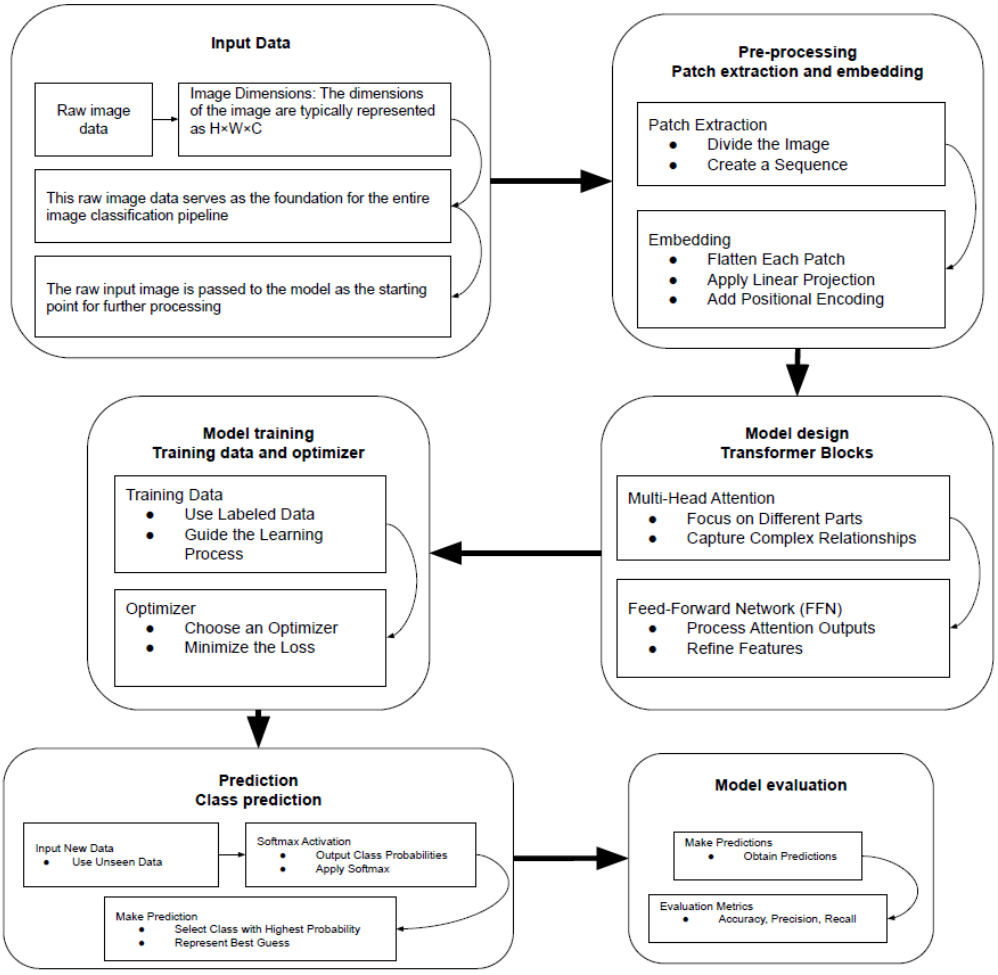

Figure 1 System structure

Phase 1: Input Image ($H \times W \times C$)

The procedure starts with segmenting the input image into smaller patches, and each patch is then linearly transformed into a high-dimensional vector space. Positional encoding is incorporated to maintain spatial details, and these patch embeddings are processed through several transformer layers. Every block is made up of multi-head self-attention mechanisms, allowing the model to concentrate on various segments of the

image at once, along with feed-forward networks that enhance the acquired features.
Unprocessed Image Data: The model receives unprocessed image data, which comprises pixel values that depict the content of the image.
Image Size: The size of the image is usually expressed as $H \times W \times C$, where:

- H represents the image's height measured in pixels.
- W represents the image's width measured in pixels.
- C represents the quantity of color channels (e.g., 3 for RGB images).

Goal of the Input: This unprocessed image data acts as the basis for the complete image classification system, offering the visual material that the model will utilize to recognize and analyze patterns.
Model Processing: The unprocessed input image is provided to the model as the initial stage for additional processing, including patch extraction, embedding, and finally classification.
This phase signifies the primary stage in which the image is introduced to the model for further modifications and examination.

Phase 2: Pre-treatment (Patch Extraction and Embedding):
Patch Retrieval
Segregate the Image: Break the image into smaller sections of predetermined dimensions (e.g., 16x16 pixels).
Generate a Sequence: Transform the image into a sequence of smaller patches, with each patch being processed separately.

Embedding
Flatten Every Patch: Transform each patch into a one-dimensional vector.
Implement Linear Projection: Send each flattened patch through a linear projection (embedding) to convert it into a vector with higher dimensions.
Incorporate Positional Encoding: Introduce positional encoding to the patch embeddings to maintain spatial information, enabling the model to comprehend the relative locations of patches within the image.

