

Evaluating conditional convolution for magnification-aware histopathological image classification

Wissal El Habti, Abdellah Azmani

Intelligent Automation and BioMed Genomics Laboratory, Faculty of Science and Technology of Tangier,
Abdelmalek Essaadi University, Tangier, Morocco

Article Info

Article history:

Received Jul 16, 2025

Revised Jul 21, 2026

Accepted Aug 12, 2026

Keywords:

Breast cancer classification
Class imbalance
Conditionally parameterized
convolutions
Histopathology
Magnification-aware CNNs
Medical image analysis

ABSTRACT

Histopathological breast cancer images exhibit multi-scale variability and class imbalance, challenging automated classification. This study integrates conditionally parameterized convolutions (CondConv) into convolutional neural networks (CNNs) to enhance accuracy and robustness across heterogeneous magnifications. This study proposes CondConv-augmented mobile network version 2 (MobileNetV2) and densely connected convolutional network 121 (DenseNet121), which dynamically adapt convolutional kernels to input characteristics, improving generalization across spatial resolutions. This is the first evaluation of CondConv-based architectures for multi-magnification histopathological classification. This study evaluates these architectures at 40×, 100×, 200×, and 400× using imbalanced and class-rebalanced datasets. Performance was measured via convergence analysis, macro-averaged metrics, and statistical validation with Friedman tests, Nemenyi post-hoc comparisons, and Wilcoxon signed-rank tests. CondConv-enhanced networks matched or surpassed static-kernel variants, with notable gains at 100× and 200× where scale variability peaks. These improvements persisted under class imbalance, demonstrating resilience to distributional skew. Although differences narrowed at 400× due to feature saturation, statistical analyses confirmed significant enhancements at lower magnifications, especially for CondMobileNetV2.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Wissal El Habti

Intelligent Automation and BioMed Genomics Laboratory, Faculty of Science and Technology of Tangier

Abdelmalek Essaadi University

Tangier, Morocco

Email: wissal.elhabti@etu.uae.ac.ma

1. INTRODUCTION

Breast cancer remains the most commonly diagnosed malignancy among women worldwide, accounting for 2.3 million new cases in 2022 alone, and contributing to over 670,000 cancer-related deaths globally across all types [1]. Within the continuum of cancer care, the diagnostic phase plays a pivotal role—serving as the gateway to staging, treatment planning, and long-term prognosis. Among the array of diagnostic methods, histopathological examination occupies a definitive position in the cancer patient lifecycle, providing microscopic validation of malignancy and guiding critical therapeutic decisions. Unlike imaging modalities such as mammography, ultrasound, or magnetic resonance imaging (MRI), histopathological analysis directly reveals the cellular and architectural changes characteristic of tumor progression, with minimal radiative risk to the patient [2].

Despite its central role in cancer diagnostics, histopathological slide interpretation remains a time-consuming process that demands expert pathologist input. The high volume of biopsy samples and the

nuanced morphological variations across cases make manual classification prone to fatigue-induced errors and inter-observer inconsistencies [3]–[6]. These limitations have prompted growing interest in computer-aided diagnostic (CAD) systems that incorporate artificial intelligence, particularly deep learning models [7]. Within histopathology, convolutional neural networks (CNNs) have shown particular effectiveness in learning discriminative morphological representations and automating tissue classification tasks [8], [9].

However, automated histopathological classification is challenged by the wide variation in image resolution across magnification levels. In datasets like breast cancer histopathological database (BreakHis), each sample is captured at multiple optical zoom levels (40×, 100×, 200×, and 400×), with each scale revealing distinct spatial and contextual cues. Conventional CNNs such as visual geometry group (VGG) and densely connected convolutional network (DenseNet) operate using fixed convolutional kernels, which lack the ability to adapt their receptive fields based on the input resolution. This inflexibility can lead to inconsistent performance when models are exposed to resolution shifts between training and testing data [10]. Additionally, CNN training is sensitive to dataset imbalance and annotation scarcity, leading to overfitting and reduced generalization [8], [10].

