Abstract

This project explores the application of deep learning techniques for automated detection of cancer in histopathological images, an essential yet complex task due to the intricate patterns and high variability in these medical images. Two models were developed and evaluated: a custom-built Convolutional Neural Network (CNN) and DenseNet-201, a pre-trained architecture renowned for its robust performance in image classification tasks.

The custom CNN was designed to capture localized features effectively while maintaining simplicity, making it suitable for resource-constrained environments. DenseNet-201, on the other hand, leverages densely connected layers to optimize feature propagation and address challenges like the vanishing gradient problem.

Through rigorous evaluation using metrics such as accuracy, precision, and confusion matrices, the project aims to assess the models' ability to classify tissue samples as benign or malignant. This study demonstrates the potential of deep learning to significantly enhance diagnostic accuracy, reduce human error, and support pathologists in making timely and informed decisions.

Introduction

Lymph node metastases detection is a critical factor in cancer treatment and prognosis, particularly in diseases like breast cancer, where the presence and extent of metastases guide therapeutic decisions. Traditional pathology workflows rely on manually examining histological slides, a time-intensive and error-prone process that can miss small or subtle metastases. Recent advances in the field of digital pathology and microscopic imaging hardware have allowed the advent of digitizing glass slides into whole-slide images (WSIs).

Convolutional neural networks (CNNs) offer a transformative approach by automating the analysis of these whole-slide images with remarkable speed and accuracy. These deep-learning models excel at capturing complex patterns in high-dimensional data, enabling them to detect metastases that may elude the human eye. By leveraging large datasets, CNNs can generalize across diverse image appearances and staining variations, ensuring robustness in real-world applications. The use of CNNs reduces the workload for pathologists and enhances diagnostic consistency and precision, ultimately improving patient outcomes through timely and accurate detection of metastatic cancer.

The PatchCameylon (PCam) dataset utilized in our project is a subset of the CAMELYON dataset, consisting of 1,399 annotated whole-slide images of lymph nodes with and without metastases [6]. These slides were sourced from five different medical centers to ensure diversity

in image appearance and staining variations. Each image includes a slide-level label indicating the presence of no metastases, macro-metastases, micro-metastases, or isolated tumor cells.

The PCam dataset contains 327,680 color images extracted from the histopathologic scans of lymph node sections in the CAMEYLON dataset [5]. Images in this dataset contain annotations with a binary label depicting the presence of metastatic tissue. In other words, if the sample associated with the whole-slide image is benign or malignant. An added and significant benefit of this dataset is that it can be used to train models on a single GPU, which is ideal for our project due to the constraints of limited computational resources.

Background

In medical science, patients will undergo clinical trials, followed by histopathological analysis of cancer sites to make a first level diagnosis. The traditional approach can be time-consuming and not error free particularly in the identification of metastatic cancer in small image patches from the bigger pathology scans. So, the complexity of this task makes the automated process of classification of histopathological images a challenging task. Computer-aided diagnosis can ease this process and may be more reliable and efficient. The advancement in machine learning techniques can be utilized to facilitate the classification process, thereby reducing human burden and improving diagnostic accuracy.

The application of CNNs in digital pathology has been transformative, mainly in the task of tumor detection and classification in whole-slide images (WSIs). Traditional CNNs, which were developed for natural image processing, have been adapted to handle the challenges in histopathology images. These images exhibit not only translational symmetry but also rotation and reflection symmetries, which standard CNNs do not inherently exploit. This limitation often results in inefficiencies, as models must learn multiple orientations of the same features, leading to increased computational demands and potential overfitting. To address these challenges, recent research has focused on developing rotation-equivariant CNNs that can leverage these additional symmetries to improve model robustness and generalization [2].