Phase 3: Model Design (Transformer Blocks)
The transformer block allows the model to identify high-level features and grasp long-range dependencies in the image. This enhanced comprehension of the input data enables the model to enhance its effectiveness in intricate image classification challenges.
This systematic method mathematically represents how transformers handle image data, providing an efficient way to model global dependencies and achieving outstanding results in tasks such as image classification and segmentation.
The algorithm consists of four stages known as input representation, transformer block, global pooling, and classification.

STEP 1: Representation of input
The input is presented as a series of features obtained from image patches or alternative processing methods. For an image input $X \in R^{H\times W\times C}$
Flatten the image into N patches, each of size $P \times PN = \frac{H.W}{P^2}, and\ each\ patch\ x_i \epsilon\ R^{P^2C}$
Project patches into a latent space using a dense layer(Linear project):

$$z_0 = [x_1W_p, x_2W_p, .., x_NW_p] + E_{pos}$$

Where W_p represents the trainable projection matrix and E_{pos} denotes the positional encoding incorporated to maintain spatial details.

STEP 2: Transformer Block (its repeat for each block
(a) Multi-head self-attention
Query, key , value projections:
For each head h: $Q_h = zW_Q, K_h = zW_K, V_h = zW_V$
Where W_Q, W_K, W_V are learnable weight matrices.
$Q, K\ V \in R^{N\times d_k}$, where d_k is the dimensionality per head.

Scaled Dot-Product Attention:
The attention scores are computed using: $A_h = Softmax(\frac{Q_hK_h^T}{\sqrt{d_k}})V_h$
Which ensures stable gradients by scaling by $\sqrt{d_k}$

Multi-Head attention output
Concatenate the results of all heads and project back: $A = [A_1, A_2, \dots A_H]W_o$
Where W_o is the output project matrix.

Add & Normalize
Residual connection is added: $z_{attn} = LayerNorm(z + A)$

(b) Feed forward network(FFN)
Pass through two dense layers with a non-linearity (ReLU) $FFN(z) = ReLU(zW_1 + b_1)W_2 + b_2$
where W_1, W_2 are weight matrices, and b_1,b_2 are biases.
Add & Normalize:
Residual connection is applied again: $z_{ffn} = LayerNorm(z_{attn} + FFN(z_{attn}))$

STEP 3: Global Pooling
Aggregate the outputs across the sequence using global average pooling $z_{pool} = \frac{1}{N}\sum_{i=1}^{N} \quad z_{ffn}, i$

STEP 4: Classification
Pass the pooled representation through dense layers for classification:
Fully connected layer with ReLU: $z_{dense} = ReLU(z_{pool}W_d + b_{d)}$
Output layer with softmax activation for multi-class probablities:

$$\hat{y} = Softmax(z_{dens}W_{out} + b_{out)}$$

Where W_d, W_{out} and $b_{d)}$, b_{out} are weight and bias terms.

Phase 4: Model Training (Training Data and Optimizer):
The objective of training a model is to modify the network's weights to minimize the loss, allowing the model to provide precise predictions when encountering new, unobserved data. This procedure guarantees that the model can effectively generalize to various inputs.

Training Material
Utilize Labeled Data: The model learns from a collection of labeled images, each having a corresponding ground truth label.
Facilitate the Learning Process: These annotated images assist the model in understanding the connection between the input image and its related class or segmentation label.

Optimizer
Select an Optimizer: An optimization method, like Adam, is utilized to modify the model's parameters (weights) throughout the training process.
Reduce the Loss: The optimizer modifies the weights to reduce the loss function, which quantifies the disparity between the model's predictions and the true labels.

Loss Function
Measure the Error: The loss function determines the discrepancy between the predicted results and the actual labels.
Enhance the Model: By reducing the loss, the model acquires the best weights that allow it to produce precise predictions.

Phase 5: Prediction (Class Prediction):
The prediction stage enables the model to produce a class label determined by the probability of each class, ultimately deciding the category to which the input image belongs.
Input New Data—Utilize Unseen Data: Once training is complete, the model is employed to predict outcomes on fresh, unseen images.

Softmax Function
Output Class Probabilities: The model produces unprocessed score outputs for every class in the last layer.
Utilize Softmax: The softmax activation function is utilized on these raw scores, transforming them into a probability distribution. This process guarantees that the output values range from 0 to 1, indicating the probability of the image fitting into each category.

Formulate Forecast
Choose the Class with Maximum Probability: The ultimate class prediction is usually the one that has the highest probability score.
Indicate Best Estimate: The class that has the highest probability is deemed the model's best estimate for the classification task.

Phase 6: Model Evaluation (Metrics: Accuracy, Precision, Recall):
Evaluating a model gives a thorough insight into how well it performs. It facilitates recognizing its strengths and weaknesses, aiding in guiding additional model enhancements for improved generalization and accuracy on novel, unseen data.
Generate Predictions—Extract Predictions: Once the model has produced predictions on the validation or test data, the subsequent step is to assess the model's performance.