Although some studies incorporate attention mechanisms or transfer learning to compensate for these issues [10]–[12], most still rely on static convolutional assumptions and evaluate models under aggregate performance metrics without stratifying by magnification. This leaves a critical gap in understanding how CNNs behave under real-world variability in image granularity. Notably, conditionally parameterized convolutions (CondConv), introduced in [13], offer a promising direction by enabling dynamic routing of input data through multiple expert convolutional filters.

A keyword-based search on Google Scholar using “conditionally parameterized convolutions” AND “cancer” retrieved approximately 60 studies, the majority of which apply CondConv to segmentation, localization, or detection tasks. For instance, Zhang *et al.* [14] introduced dynamic on-demand network (DoDNet) for multi-organ and tumor segmentation using dynamic filters guided by task-aware priors. In cervical cytology, task decomposing and cell comparing network (TDCC-Net) [15] applied task decomposition and CondConv-driven contrastive learning for abnormal cell detection. Li *et al.* [16] embedded adaptive kernels into a dynamic hierarchical transformer network (DHT-Net), improving liver and tumor segmentation by enhancing spatial representation. These studies demonstrate the utility of CondConv in spatially localized tasks. However, to our knowledge, no prior work has systematically evaluated CondConv in histopathological image classification, particularly with magnification stratification. This study addresses this gap by integrating CondConv into two widely used CNNs—mobile network version 2 (MobileNetV2) and DenseNet121—and evaluating their classification performance across four magnifications (40×, 100×, 200×, and 400×) using the BreakHis dataset. The investigation is guided by one primary research question: does the integration of CondConv into CNNs improve classification performance and robustness across magnification levels and data imbalance scenarios in histopathological breast cancer images?

To address this question, this research was divided into several sections. Section 2 outlines the experimental methodology, which includes data collection and preprocessing; the integration of CondConv into MobileNetV2 and DenseNet121; and the formulation of evaluation protocols. Section 3 reports and discusses the results, beginning with quantitative evaluations; followed by statistical analysis stratified by magnification level; including Friedman, Nemenyi, and Wilcoxon tests. Lastly, section 4 concludes this research.

2. MATERIAL AND METHOD

2.1. Data collection and preprocessing

The BreakHis dataset was employed for training and evaluation, comprising 7,909 histopathological images of breast tumors collected from 82 individuals. Each sample is annotated as either benign or malignant and captured at one of four magnification levels: 40×, 100×, 200×, or 400× [17]. Representative samples across classes and magnifications are illustrated in Figure 1, highlighting the diversity in tissue morphology. Specifically, Figure 1(a) presents representative malignant histopathological images at 40×, 100×, 200×, and 400× magnifications, while Figure 1(b) shows corresponding benign samples across the same magnification levels. The dataset was restructured for preprocessing into dedicated training, validation, and test folders per magnification. Data augmentation was applied exclusively to the training set and included random horizontal and vertical flips, affine rotations, color jittering, and scaled cropping, all designed to increase variability and enhance generalization. Images were uniformly resized to 224×224 pixels and normalized using the standard mean and standard deviation values from ImageNet. The dataset was then loaded into PyTorch data loaders for batch-wise training and evaluation, facilitating reproducibility and controlled experimentation.

