Evolving Diagnostics: Incremental and Transfer Learning in Histopathology Image Analysis for Cancer Detection

ANKUR YADAV, The University of Texas at Dallas, USA

ACM Reference Format:
Ankur Yadav. 2024. Evolving Diagnostics: Incremental and Transfer Learning in Histopathology Image Analysis for Cancer Detection. 1, 1 (May 2024), 13 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In the realm of medical diagnostics, utilizing histopathology images for cancer detection is a critical frontier where precision and accuracy significantly impact patient outcomes. Histopathology, the examination of biological tissues to observe diseased cells and tissues at a microscopic level, is essential for identifying and classifying various forms of cancer. The complexity of these images, coupled with the subtlety of pathological variations among different cancers, presents substantial challenges. Traditional diagnostic methods depend heavily on the expertise of pathologists, leading to potential subjectivity and variability in diagnoses.

To overcome these challenges, this project, "Evolving Diagnostics: Incremental and Transfer Learning in Histopathology Image Analysis for Cancer Detection," leverages advanced machine learning techniques. Transfer learning enables the use of pre-trained models from non-medical contexts, adapted to the specific requirements of medical image analysis. This method enhances the model's ability to generalize from limited data, which is often a bottleneck given the scarcity and high cost of annotated medical images.

Additionally, the dynamic nature of medical research, with its continual evolution of data and insights, necessitates models that can adapt without extensive retraining. Incremental learning meets this need by allowing the model to assimilate new data while preserving previously learned information, ensuring the diagnostic system remains current with the latest medical advancements. This project aims to demonstrate that integrating transfer and incremental learning can substantially improve the accuracy, efficiency, and adaptability of cancer diagnostics tools, thereby revolutionizing how histopathology images are interpreted and used in clinical settings.

2 BACKGROUND

2.1 Foundations of Histopathology Image Analysis

Histopathology, the microscopic examination of tissue in order to study the manifestations of disease, stands as a cornerstone in cancer diagnostics. Traditionally, this analysis has been performed manually by pathologists, a process that requires extensive time and expertise. The advent of digital pathology, which involves scanning traditional glass slides to produce digital images, has opened new avenues for applying image analysis techniques. One of the primary challenges in

Author's Contact Information: Ankur Yadav, ankur.yadav@utdallas.com, The University of Texas at Dallas, Richardson, Texas, USA.

ACM XXXX-XXXX/2024/5-ART
https://doi.org/10.1145/nnnnnnn.nnnnnnn

histopathology image analysis is the high variability in tissue appearances, which can vary widely not only between different types of cancers but also within the same tumor due to heterogeneity. Another challenge is the sheer scale of data, as each slide can generate gigabytes of data, making manual analysis labor-intensive and prone to human error. These factors underscore the necessity for advanced computational methods that can enhance the precision, speed, and reliability of cancer diagnostics.

2.2 Evolution of Machine Learning in Medical Imaging

The integration of machine learning into medical imaging marks a significant evolution in the field, transforming diagnostic techniques across various specialties. Initially, traditional machine learning algorithms like Support Vector Machines (SVMs) and Random Forests were employed to analyze medical images, providing a baseline for automated diagnostic systems. However, the introduction of deep learning has catalyzed a paradigm shift, particularly with the development of convolutional neural networks (CNNs). These deep learning models have demonstrated remarkable success in image classification tasks due to their ability to learn complex patterns and features directly from large volumes of data without the need for manual feature extraction. This capability has been particularly transformative in medical imaging, where CNNs and other deep architectures are now routinely used to detect anomalies, segment regions of interest, and predict disease progression with high accuracy. The adoption of these advanced models has not only increased diagnostic accuracy but also significantly reduced the time required for analysis, supporting faster and more efficient patient care.

2.3 Need for Incremental Learning in Biomedical Images

Incremental learning has emerged as a crucial strategy in biomedical imaging due to the dynamic nature of medical data and the continuous advancement in medical knowledge. Traditional machine learning models require retraining from scratch whenever new data is introduced, which is not only computationally expensive but also impractical in clinical settings where new data accumulates rapidly. Incremental learning addresses this challenge by enabling models to learn from new data without forgetting previously acquired knowledge, thus maintaining their relevance over time. This is particularly vital in histopathology, where new cancer types may be discovered, and existing classifications updated. By integrating incremental learning, diagnostic models can adapt to these changes with minimal overhead, ensuring that they continue to provide accurate and timely insights. Furthermore, this adaptability is essential for personalized medicine, as it allows models to be updated with patient-specific data, enhancing the precision of diagnostics and treatments.