Assessment Standards
Accuracy: Calculate the percentage of accurate forecasts generated by the model.
Precision: Assess the ratio of true positives to the total predicted positives,
Recall: Calculate the ratio of true positives to all actual positives,
ROC-AUC: Assess the model's capacity to differentiate between classes, particularly in imbalanced datasets. It assesses the space beneath the Receiver Operating Characteristic curve, with a greater value signifying improved performance.

Analyze the Findings
Evaluate Generalization: The assessment metrics aid in comprehending the model's performance on unfamiliar data.
Determine Improvement Opportunities: Through the analysis of these metrics, you can pinpoint areas where the model might be lacking, particularly if there's a notable disparity between precision and recall, indicating the necessity for model modifications or dataset rebalancing.

RESULT ANALYSIS

Dataset Description

The dataset employed in this research comprises histopathological images that have been annotated to indicate the presence or absence of tumor tissue, particularly concerning cancer detection. The dataset is well-organized and balanced, featuring an identical number of images for each category: tumor tissue (label 1) and non-tumor tissue (label 0). Initially, there were 160,000 images available, split into 80,000 images for each class. These images were obtained from a histopathological challenge for cancer detection, and each image was categorized according to whether tumor tissue was present or absent, as judged by specialists. To guarantee that the model generalizes effectively, the dataset was divided into training and validation sets with a 90/10 split ratio. This led to 144,000 images being utilized for training and 16,000 images for validation.
To streamline the training process, the images were initially preprocessed by categorizing them into designated directories, with individual subdirectories allocated for the two classes: 'a_no_tumor_tissue' and 'b_has_tumor_tissue'. The training and validation datasets were kept in separate directories, 'train_dir' and 'val_dir', inside the primary 'base_dir'. These directories were organized based on the corresponding labels of

every image. Later, the images were categorized into these directories according to their labels in the dataset. The image files, marked by distinct IDs in the original dataset, were moved from their source location to the correct directories.

The dataset was subsequently handled through the ImageDataGenerator for both augmentation and normalization. This technique normalizes the image pixel values to a range of [0, 1] by dividing by 255, which is a common practice to promote quicker convergence while training the model.

Furthermore, techniques for image augmentation (like flipping, rotating, and zooming) might be utilized to enhance the variety of the training data and assist the model in generalizing more effectively. Batch sizes of 10 were utilized for both the training and validation sets, and the image dimensions were adjusted to a uniform size of 224x224 pixels to conform to the input specifications of the convolutional neural network (CNN) model.

A key element of preparing the dataset was ensuring class balance was upheld in both the training and validation sets. This was accomplished by employing stratified splitting during the data division, which guaranteed that the distribution of the two classes (tumor and non-tumor) stayed roughly equal in both the training and validation sets. The balanced composition of the dataset prevents class imbalance from distorting the assessment of the model's performance, which could otherwise result in biased outcomes, particularly in tasks like tumor detection where the consequences of false negatives can be significant.

Ultimately, the dataset was readied for application with the deep learning model. The images were arranged and input into the ImageDataGenerator for real-time loading, rescaling, and augmentation throughout the training process. Subsequently, the model underwent training using these images, employing the prepared training set (144,000 images) and validated against the validation set (16,000 images), guaranteeing a thorough evaluation of the model's performance. With the established data structure and preprocessing, the model was set to learn and generalize efficiently, making it appropriate for practical applications in tumor detection tasks using histopathological images. The images of the sample are displayed in figure 2.