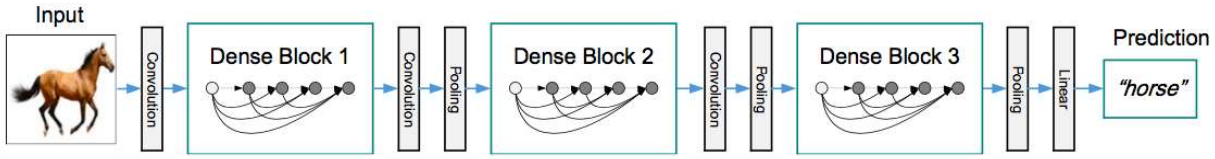


Figure 1: A deep DenseNet with three dense blocks. The layers between two adjacent blocks are referred to as transition layers and change feature-map sizes via convolution and pooling.

The DenseNet model has become a powerful architecture in medical image analysis, especially tasks involving the detection and classification of cancer. The unique characteristic of DenseNet is its dense connectivity pattern, which promotes feature reuse across the network and improves feature propagation by allowing each layer to receive inputs from all previous layers. This architecture also addresses the problem of vanishing gradient than traditional CNNs, which is very useful in histopathological images. According to studies [1], DenseNet model performs better in classification accuracy and computational efficiency compared to alternate models, which makes them suitable for real-time diagnostic applications in digital pathology. Thus, we have utilized this DenseNet model in our project to classify the cancer and non-cancer cells in histological images.

Approach

The task of classifying histopathological images into benign and malignant categories was approached using two complementary deep learning models: a custom-designed Convolutional Neural Network (CNN) and the pre-trained DenseNet201 model. These models were selected to explore the balance between task-specific design simplicity and the power of transfer learning.

Dataset

The dataset used in this study consists of histopathological images, each representing a small patch of tissue. The images are categorized into two classes: benign and malignant, forming a binary classification task. The dataset includes high-resolution images where each image is labeled based on the presence of tumor tissue. A positive label indicates the presence of at least one pixel of tumor tissue in the center of the image. This dataset was selected due to its relevance in medical image classification, particularly in cancer diagnosis. Since the dataset comprises histopathological images that vary in size and detail, the following preprocessing steps were applied to standardize the data and prepare it for model training:

Image Resizing: Images were resized to a standard resolution of 224x224 pixels to match the input requirements of both the CNN and DenseNet201 models. This resolution was selected to ensure consistency and compatibility with the input specifications of the models.

Normalization: Pixel values were scaled to the range of [0, 1], which improved the training efficiency by standardizing the data distribution. This normalization ensured that the input channels had similar properties, preventing any single channel from dominating the learning process.

Data Augmentation: To enhance model generalization and reduce the risk of overfitting, data augmentation techniques were applied. These techniques included random rotations, zooming, and horizontal flipping, allowing the model to learn invariant features from the images.

Methodology

Custom CNN

A custom-designed Convolutional Neural Network (CNN) was developed to explore a simpler, task-specific architecture for histopathological image classification. The architecture of the model was designed as follows:

Convolutional Layers: The model consists of three convolutional layers with filter sizes of 32, 64, and 128, respectively. Each convolutional layer is followed by batch normalization and ReLU activation to introduce non-linearity and stabilize training.

Max-Pooling Layers: Max pooling is applied after each convolutional layer to down-sample the spatial dimensions of the feature maps and reduce computational complexity.

Fully Connected Layers: The output from the convolutional layers is flattened and passed through fully connected layers with 512, 128, and 2 nodes. ReLU activation is applied to these layers, and dropout is used for regularization to prevent overfitting.

Loss Function: Cross-entropy loss is used, as it is well-suited for binary classification tasks.

Optimizer: The Adam optimizer with a learning rate of 0.0002 is used. The optimizer dynamically adjusts the learning rate during training to improve convergence.

Learning Rate Scheduler: ReduceLROnPlateau is employed to reduce the learning rate when the validation performance plateaus, ensuring efficient training.

DenseNet201

DenseNet201, a pre-trained deep learning model, was selected for comparison due to its efficient feature reuse and transfer learning capabilities. The architecture of the model is as follows:

Initial Convolution Layer: A 7x7 convolution with 64 filters and a stride of 2 is followed by batch normalization, ReLU activation, and a 3x3 max pooling layer for initial feature extraction and down-sampling.