2.4 Related Work

The integration of machine learning into histopathology has notably advanced through the adoption of sophisticated models designed to handle the complexities of medical images. Recent research underscores the importance of strategies that mitigate the challenges posed by the dynamic nature of medical datasets and the immense scale of histopathological images. For instance, the MICIL algorithm, introduced in a recent study, leverages multiple-instance and class-incremental learning to address catastrophic forgetting—a critical issue when training models sequentially with new batches of data, particularly in the context of skin cancer WSIs [6]. This approach reflects a broader trend in employing advanced learning techniques such as knowledge distillation and data rehearsal to enhance model stability and accuracy in ongoing medical tasks [6].

Similarly, the incremental learning paradigm has been effectively utilized in other contexts, such as in the work presented on the BreakHis dataset, which emphasizes the efficacy of curriculum incremental learning. This method strategically increases the complexity of data during training,

improving model performance through a well-structured learning progression [5]. Additionally, research involving incremental boosting convolution networks for breast cancer diagnosis illustrates the potential of incrementally enhancing models to better classify complex image features, offering significant improvements over traditional training methods [12]. Another study contrasts the effectiveness of training deep convolutional neural networks from scratch against using pre-trained models, revealing that fine-tuned pre-trained networks can achieve competitive performance, thereby supporting the feasibility of transfer learning in histopathology image analysis [4]. These studies collectively inform the development of more robust, efficient, and adaptable diagnostic tools, paving the way for their implementation in clinical settings.

3 FORMAL PROBLEM DEFINITION

The project "Evolving Diagnostics: Incremental and Transfer Learning in Histopathology Image Analysis for Cancer Detection" aims to enhance the diagnostic capabilities of machine learning models by utilizing both transfer and incremental learning strategies with histopathology images. These objectives are designed to not only validate the effectiveness of these methodologies but also to ensure the model's adaptability and long-term utility in clinical settings.

3.1 Objective 1: Transfer Learning Model Development

Goal: Develop a transfer learning model by training the EfficientNet architecture, chosen for its scalability and efficiency, on a biomedical dataset containing images from four different tumor types. This model will serve as a robust feature extractor tailored for medical image analysis.
Application: Evaluate the model's generalization capabilities by testing it on two additional datasets not used during the initial training phase. This step will demonstrate the practical utility of the transfer learning approach across diverse medical imaging contexts.

3.2 Objective 2: Incremental Learning Implementation

Initial Training: Start by training the model with 50% of the dataset to establish a baseline for performance.
Incremental Updates: Continue with five incremental training phases, each using an additional 10% of the dataset. During these phases, only the last three blocks of the EfficientNet architecture will be retrained, focusing on refining the model's higher-level feature representations.
Performance Evaluation: After each incremental phase, evaluate the model's performance on a validation set to monitor changes in diagnostic accuracy. This process will highlight the effectiveness of incremental learning in enhancing the model's capabilities with the integration of new data.

These objectives underscore our commitment to demonstrating that advanced machine learning techniques, specifically transfer and incremental learning, can significantly improve the precision and adaptability of tools used in cancer diagnostics. The project will provide comprehensive insights into how these methodologies can be effectively applied to revolutionize medical imaging analysis.

4 EXAMPLE

In the medical field, particularly in cancer diagnostics, data continually evolves as new cases are diagnosed and treatment methodologies progress. This dynamic nature presents a significant challenge for traditional machine learning models, which are typically trained on static datasets. For instance, models trained on the ImageNet dataset are optimized for recognizing generic objects and may not perform adequately when applied directly to histopathology images. These images, used to diagnose diseases at the cellular level, exhibit unique features that are significantly different from the objects found in typical image recognition datasets.

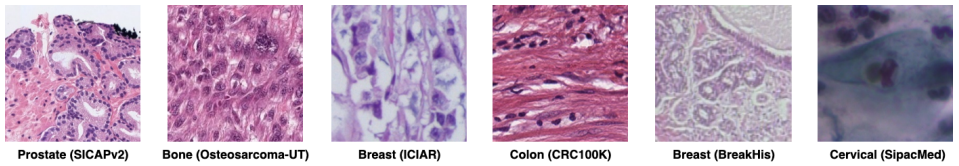

Fig. 1. Representative histopathological image samples from different cancer datasets. From left to right: Prostate tissue from SICAPv2, an osteosarcoma-affected bone from Osteosarcoma-UT, breast tissue from ICIAR, colorectal cancer from CRC100K, breast cancer from BreakHis, and cervical cells from SipacMed. These images illustrate the diverse morphological characteristics and staining patterns the HistoNet model trains to analyze and classify.