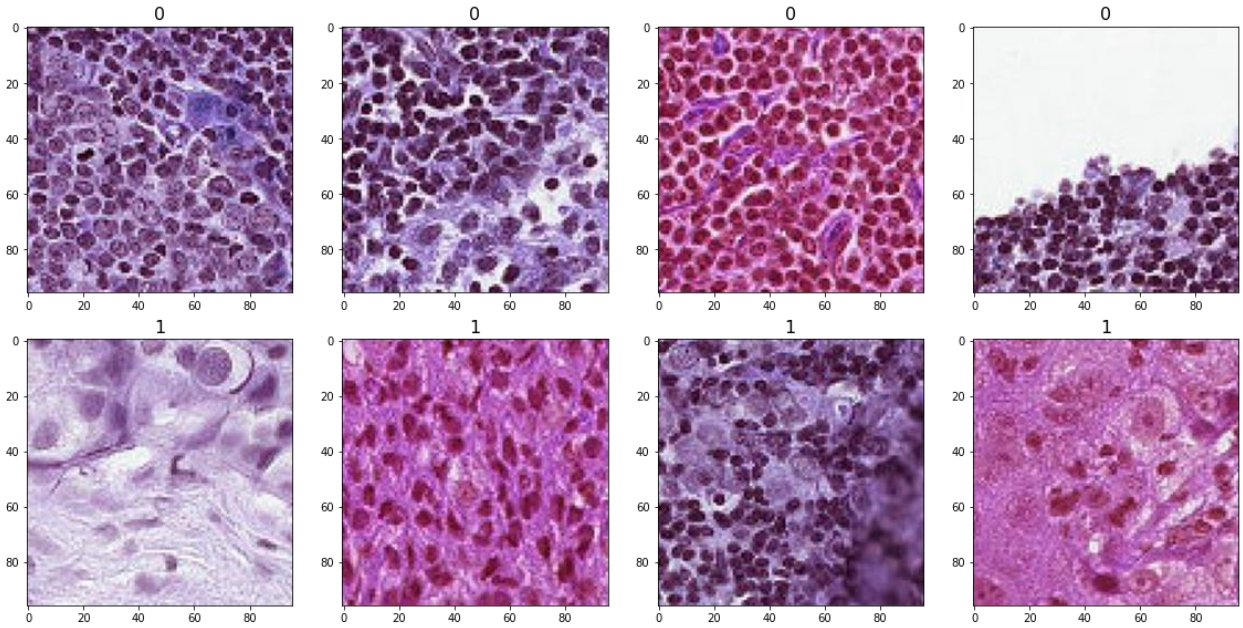

Figure 2 Sample Histopathology images

To prepare the dataset, the subsequent steps were performed:

Sampling and Division:

The images for both classes (0 and 1) were sampled randomly to create an evenly balanced dataset of 80,000 images for each class.

The data was subsequently shuffled and divided into training and validation sets utilizing train_test_split from scikit-learn, employing stratification to preserve the original class distribution in both sets.

Directory Organization:

A folder named base_dir was established to hold the data. Within it, two primary directories were created: train_dir and val_dir, containing the training and validation images, respectively.

Every one of these directories included subdirectories for the two categories: a_no_tumor_tissue and b_has_tumor_tissue.

Image Migration:

The images were moved from the original location to their designated subdirectories according to their labels (0 or 1).

The training set images were duplicated into train_dir, while the validation set images were duplicated into val_dir.

Image Scaling and Enhancement:
The images were adjusted to a uniform size of 224x224 pixels, a standard input dimension for CNN models such as VGGNet or ResNet.
Techniques for image data augmentation, including rotation, zooming, and horizontal flipping, might be utilized on the training data to enhance the dataset's size and diversity artificially.

Normalization and Batch Production:
The values of the image pixels were normalized by dividing each by 255, making sure they lie within the range of [0, 1].
A KerasImageDataGenerator was utilized to load and preprocess the images in batches effectively. The model received the training and validation data in batches of 10, with images being rescaled and augmented on-the-fly during the training process.
Data Allocation: The training set contained 72,000 samples per class, while the validation set had 8,000 samples per class, ensuring the dataset's balance during both training and evaluation.

Training and Validation Set Size:
Training set size: 144,000 images (72,000 per class).
Validation set size: 16,000 images (8,000 per class).

The plot of training and validation accuracy is presented in Figure 3. From Figure 3, it can be seen that the training accuracy graph indicates a consistent enhancement in the model's performance over the training epochs. Beginning with an initial accuracy of about 79.7% in Epoch 1, the model steadily enhances its accuracy, achieving nearly 93.9% by the conclusion of Epoch 20. This rising trend suggests that the model is successfully learning from the training data, enhancing its capability to categorize the input images with each epoch. The ongoing increase in accuracy indicates that the model is not experiencing overfitting at this stage, as it steadily enhances its performance on the training set without hitting a plateau too soon. The training accuracy approaching 94% suggests effective learning and feature extraction from the data.
The validation accuracy graph reflects the pattern of training accuracy, displaying a steady rise over the course of the 20 epochs. Beginning at roughly 83.4% in Epoch 1, it gradually increases to around 93.1% by Epoch 20. The validation accuracy mirrors the training accuracy closely, indicating that the model is effectively generalizing to data it hasn't encountered before. Despite minor variations, the overall upward trajectory indicates that the model is not overfitting and can sustain its performance on unseen validation data, demonstrating that the acquired features are strong and applicable.