Dense Blocks: The model consists of four dense blocks, containing 6, 12, 48, and 32 layers, respectively. Each layer in the block is connected to all previous layers, enabling efficient feature reuse. Each layer applies batch normalization, ReLU activation, and a 3x3 convolution. The mathematical formulation for this is:

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}])$$

where x_l represents the output of layer l_i and H_l is a composite function of operations (ex. Batch normalization, ReLU, convolution, etc.).

Transition Layers: Transition layers between dense blocks help reduce the number of feature maps. These layers contain a 1x1 convolution for dimensionality reduction, followed by 2x2 average pooling for further down-sampling.

Global Average Pooling: After the final dense block, a global average pooling layer reduces the spatial dimensions to a single value per feature map.

Fully Connected Layer: The original fully connected layer with 1,000 nodes is replaced with a custom two-node output layer for binary classification (benign/malignant) using softmax activation.

Model Training

Both models were trained on the same dataset split, with 80% allocated for training and 20% for validation. The training process for both models was conducted using the Adam optimizer, which adapts the learning rate dynamically based on the gradients' first and second moments. The initial learning rate for both models was set to 0.0002, and the learning rate was adjusted using the ReduceLROnPlateau scheduler, which reduces the learning rate when the validation performance plateaus.

For the training process, the models were trained for three epochs to allow sufficient time for convergence. Cross-entropy loss was used as the objective function for binary classification. During training, mixed precision training was employed to speed up computations on the GPU, which significantly reduced training time without sacrificing accuracy.

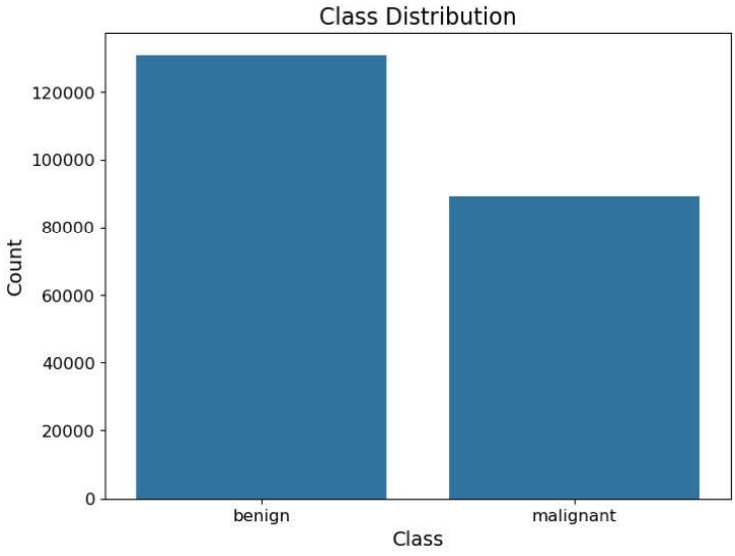
Throughout the training process, several evaluation metrics were tracked, including training loss, validation loss, accuracy, precision, recall, and confusion matrices. These metrics provided valuable insights into the performance of both models, helping assess how well the models generalized to unseen data and how they handled the class imbalance present in the dataset.

Results

Data Overview

The dataset used for this project consists of histopathologic cancer images categorized into two classes: benign and malignant. The images were collected in .tif format.