To illustrate this challenge, consider the use of ImageNet-trained models for feature extraction in histopathology. While these models are adept at identifying everyday objects, their ability to discern subtle variations in cell morphology—which are crucial for accurate cancer diagnosis—is limited. This discrepancy highlights the need for specialized feature extraction models that are tailored to the nuances of medical imaging.

By implementing transfer learning and incremental learning, we can develop a model that not only specializes in histopathology image analysis but also evolves over time. For example, as new patient data becomes available annually, this model can be incrementally trained to incorporate new insights without starting from scratch. This approach not only enhances the model's accuracy and adaptability but also ensures its longevity and relevance in clinical settings.

Furthermore, the versatility of this model allows it to serve as a generic feature extractor for various types of cancer data. Once the model is adept at extracting relevant features from histopathology images, it can be fine-tuned with smaller, specific datasets for different types of cancers. This methodology provides a robust foundation for developing specialized diagnostic tools across multiple cancer types, illustrating a scalable and efficient approach to medical image analysis in the ever-evolving landscape of healthcare.

5 APPROACH/FRAMEWORK

Describe the methodology or framework that you are using in your project, including any theoretical or conceptual models.

5.1 Dataset Selection

Our project strategically utilizes various publicly available benchmark datasets for the automated detection and classification of cancer through histopathological (HP) image analysis. These datasets are essential for training and validating HistoNet, our specialized feature extractor, designed to recognize and classify diverse morphological characteristics and staining patterns typical of different cancer tissues. We categorize these datasets into two groups: those used for both training and validation, and those reserved exclusively for validation to assess the model's generalizability and performance.

5.1.1 Training and Validation Datasets. For the development and continuous refinement of HistoNet, we have selected datasets that provide a comprehensive representation of tumor tissues across different types and stages of cancer. These datasets include:

- **Breast Cancer: BACH Dataset** [1] – This dataset, presented at the ICIAR 2018 Grand Challenge, consists of a diverse range of histopathological images indicative of various stages of breast cancer.

- **Prostate Cancer: SICAPv2 Dataset** [8] – Contains annotated prostate histology Whole Slide Images (WSIs) with detailed global Gleason scores and patch-level Gleason grades.
- **Bone Cancer: UT-Osteosarcoma Dataset** [2] – Offers a collection of H&E stained osteosarcoma histology images from the UTSW Medical Center.
- **Colon Cancer: CRC100K** [3] – Features 100,000 histological images of colorectal cancer and healthy tissue, aiding comprehensive tissue classification.

5.1.2 Validation-Only Datasets. To further assess the generalizability and effectiveness of HistoNet in processing unseen data, we utilize additional datasets exclusively for the validation phase:

- **Breast Cancer: BreakHis Dataset** [9] – Comprises microscopic images of breast tumor tissue labeled as benign or malignant, across various magnifications, which HistoNet has not been exposed to during training.
- **Cervical Cancer: SipacMed** [7] – Specifics regarding the SipacMed dataset are less documented, but it is intended to validate the transferability of the features extracted by HistoNet.

5.2 Data Preparation and Utilization

The efficacy of machine learning models in image-based analysis is critically dependent on the quality and uniformity of the input data. To address this, a detailed data preparation protocol has been established, ensuring each histopathological image is optimally pre-processed for effective deep learning analysis. This subsection outlines the key steps undertaken in our data preparation process:

Resizing: All images are resized to a uniform resolution of 224x224 pixels. This resolution was selected to balance the need for retaining essential histological details for accurate disease identification against the demands of computational efficiency. Standardizing the image size ensures uniformity across various datasets, facilitating more consistent model training.

Normalization: To compensate for variations in staining techniques across different laboratories and over time, a rigorous normalization process is applied to each image. This step adjusts staining intensity and color balance, helping the model focus on pertinent morphological features rather than discrepancies in stain quality.

Augmentation: Data augmentation techniques such as random rotations, flips, and scalings are employed to artificially expand the dataset. This practice helps the models learn to recognize and classify histopathological features accurately, regardless of how the tissue is oriented or presented in the image. Augmentation is crucial for enhancing the robustness and generalization ability of the models, particularly in medical imaging where sample presentation can vary significantly.

These preparatory measures are meticulously applied across our datasets, providing a robust foundation for our models to learn effectively and perform precise classifications of a wide array of histopathological images. Our commitment to rigorous data standardization and enhancement exemplifies our dedication to advancing computational pathology through high-performing, efficient models.

5.3 Hardware and Software Configuration

The experiments were conducted using a high-performance computing setup equipped with NVIDIA RTX 2060 Super GPUs, each boasting 8GB of VRAM. This powerful hardware configuration is essential for handling the intensive computational demands associated with training deep learning models on large histopathological image datasets. Additionally, the system included 16 GB of RAM to ensure efficient data processing and multitasking during model training and evaluation.