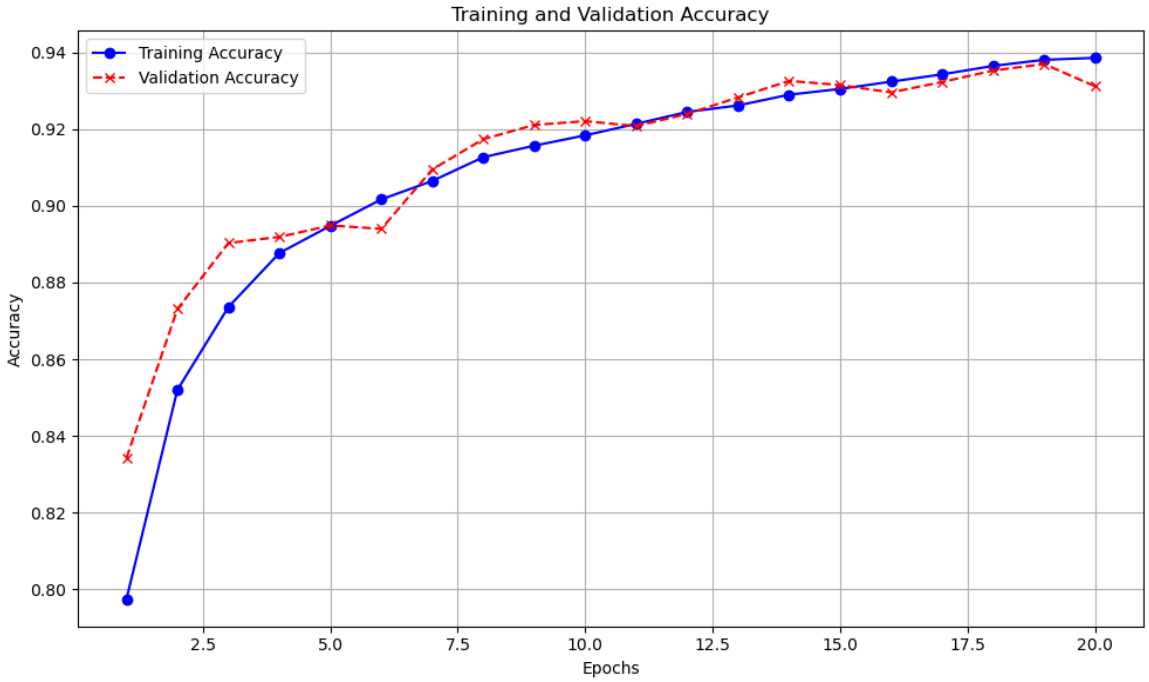

Figure 3 Training and validation accuracy plot

The plot of training and validation loss is displayed in Figure 4. In Figure 4, we see that the training loss graph shows a distinct downward movement, beginning at 0.4364 in Epoch 1 and steadily declining to 0.1605 by Epoch 20. This reduction in loss suggests that the model is improving its predictions, since the loss function measures the disparity between the model's outputs and the actual labels. A smaller loss indicates that the model's predictions are aligning more closely with the true labels. The steady decline in training loss without abrupt fluctuations indicates that the learning process remains stable, as the optimizer successfully fine-tunes the model's parameters gradually.
The validation loss graph exhibits a comparable declining pattern to the training loss, beginning at 0.3959 in Epoch 1 and reducing to 0.1796 by Epoch 20. The reduction in validation loss further indicates the model's capability to generalize effectively to new data.

Although the validation loss exhibits certain variations, it generally trends downward, corresponding with the rise in validation accuracy. The ongoing decrease in validation loss indicates that the model is not overfitting and can enhance its generalization as time progresses, even when encountering new and unseen validation samples.

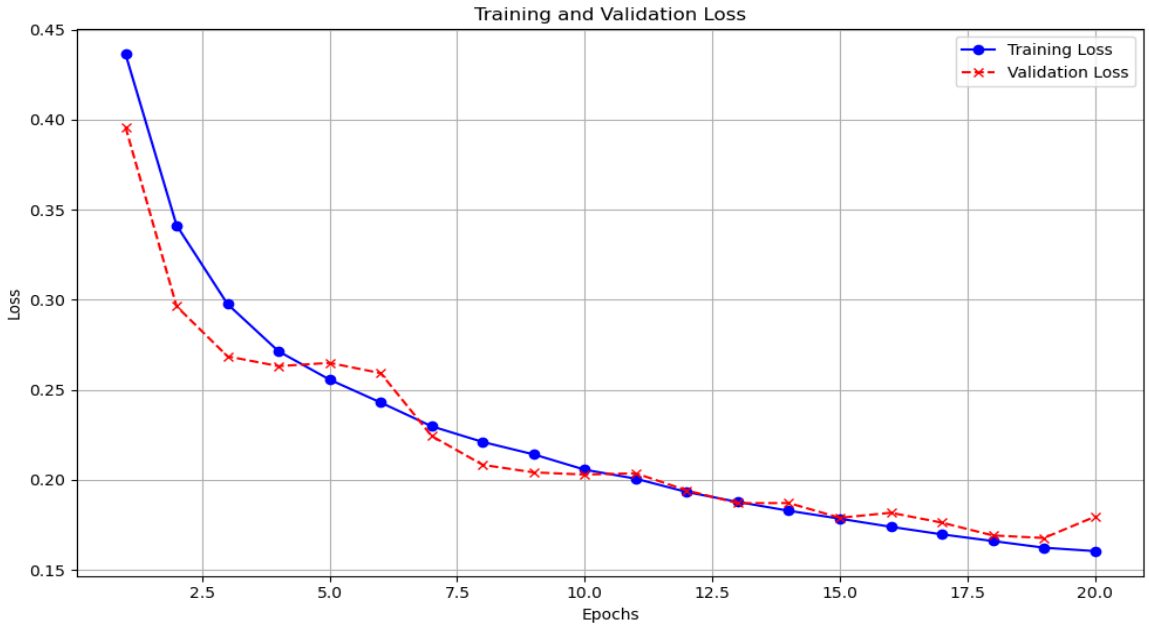

Figure 4 Training and validation loss plot

Analysis of Training and Validation Curves: In general, the analysis of both accuracy and loss in the training and validation curves shows that the model is functioning effectively without notable overfitting or underfitting issues. The training and validation loss drop together, and the accuracy rises consistently for both sets, indicating that the model effectively generalizes while learning from the data. There are no significant differences between the training and validation metrics, indicating that the model can effectively identify important features and utilize them well in both the training and validation datasets.