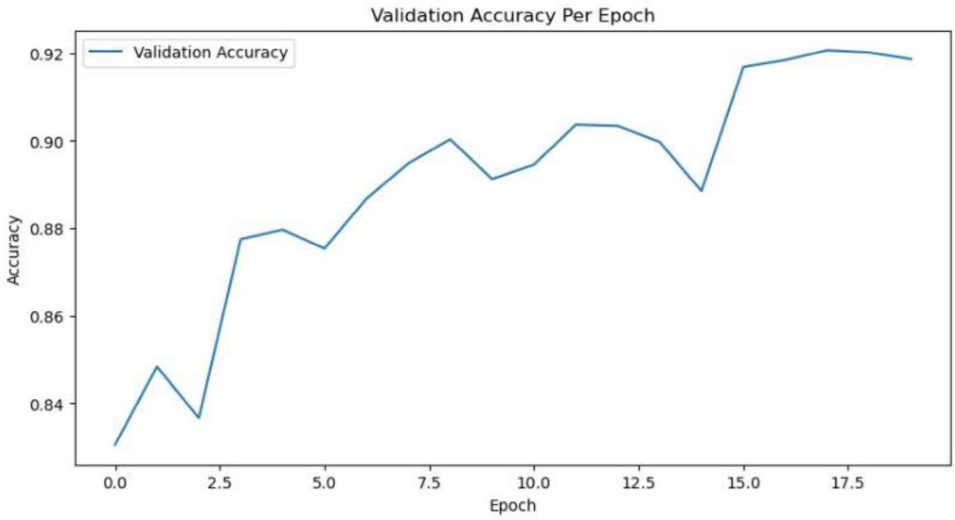
The dataset exhibited a slight imbalance, with more benign samples than malignant ones but the ratio does not affect the model in prediction.



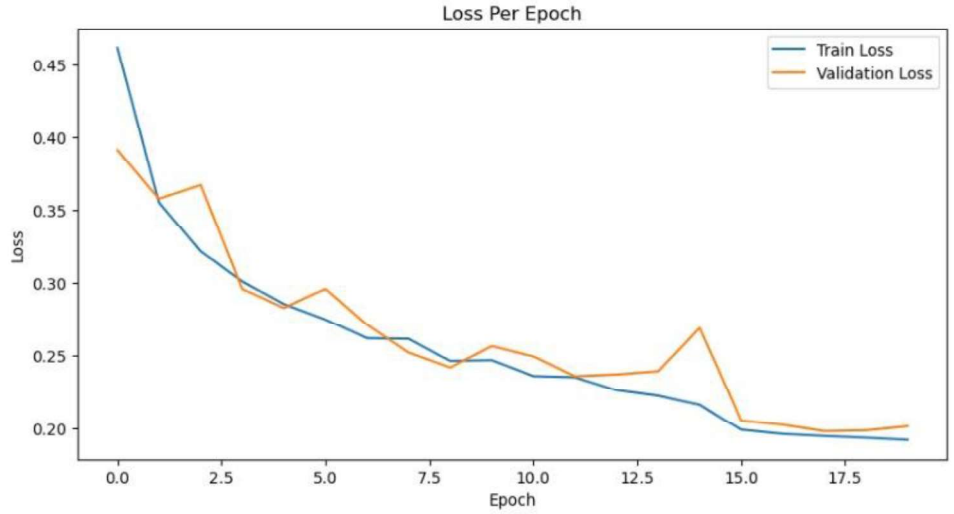
Model Performance

Custom CNN

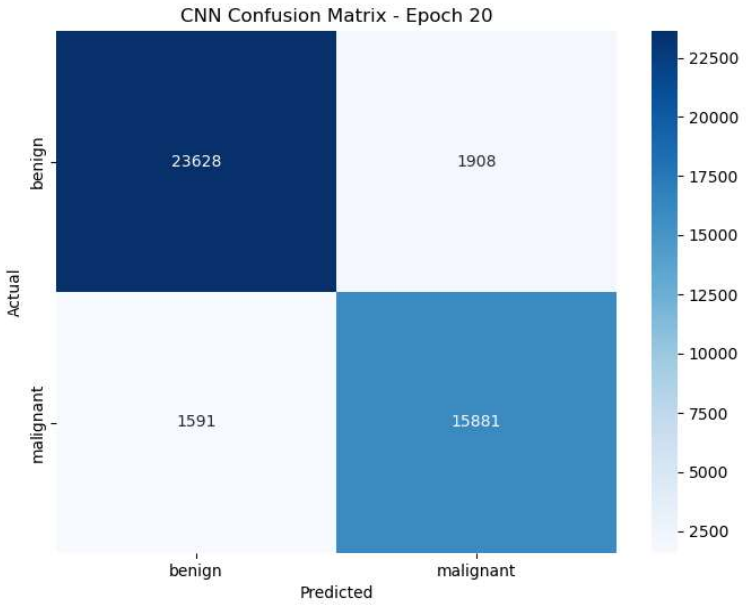
The custom CNN model achieved validation accuracy of 91.9 %, which plateaued after 15 epochs, suggesting the model struggles to generalize effectively.