For software, PyTorch was selected as the primary framework for developing our deep learning models due to its flexibility and extensive support for advanced modeling techniques. Alongside

Table 1. Distribution of image samples across different cancer datasets for training and testing.

Dataset	Classes	Train	Test
Healthy vs Tumor	Healthy	24886	10714
	Tumor	51406	22010
	Total	**76292**	**32715**
Bone (Osteosarcoma-UT)	Healthy	5982	2071
	Viable	4180	1782
	Non-Viable	5424	2325
	Total	**15586**	**6720**
Breast (ICIAR)	Healthy	32347	14001
	Benign	3324	1428
	In Situ	3278	1402
	Invasive	3339	1427
	Total	**13188**	**5658**
Colon (CRC100K)	Healthy	6134	2627
	STR	7313	3133
	TUM	10022	4295
	Total	**23469**	**10055**
Prostate (SICAPv2)	Healthy	9523	4073
	G3	4277	1834
	G4	8491	3622
	G5	1758	753
	Total	**24049**	**10882**
Breast (BreakHis)	Benign	1989	859
	Malignant	5235	2245
	Total	**7224**	**3104**
Cervical (SipacMed)	im_Dysk.	10686	4641
	im_Koil.	10743	4673
	im_Meta.	12239	5251
	im_Para.	5444	2382
	im_Supe.	5582	2380
	Total	**44876**	**19277**

PyTorch, we employed NumPy, SciPy, and scikit-learn for effective data manipulation, statistical analysis, and classical machine learning operations. This suite of software tools provides a robust environment for the preprocessing of data and the execution of complex model training algorithms, thereby supporting our goal of achieving high accuracy and reliability in our computational pathology experiments.

5.4 Model Architectures

HistoNet Architecture: The HistoNet model builds upon the EfficientNet B0 framework, specifically tailored and fine-tuned for the unique requirements of histopathological image analysis. This initial deep-learning model is designed to discern intricate details within histopathological images, essential for accurate cancer diagnosis. Through rigorous training on a diverse dataset encompassing various cancer types, HistoNet specializes in extracting features that highlight the textural and structural nuances crucial for the interpretation of medical images.

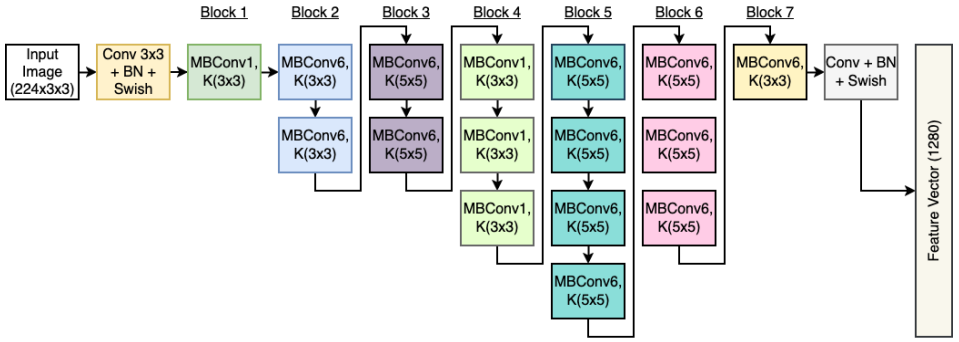

Fig. 2. **Architecture of EfficientNet-B0**, featuring seven blocks of varying configurations. The network begins with a standard convolutional layer, followed by a series of mobile inverted bottleneck convolutions (MBConv) with kernel sizes denoted as K(3x3) or K(5x5). The final feature vector has a dimensionality of 1280, extracted post a concluding convolutional layer with batch normalization and Swish activation.

Individual Model Training: The dataset utilized for training and testing HistoNet was split into 70% for training and 30% for testing to ensure robust model evaluation. For tasks that required fine-tuning of the model, the training dataset was further divided, using an 80:20 split to create a validation subset. The optimization of the model was carried out using the Adam optimizer for Multilayer Perceptrons (MLPs), while the XGBoost models were configured using their default settings. Hyperparameters were meticulously refined based on their performance on the validation subset, ensuring optimal model behavior.

5.5 Training Protocols

The methodology behind training our models is pivotal to our study's aim of enhancing histopathological image classification. This section highlights the streamlined procedures from data preprocessing to model evaluation.

Data Preprocessing: To ensure consistency across all inputs, histopathological images were resized to 224x224 pixels. To augment the diversity of our dataset and bolster model robustness, we applied data augmentation techniques such as rotations and flips.