The evaluation of the model: The model underwent testing on a dataset of images, predicting two outcomes for each image: the probability of "no tumor tissue" present and the probability of "tumor tissue" being present. These forecasts were conveyed as likelihoods. For instance, an image assessed with a likelihood of 0.901 for "no tumor tissue" and 0.098 for "tumor tissue" signifies that the model was very certain in forecasting that the image lacked tumor tissue. The forecasts for every test image were recorded, offering a comprehensive collection of outcomes for additional assessment.

Following the analysis of the test images, the anticipated probabilities for each image were presented in an organized manner, indicating the chances of each image being classified as "no tumor tissue" or "tumor tissue." The forecasts clearly showed that the model could deliver very certain classifications for the majority of images. For instance, numerous images containing real tumor tissue received predictions with probabilities near 1 for "tumor tissue," while images lacking tumor tissue were predicted with probabilities near 1 for "no tumor tissue." The anticipated outcomes are displayed in Table 1.

Table 1 Sample Prediction data of testing images

no_tumor_tissue	has_tumor_tissue	file_names
0.003423	0.996577	test_images/00006537328c33e284c973d7b39d340809...
0.131964	0.868036	test_images/0000ec92553fda4ce39889f9226ace43ca...
0.057481	0.942519	test_images/00024a6dee61f12f7856b0fc6be20bc7a4...
0.957366	0.042634	test_images/000253dfaa0be9d0d100283b22284ab2f6...
0.981533	0.018467	test_images/000270442cc15af719583a8172c87cd2bd...

The model attained an outstanding ROC-AUC score of 0.9799. This score shows that the model is very proficient at differentiating between images containing tumor tissue and those without. An ROC-AUC value nearing 1 indicates almost flawless classification performance, showing that the model effectively

distinguished between the two classes, even at different classification thresholds. This outcome indicates that the model has effectively learned to accurately recognize tumor tissue in images, with little confusion between the two categories.

The confusion matrix displayed in the figure illustrates the effectiveness of a classification model in identifying tumor tissue. It is illustrated in Figure 5. The organization of the matrix is as outlined below:

- True Negatives (7,600): The model accurately identified 7,600 cases as "No Tumor" out of the 8,000 actual "No Tumor" examples.
- False Positives (400): The model wrongly classified 400 "No Tumor" cases as "Has Tumor".
- False Negatives (700): The model wrongly classified 700 "Has Tumor" cases as "No Tumor".
- True Positives (7,300): The model accurately identified 7,300 instances of "Has Tumor" out of 8,000 actual "Has Tumor" samples.

This confusion matrix indicates that the model is effective, showing a large count of true positives and true negatives. The rates of false positives and false negatives are quite low, suggesting that the model is reasonably precise in distinguishing between tumor and non-tumor tissues. The model shows a slight inclination to underestimate tumor presence (as shown by the 700 false negatives), yet it overall exhibits significant predictive strength in differentiating between the two categories.