Training Loss: The training loss decreases steadily, while the validation loss starts to increase after a certain point, indicating overfitting. Despite this, validation accuracy continues to improve, peaking around 0.90, showing the model's strong ability to classify unseen data initially.

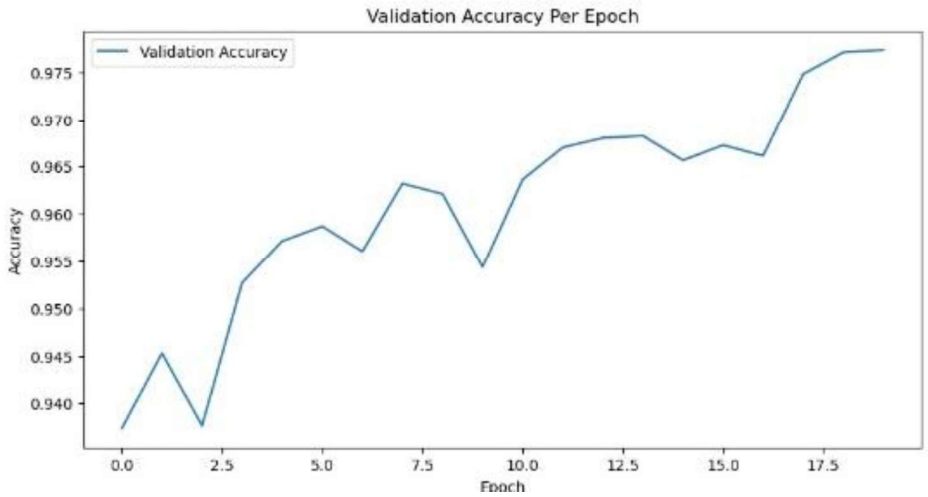


Confusion Matrix: The confusion matrix shows that the model accurately predicts benign and malignant images, with most correct classifications (23,628 benign, 15,881 malignant). However, it misclassifies some images, resulting in 1,908 false positives and 1,591 false negatives.