Feature Extraction: Deep learning-based features were extracted using HistoNet. This approach focuses solely on leveraging the computational power and sophisticated pattern recognition capabilities of deep learning to analyze complex image data effectively.

Model Training: Our dataset was split into 70% for training and 30% for testing. For tasks requiring model fine-tuning, the training set was further divided using an 80:20 split for validation purposes. Model optimization was conducted with the Adam optimizer for MLPs, while XGBoost models utilized their default configurations. Hyperparameters were refined based on the performance in the validation subset.

Epochs and Evaluation: The training process was tailored to each experiment, incorporating early stopping based on validation loss to avoid overfitting. Models' effectiveness was assessed through metrics such as accuracy, precision, recall, and F1 score, offering a comprehensive evaluation of their classification capabilities.

This protocol underlines our commitment to developing and validating models that significantly improve the accuracy and efficiency of detecting and classifying cancer types in histopathological images.

5.6 Selection of EfficientNet B0 for Feature Extraction Training

Our study prioritizes EfficientNet B0 for its balance between high performance and computational efficiency. This model, the simplest in the EfficientNet family, offers state-of-the-art accuracy on the ImageNet dataset with comparatively lower complexity and computational demand.
Performance and Complexity: EfficientNet B0 is chosen for its innovative scaling method that balances speed and accuracy. This feature is crucial for the detailed analysis required in histopathological image examination, making it an ideal model for our purposes.
Dataset Size Consideration: The lower complexity of the architecture is advantageous for our dataset, which is smaller than ImageNet, mitigating the risk of overfitting while still capitalizing on the model's robust feature extraction abilities.
Computational Feasibility: Given hardware constraints and the need for iterative testing, the EfficientNet B0's design, requiring fewer computational resources, facilitates broader experimentation without sacrificing training and evaluation efficiency.
Rationale for Retraining: Retraining the model on histopathological images, rather than using it as pre-trained on ImageNet, customizes its feature extraction to focus on cancer-specific characteristics, significantly enhancing diagnostic precision.

This selection process underscores our commitment to performing histopathological analysis through a model that has high-performance, low-complexity, and is computationally efficient.

6 EXPERIMENTS

6.1 Experiment 1: Assessing Feature Extraction Performance with EfficientNet B0

6.1.1 Objective. The primary objective of this experiment is to evaluate the efficacy of the EfficientNet B0 model, retrained specifically for histopathological images, in extracting relevant features compared to other pre-existing transfer learning models.

6.1.2 Methodology. EfficientNet B0 will be retrained using a compilation of four distinct histopathology image datasets, each representing different types of cancer. This retraining aims to tailor the model's feature extraction capabilities specifically to the nuances of histopathological data. After retraining, we will utilize the HistoNet architecture to extract features from the images. The extracted features will then be compared against those obtained using other standard transfer learning models that are commonly employed in medical image analysis. The comparison will focus on the quality and relevance of the features for cancer diagnosis.

6.1.3 Data Handling. The datasets used for this experiment include breast cancer, prostate cancer, bone cancer, and colon cancer images. These datasets will provide a comprehensive base for assessing the model's ability to generalize across various forms of cancerous tissues. Performance will be evaluated using standard metrics such as accuracy, precision, recall, and the F1 score to determine the effectiveness of each model in feature extraction tasks.

6.1.4 Results. In a comprehensive benchmarking exercise, the feature extraction capabilities of HistoNet were evaluated against well-known transfer learning models such as VGG16, VGG19, InceptionV3 [10], and the EfficientNet series [11] (B0, B1, B2). The evaluation covered the BreakHis and SipacMed datasets, using both MLP and XGBoost (XGB) classifiers to assess the performance of the models (see Table 2).

On the BreakHis dataset, HistoNet demonstrated exceptional performance when paired with an MLP classifier, achieving a notable test accuracy of 95.68%. This performance was significantly higher than that of its counterparts, with EfficientNetB0 achieving a test accuracy of 94.23% under the same conditions. With the XGB classifier, HistoNet continued to lead, achieving a test accuracy of 92.91%, followed by EfficientNetB0 at 91.98%.