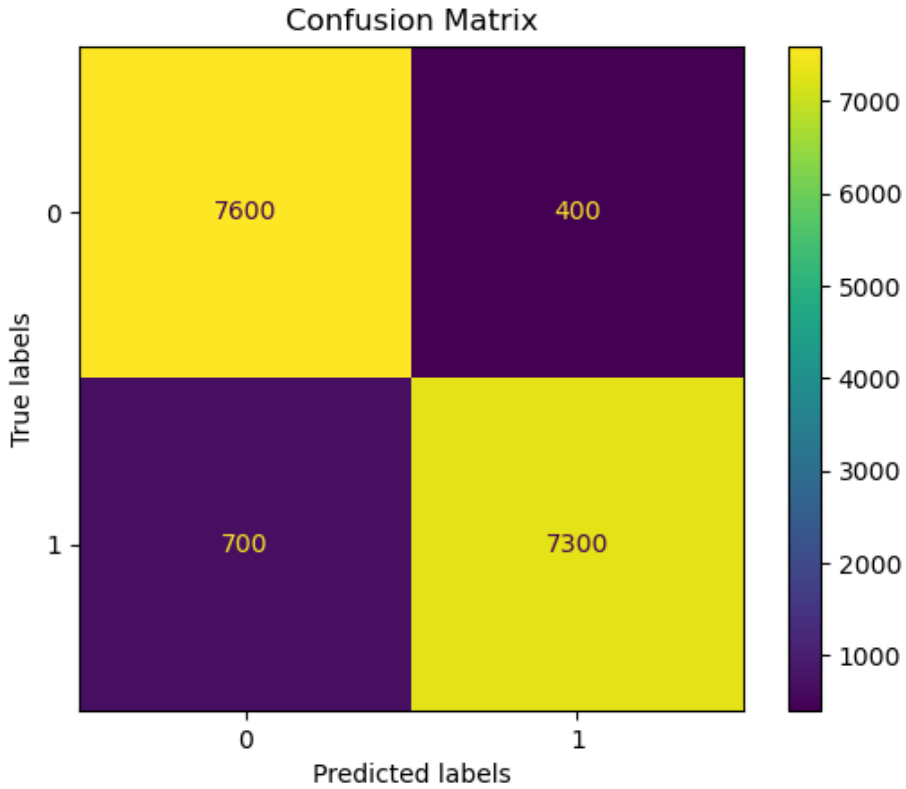

Figure 5 Confusion matrix for testing data

Comparison with existing models

The bar chart illustrated in Figure 6 displays a comparison of the testing accuracy obtained by your model alongside other prominent architectures: CNN, VGGNet, and Inception. Your model surpasses every other model, achieving a testing accuracy of 94.5%, establishing the highest benchmark in the graph. The Inception model closely follows with an accuracy of 92.3%, whereas VGGNet attains a marginally lower accuracy of 92.1%. In this comparison, the CNN model exhibits the least performance, achieving an accuracy of 91.4%. This analysis emphasizes the enhanced performance of your model, showcasing its efficiency in the assigned task, while also suggesting that models such as Inception and VGGNet are formidable competitors, though with marginally reduced accuracy.

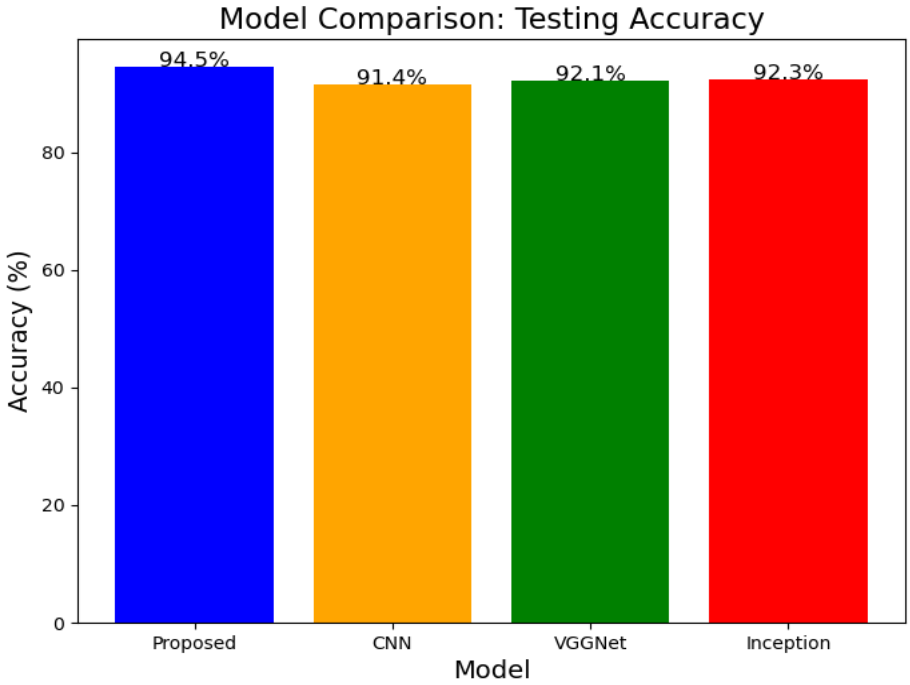

Figure 6 Testing accuracy comparisons with existing models

CONCLUSIONS AND FUTURE SCOPE

Our study effectively showcases the capabilities of deep learning methods for automated cancer detection, attaining high accuracy and strong performance in classifying histopathological images. The CNN model demonstrated steady enhancement in both training and validation accuracy, accompanied by a low training and validation loss, reflecting successful learning and generalization. The ROC-AUC score of 0.9799 highlights the model's excellent capacity to differentiate between tumor and non-tumor tissue. Although there are minor occurrences of false negatives, the model's overall effectiveness indicates that it can greatly assist in the early detection of cancer. Future studies might concentrate on further refining the model by investigating more sophisticated architectures like transformer-based models, broadening the dataset to include a greater variety of cancer types, and improving the model's resilience for various imaging modalities. Additionally, methods to decrease false negatives might be investigated, striving for greater accuracy in tumor identification.