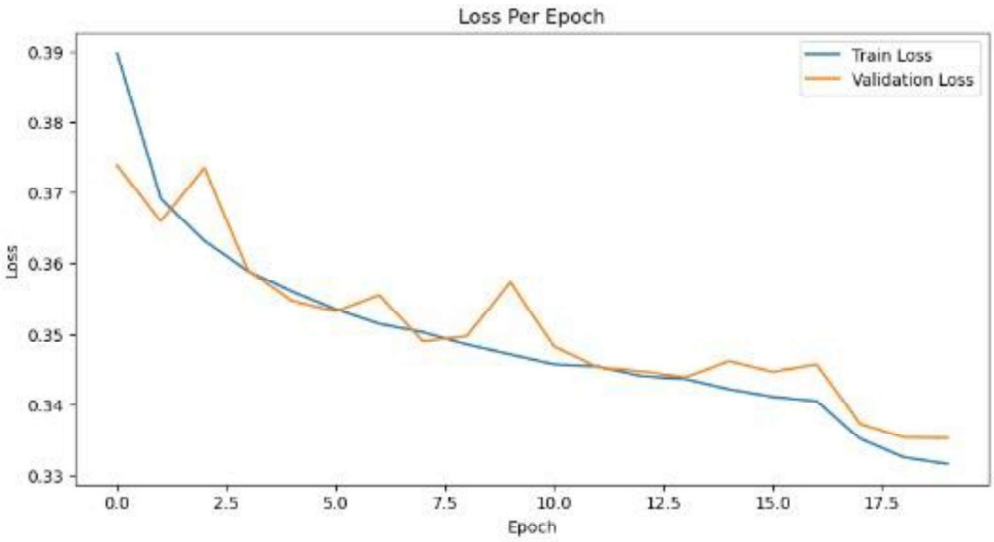


DenseNet201

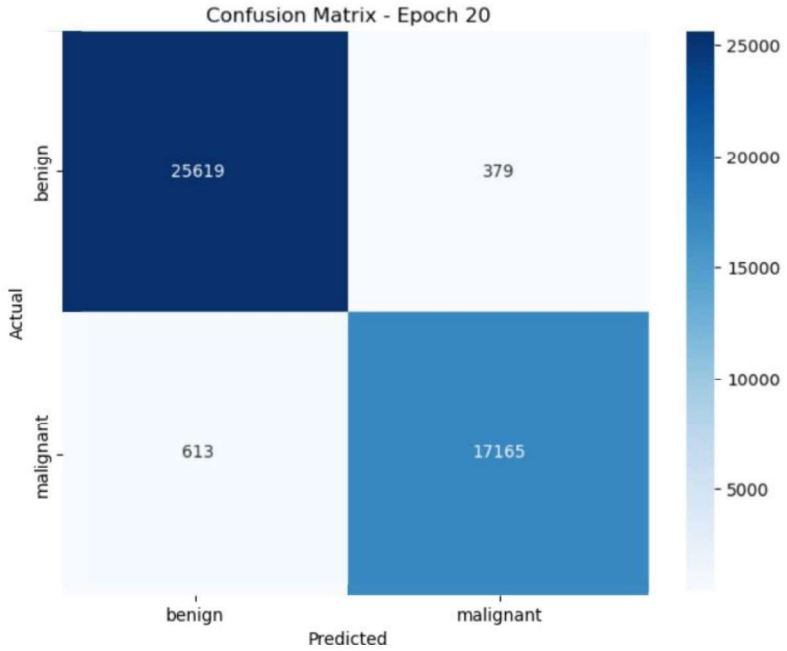
DenseNet201, utilizing pre-trained weights, outperformed our custom CNN in all aspects. Achieved an impressive 97.3% validation accuracy, benefiting from transfer learning and faster convergence.

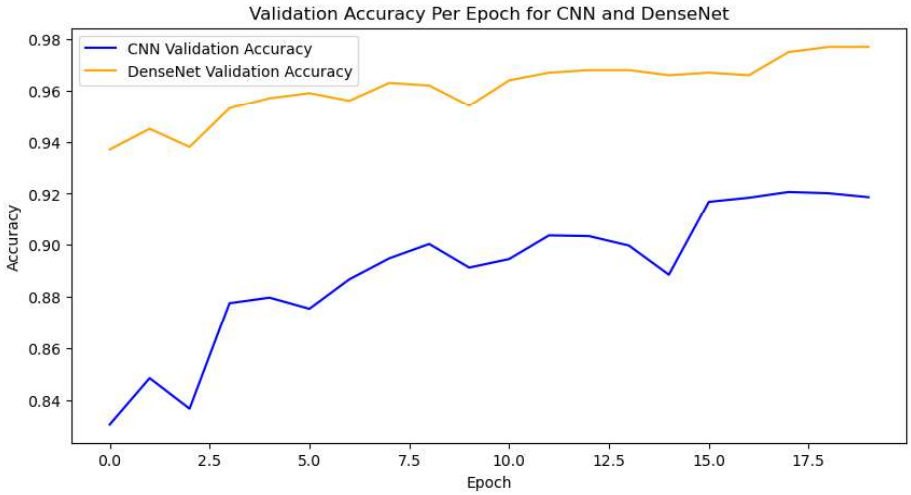


Training Loss: Both training and validation losses decreased steadily, reaching a minimum of 0.33, indicating efficient learning and minimal overfitting.