Table 2. Comparison with other Transfer Learning Models

Dataset	Trained Model	Features from	Number or Features	Accuracy Train	Accuracy Test
BreakHis	XGB	VGG16	512	1.0000	0.9146
		VGG19	512	1.0000	0.9027
		InceptionV3	2048	1.0000	0.8950
		EfficientNetB0	1280	1.0000	0.9198
		EfficientNetB1	1280	1.0000	0.9053
		EfficientNetB2	1408	1.0000	0.9008
		HistoNet	1280	1.0000	**0.9291**
	MLP	VGG16	512	1.0000	0.9325
		VGG19	512	1.0000	0.9285
		InceptionV3	2048	1.0000	0.9214
		EfficientNetB0	1280	1.0000	0.9423
		EfficientNetB1	1280	1.0000	0.9346
		EfficientNetB2	1408	1.0000	0.9301
		HistoNet	1280	1.0000	**0.9568**
SipacMed	XGB	VGG16	512	0.9944	0.8504
		VGG19	512	0.9912	0.8368
		InceptionV3	2048	0.9958	0.7713
		EfficientNetB0	1280	0.9972	0.8321
		EfficientNetB1	1280	0.9952	0.8015
		EfficientNetB2	1408	0.9966	0.7731
		HistoNet	1280	0.9963	**0.8705**
	MLP	VGG16	512	0.9979	0.8878
		VGG19	512	0.9924	0.8681
		InceptionV3	2048	0.9996	0.8177
		EfficientNetB0	1280	0.9914	0.8840
		EfficientNetB1	1280	0.9970	0.8584
		EfficientNetB2	1408	0.9992	0.8459
		HistoNet	1280	0.9550	**0.8935**

In our evaluation on the SipacMed dataset, HistoNet demonstrated a robust capability in feature extraction with an accuracy of 87.05% when using the XGBoost (XGB) classifier and 89.35% accuracy with the Multi-Layer Perceptron (MLP) classifier. These results underscore HistoNet's effectiveness in handling histopathological image analysis and its superiority over traditional methods in this specific application.

This comparative analysis clearly demonstrates the superior feature extraction ability of HistoNet, underscoring the advantage of integrating advanced feature extraction techniques with traditional transfer learning models to enhance classification accuracies in cancer diagnosis through histopathological imagery.

6.2 Experiment 2: Optimization of Incremental Learning

6.2.1 Objective. The objective of this experiment is to determine the optimal number of layers (or blocks) of the EfficientNet B0 model to retrain during the incremental learning process to effectively classify histopathological tiles as healthy or tumorous. Using the four designated datasets,

this experiment seeks to refine the model's accuracy in distinguishing between these critical classifications while minimizing the retraining overhead. This approach ensures that the model remains both efficient and capable of adapting to new data, which is essential for maintaining high performance in clinical diagnostics.

6.2.2 Methodology. Initially, the entire EfficientNet B0 model will be trained using 50% of the dataset to establish a baseline for feature extraction capabilities. Subsequently, incremental training will be conducted in phases, where different configurations of the top layers of the model are retrained. Specifically, the following configurations will be tested:

- Retraining only the top 1 block
- Retraining the top 2 blocks
- Retraining the top 3 blocks

In each iteration, the remaining dataset will serve as a validation holdout set to record the validation accuracy. This strategy allows for precise measurement of the impact of retraining different numbers of blocks on the model's ability to adapt to new information, ensuring the model remains effective without unnecessary computational expense.

6.2.3 Data Handling. The same datasets as described in previous experiments will be used, with each phase of incremental training employing a new subset of data that simulates the acquisition of new patient data. This mimics real-world scenarios where models must adapt to evolving information without full retraining.

6.2.4 Metrics for Evaluation. The primary metric for evaluation will be validation accuracy. This metric will help determine the effectiveness of each layer retraining configuration in improving the model's ability to classify new data accurately. This experiment is designed to elucidate the most efficient and effective strategy for incremental learning in deep learning models, particularly in the context of ongoing data evolution typical in medical fields like oncology.

6.2.5 Results. The initial phase of our experiment involved training the entire EfficientNet B0 model using 50% of the dataset. This comprehensive training was conducted over 25 epochs to establish a robust baseline for feature extraction across histopathological images, distinguishing between healthy and tumor tiles.

Following the baseline establishment, we embarked on a series of incremental training sessions. These sessions were structured to refine the model by retraining different numbers of the top blocks. Specifically, we implemented seven distinct methods, incrementally training with:

- only the top 1 block,
- the top 2 blocks,
- the top 3 blocks, and so forth up to the top 7 blocks.

Each method involved retraining with an additional 10% of the dataset, repeated five times to ensure consistency and reliability of the results.