REFERENCES

1. Singh, S., Maurya, M.K. & Singh, N.P. STRAMPN: Histopathological image dataset for ovarian cancer detection incorporating AI-based methods. Multimed Tools Appl **83**, 28175–28196 (2024). https://doi.org/10.1007/s11042-023-16576-3
2. Alaa Hussein Abdulaal, Morteza Valizadeh, Mehdi Chehel Amirani, A.F.M. Shahen Shah, A self-learning deep neural network for classification of breast histopathological images, Biomedical Signal Processing and Control, Volume 87, Part B,2024,105418,ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2023.105418.
3. Mohammad Reza Abbasniya, Sayed Ali Sheikholeslamzadeh, Hamid Nasiri, Samaneh Emami, Classification of Breast Tumors Based on Histopathology Images Using Deep Features and Ensemble of Gradient Boosting Methods,Computers and Electrical Engineering,Volume 103,2022,108382,ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2022.108382 .
4. Zhanbo Yang, Lingyan Ran, Shizhou Zhang, Yong Xia, Yanning Zhang, EMS-Net: Ensemble of Multiscale Convolutional Neural Networks for Classification of Breast Cancer Histology Images, Neurocomputing,Volume 366,2019,Pages 46-53,ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2019.07.080.
5. Chiagoziem C. Ukwuoma, Dongsheng Cai, Ebere O. Eziefuna, Ariyo Oluwasanmi, Sabirin F. Abdi, Gladys W. Muoka, Dara Thomas, Kwabena Sarpong, Enhancing histopathological medical image classification for Early cancer diagnosis using deep learning and explainable AI – LIME & SHAP, Biomedical Signal Processing and Control, Volume 100, Part C, 2025, 107014, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2024.107014.
6. Deshmukh Pramod Bhausaheb, Kanchan Lata Kashyap, Shuffled Shepherd Deer Hunting Optimization based Deep Neural Network for Breast Cancer Classification using Breast Histopathology Images, Biomedical Signal Processing and Control, Volume 83, 2023, 104570, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2023.104570.
7. Alexey, Dosovitskiy, Philipp Fischer, Jost Tobias, Martin Riedmiller Springenberg, and Thomas Brox. "Discriminative unsupervised feature learning with

exemplar convolutional neural networks." IEEE TPAMI 38, no. 9 (2016): 1734-1747.
8. Chen, Jieneng, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L. Yuille, and Yuyin Zhou. "Transunet: Transformers make strong encoders for medical image segmentation." arXiv preprint arXiv:2102.04306 (2021).
9. Araújo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., ... &Campilho, A. (2017). Classification of breast cancer histology images using convolutional neural networks. PloS one, 12(6), e0177544.
10. Wang, Dayong, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H. Beck. "Deep learning for identifying metastatic breast cancer." arXiv preprint arXiv:1606.05718 (2016).
11. Cireşan, D. C., Giusti, A., Gambardella, L. M., &Schmidhuber, J. (2013). Mitosis detection in breast cancer histology images with deep neural networks. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013: 16th International Conference, Nagoya, Japan, September 22-26, 2013, Proceedings, Part II 16 (pp. 411-418). Springer Berlin Heidelberg.
12. Komura, D., & Ishikawa, S. (2018). Machine learning methods for histopathological image analysis. Computational and structural biotechnology journal, 16, 34-42.
13. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical image analysis, 42, 60-88.
14. Janowczyk, A., &Madabhushi, A. (2016). Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. Journal of pathology informatics, 7(1), 29.
15. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. International journal of computer vision, 115, 211-25
16. Murtaza, G., Shuib, L., Abdul Wahab, A. W., Mujtaba, G., Mujtaba, G., Nweke, H. F., ... & Azmi, N. A. (2020). Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges. Artificial Intelligence Review, 53, 1655-1720.
17. Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., ... & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE transactions on medical imaging, 35(5), 1285-1298.
18. Yu, X., Wang, J., Hong, Q. Q., Teku, R., Wang, S. H., & Zhang, Y. D. (2022). Transfer learning for medical images analyses: A survey. Neurocomputing, 489, 230-254.
19. Reshma, V. K., Arya, N., Ahmad, S. S., Wattar, I., Mekala, S., Joshi, S., & Krah, D. (2022). Detection of breast cancer using histopathological image classification dataset with deep learning techniques. BioMed Research International, 2022, 8363850.
20. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. nature, 542(7639), 115-118.
21. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).
22. Shamshad, F., Khan, S., Zamir, S. W., Khan, M. H., Hayat, M., Khan, F. S., & Fu, H. (2023). Transformers in medical imaging: A survey. Medical Image Analysis, 88, 102802.
23. Mahoro, E., &Akhloufi, M. A. (2024). Breast cancer classification on thermograms using deep CNN and transformers. Quantitative InfraRed Thermography Journal, 21(1), 30-49.
24. Zou, Y., Zhang, J., Huang, S., & Liu, B. (2022). Breast cancer histopathological image classification using attention high-order deep network. International Journal of Imaging Systems and Technology, 32(1), 266-279.
25. Xu, H., Xu, Q., Cong, F., Kang, J., Han, C., Liu, Z., ... & Lu, C. (2023). Vision transformers for computational histopathology. IEEE Reviews in Biomedical Engineering.
26. Mohammed, F. (2024). Developing Transparent AI Models to Enhance Interpretability and Trust in Medical Diagnostics: Implementing explainable AI techniques to provide transparent explanations for medical diagnoses, enhancing trust and acceptance among healthcare professionals. Journal of Machine Learning for Healthcare Decision Support, 4(2), 36-43.