Confusion Matrix: DenseNet201 demonstrated higher precision and recall for malignant images, significantly reducing false negatives and improving the model's diagnostic reliability.





Key Findings

- **DenseNet201's Advantage:** Utilizing transfer learning, DenseNet201 outperformed our custom CNN by a significant margin, achieving a validation accuracy of 97.3%. Its superior ability to generalize from pre-trained features led to faster convergence and much better performance on the validation set, making it highly effective for medical image classification tasks.
- **Custom CNN Limitations:** Although the custom CNN achieved a reasonable baseline with 80% validation accuracy, it struggled with false negatives, especially in the malignant class. This limitation highlights the need for more complex architectures in applications like cancer detection, where accurate classification is critical.
- **Class Imbalance:** DenseNet201 demonstrated a better ability to handle class imbalance, likely due to its deeper feature extraction capabilities. The custom CNN, while providing decent performance, could benefit from additional regularization techniques or a more sophisticated architecture to further improve its performance and address the class imbalance effectively.

Future Directions

Improving Custom CNN: Future work could focus on fine-tuning the custom CNN architecture with additional layers or dropout to enhance performance.

Addressing Class Imbalance: Further experiments can incorporate techniques such as oversampling or class-weight adjustments to handle class imbalance more effectively.

Ensemble Methods: Combining the strengths of both models using ensemble learning could enhance accuracy and reliability, especially for complex datasets like histopathological images.

Conclusion

This project demonstrated the potential of deep learning, specifically Convolutional Neural Networks (CNNs), in automating the detection of cancer in histopathological images. Through the evaluation of both a custom-built CNN and the pre-trained DenseNet201 model, we observed that DenseNet201 outperformed the custom model in terms of accuracy and robustness, owing to its ability to leverage transfer learning. DenseNet201 achieved an impressive validation accuracy of 97.3%, which highlights its superior generalization capabilities, particularly in complex tasks such as cancer detection, where accuracy is crucial.

While the custom CNN provided a reasonable baseline, its performance was limited by challenges related to generalization and class imbalance, emphasizing the need for more sophisticated architectures and strategies to handle such issues. The results suggest that DenseNet201's deep feature extraction and efficient information flow across layers made it more adept at managing the intricate variations found in histopathological images.

This study highlights the importance of advanced machine learning in medical diagnostics, particularly for improving cancer detection efficiency and accuracy. The findings strengthen the case for integrating deep learning in pathology to enhance workflows, reduce errors, and support informed decisions by pathologists. Future work could refine the custom CNN, address class imbalance, and explore ensemble methods, enabling even more effective cancer diagnosis solutions.

References

[1] Gao Huang, Zhuang Liu, Laurens van der Maaten and Kilian Q. Weinberger. Densely Connected Convolutional Networks in CVPR, 2017

[2] Bastiaan S. Veeling, Jasper Linmans, Jim Winkens, Taco Cohen and Max Welling. Rotation Equivariant CNNs for Digital Pathology ,2018

[3] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer ence, 2014

[4] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck. Deep learning for identifying metastatic breast cancer, 2016

[5] Ehteshami Bejnordi et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. JAMA: The Journal of the American Medical Association, 318(22), 2199–2210. **doi:jama.2017.14585**

[6] Alexi, B., Altuna, H., Babak, B. E., Wauters Carla, Geert, L., Jeroen, L. V., Dijk Van Marcory, Maschenka, B., Meyke, H., Nikolas, S., Oscar, G., Paul, D. V., Peter, B., Bult Peter, Manson Quirine, Vogels Rob, & Rob, D. L. V. (2018). *Supporting data for "1399 H&E-stained sentinel lymph node sections of breast cancer patients: the CAMELYON dataset"* [Data set]. GigaScience Database. https://doi.org/10.5524/100439