This incremental training approach allowed us to systematically evaluate the impact of varying the depth of retraining on the model's performance. By comparing the validation accuracy across these configurations, we were able to identify the optimal balance between computational efficiency and diagnostic accuracy, essential for practical applications in clinical settings. The analysis of the incremental retraining results, depicted in the graph 3, reveals insightful trends regarding the model's adaptability and performance optimization. As the model is incrementally trained with additional 10% segments of the dataset, we observe varying improvements in accuracy depending on the number of top blocks retrained. Notably, configurations involving the retraining of three to six top blocks consistently demonstrate a stable and positive trend in accuracy, suggesting an

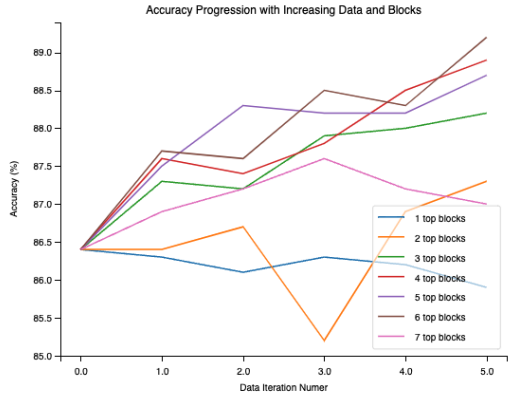

Fig. 3. Accuracy progression of the EfficientNet B0 model across five incremental data iterations, illustrating the impact of retraining varying numbers of top blocks. Each line represents a different retraining configuration, ranging from one to seven top blocks, with each increment involving an additional 10% of the training data. This graph highlights the relationship between the depth of retraining and model accuracy, with configurations of three to six top blocks showing the most consistent and stable improvements.

effective balance between learning new features and retaining previously learned information. However, the configurations with one or two top blocks exhibit more volatility, possibly indicating insufficient model adaptation to new data. The seven-block retraining shows a unique pattern of initial performance dip followed by recovery, hinting at potential overfitting issues that stabilize with further data integration. These results underscore the importance of choosing an appropriate depth of retraining to maximize performance while minimizing unnecessary computational expense and risk of overfitting.

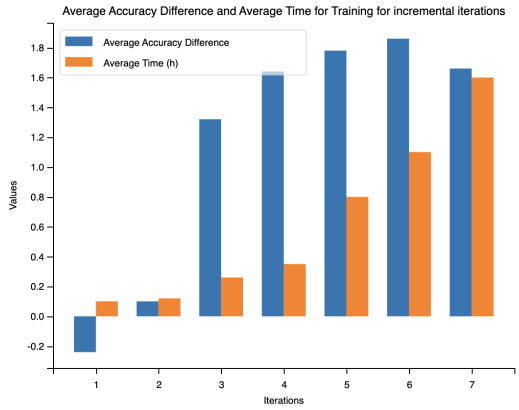

Fig. 4. Accuracy progression of the EfficientNet B0 model across five incremental data iterations, illustrating the impact of retraining varying numbers of top blocks. Each line represents a different retraining configuration, ranging from one to seven top blocks, with each increment involving an additional 10% of the training data. This graph highlights the relationship between the depth of retraining and model accuracy, with configurations of three to six top blocks showing the most consistent and stable improvements.

The plot 4 illustrates the relationship between the number of top blocks retrained in the HistoNet model and both the average accuracy improvements and the time required for each training iteration. As we increment the number of blocks retrained—from one up to seven—there is a notable increase in the training time, indicating a higher computational cost with each additional block. Interestingly, the average accuracy improvement does not consistently increase with more blocks retrained. While training more blocks tends to lead to higher initial accuracy improvements, this trend does not hold uniformly across all iterations. Specifically, retraining five blocks appears to offer a significant gain in accuracy without the highest time cost, suggesting an optimal balance at this level. Retraining seven blocks, while yielding the highest accuracy gain, requires substantially more time, indicating diminishing returns as the number of retrained blocks increases. This analysis highlights the importance of optimizing the number of layers retrained to achieve efficient and effective incremental learning in clinical diagnostic models.

7 DISCUSSION

This section reflects on the findings from our experiments with HistoNet, focusing on its optimization through incremental learning for histopathological image analysis. These insights have significant implications for the deployment of machine learning models in medical diagnostics.

7.1 Optimization of Incremental Learning

The study demonstrates the effectiveness of HistoNet's incremental learning capability, which adapts to new histopathological data without requiring a complete model retraining. The selective retraining of top layers, particularly the strategy involving five top blocks, optimized accuracy improvements while managing computational resources efficiently. This approach underlines the potential of precise, layer-specific retraining in enhancing model responsiveness and efficiency.

7.2 Implications for Clinical Practice

The findings are particularly relevant to clinical environments, where updating diagnostic tools with new medical data is crucial. HistoNet's ability to be incrementally trained means that it can evolve with minimal resource expenditure, maintaining up-to-date diagnostic accuracy. This capability is vital in clinical settings that demand both high accuracy and operational efficiency, ensuring that the models stay relevant over time.

7.3 Future Directions

Our findings open several avenues for future research to enhance and validate the incremental learning capabilities of HistoNet. One promising direction is to test HistoNet on larger and more diverse datasets, which could provide deeper insights into the model's scalability and robustness across various medical imaging contexts. Additionally, exploring the application of more complex models such as ResNet and vision transformers could offer comparative benchmarks and highlight the advantages or challenges of using different architectures for histopathological image analysis. Beyond simply adjusting the number of retrained top layers, a more granular analysis of individual layer contributions could yield significant improvements. By identifying which specific blocks within the model contribute most to incremental learning, we can optimize training protocols to focus on the most impactful layers, potentially reducing computational overhead while enhancing model accuracy. This layer-specific analysis will help in understanding the nuanced impact of different blocks on model training and data interpretation, further refining the deployment strategies for HistoNet in clinical settings.

7.4 Limitations and Challenges

While the results of this study are promising, it is important to acknowledge certain limitations that might have influenced the outcomes. One significant constraint was the limited amount of available computing power, which necessitated the selection of a smaller and potentially less complex model architecture, such as EfficientNet B0, rather than more computationally demanding models like larger ResNets or vision transformers. This limitation could affect the generalizability and scalability of our findings, as more powerful architectures may yield different insights, particularly in handling larger or more complex datasets. Future studies with access to enhanced computational resources could explore the implications of using more robust models, potentially leading to further improvements in accuracy and efficiency for clinical diagnostic tools.

In conclusion, this study reinforces the utility of targeted, efficient incremental learning strategies in medical imaging, striking a crucial balance between maintaining high diagnostic accuracy and managing computational resources effectively. The ongoing evolution of models like HistoNet is essential for leveraging the full potential of AI in enhancing medical diagnostic processes.

REFERENCES

[1] Guilherme Aresta, Teresa Araújo, Scotty Kwok, Sai Saketh Chennamsetty, Mohammed Safwan, Varghese Alex, Bahram Marami, Marcel Prastawa, Monica Chan, Michael Donovan, et al. 2019. Bach: Grand challenge on breast cancer histology images. *Medical image analysis* 56 (2019), 122–139.

[2] Harish Babu Arunachalam, Rashika Mishra, Ovidiu Daescu, Kevin Cederberg, Dinesh Rakheja, Anita Sengupta, David Leonard, Rami Hallac, and Patrick Leavey. 2019. Viable and necrotic tumor assessment from WSI of osteosarcoma using ML and DL models. *PloS one* 14, 4 (2019), e0210706.

[3] Jakob Nikolas Kather, Niels Halama, and Alexander Marx. 2018. 100,000 histological images of human colorectal cancer and healthy tissue. *Zenodo10* 5281 (2018).

[4] Brady Kieffer, Morteza Babaie, Shivam Kalra, and Hamid R Tizhoosh. 2017. Convolutional neural networks for histopathology image classification: Training vs. using pre-trained networks. In *2017 seventh international conference on image processing theory, tools and applications (IPTA)*. IEEE, 1–6.

[5] Mouna Sabrine Mayouf and Florence Dupin de Saint-Cyr. 2022. Curriculum Incremental Deep Learning on BreakHis DataSet. In *Proceedings of the 2022 8th International Conference on Computer Technology Applications*. 35–41.

[6] Pablo Meseguer, Rocío Del Amor, and Valery Naranjo. 2024. MICIL: Multiple-Instance Class-Incremental Learning for skin cancer whole slide images. *Artificial Intelligence in Medicine* (2024), 102870.

[7] Marina E. Plissiti, P. Dimitrakopoulos, G. Sfikas, Christophoros Nikou, O. Krikoni, and A. Charchanti. 2018. Sipakmed: A New Dataset for Feature and Image Based Classification of Normal and Pathological Cervical Cells in Pap Smear Images. In *ICIP*. 3144–3148. https://doi.org/10.1109/ICIP.2018.8451588

[8] Julio Silva-Rodríguez, Adrián Colomer, María A Sales, Rafael Molina, and Valery Naranjo. 2020. Going deeper through the Gleason scoring scale: An automatic end-to-end system for histology prostate grading and cribriform pattern detection. *Computer Methods and Programs in Biomedicine* 195 (2020), 105637.

[9] Fabio A Spanhol, Luiz S Oliveira, Caroline Petitjean, and Laurent Heutte. 2015. A dataset for breast cancer histopathological image classification. *Ieee transactions on biomedical engineering* 63, 7 (2015), 1455–1462.

[10] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. 2016. Rethinking the Inception Architecture for Computer Vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

[11] Mingxing Tan and Quoc V. Le. 2019. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*. 6105–6114.

[12] Duc My Vo, Ngoc-Quang Nguyen, and Sang-Woong Lee. 2019. Classification of breast cancer histology images using incremental boosting convolution networks. *Information Sciences* 482 (2019), 123–138.