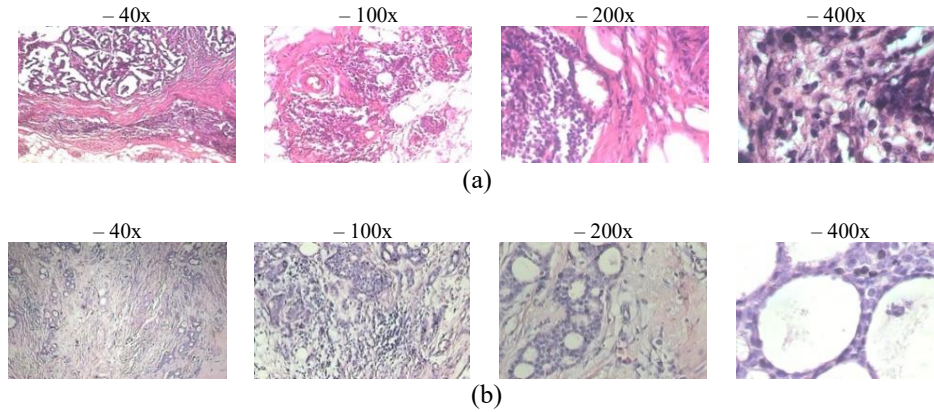


Figure 1. Histopathological image samples from the BreakHis dataset across magnification levels: (a) malignant and (b) benign

2.2. Model architecture

To address the representational limitations of static convolutional filters in conventional CNNs, CondConv were proposed [13] as a lightweight dynamic alternative that enhances model capacity with minimal inference overhead. Unlike traditional convolutions that use a single shared kernel for all inputs, CondConv layers dynamically assemble instance-specific kernels as a weighted sum of n expert filters $\{W_i\}_{i=1}^n$, guided by input-dependent routing weights $\alpha_i(x)$. The output of a CondConv layer for an input x is expressed as in (1).

$$\text{Output}(x) = \sigma((\alpha_1 \cdot W_1 + \dots + \alpha_n \cdot W_n) \times x) \quad (1)$$

Here, $\sigma(\cdot)$ denotes the activation function, and the routing weights $\alpha_i(x) = r_i(x) \in [0,1]$ are computed through a differentiable mechanism involving global average pooling, a fully connected layer, and a sigmoid activation, as defined in (2).

$$r(x) = \text{Sigmoid}(GAP(x) \cdot R) \quad (2)$$

Where $R \in \mathbb{R}^{d \times n}$ is a learned projection matrix and d is the number of channels. Crucially, this formulation is mathematically equivalent to a linear mixture of experts but remains computationally efficient by applying only a single forward pass per expert kernel per input.

To evaluate the effectiveness of CondConv in the context of histopathology image classification, two architectures of differing complexity were selected. MobileNetV2 [18] was chosen for its efficiency and compactness, leveraging inverted residuals and depthwise separable convolutions to maintain low computational cost—making it suitable for real-time and embedded systems. DenseNet121 [19], on the other hand, offers deeper representation through densely connected blocks that facilitate gradient flow and feature reuse, making it a strong candidate for high-resolution tasks. By adapting the final convolutional blocks of these architectures to use CondConv, the study assesses whether the introduction of conditional computations benefits both lightweight and deeper CNNs in distinguishing multi-scale morphological patterns inherent in histopathological breast cancer images.

2.2.1. CondMobileNetV2

MobileNetV2 [18] is a lightweight CNN architecture designed for efficient image classification on resource-constrained devices. Its design is centered around two major concepts: depthwise separable convolutions and the inverted residual with linear bottleneck structure. This architectural innovation enables a significant reduction in parameter count and computational cost without compromising performance. The core building block of MobileNetV2 consists of a three-layer module. Given an input tensor $x \in \mathbb{R}^{H \times W \times C}$, the block first applies a pointwise convolution to expand the channel dimension by a factor t (commonly $t = 6$), resulting in an intermediate representation in $\mathbb{R}^{H \times W \times C}$. This is followed by a depthwise convolution that applies a spatial filter to each channel independently as in (3).

Subsequently, a second pointwise convolution projects the output to the desired output dimensionality C' . The full transformation, including ReLU6 activations (except after the last projection), is defined as in (4). If the input and output dimensions match and a residual connection is applied. This structure efficiently preserves information while facilitating feature reuse and gradient propagation.

$$DWConv(x) = x * k_d, k_d \in \mathbb{R}^{3 \times 3 \times 1} \quad (3)$$

$$y = x + PWConv_2 \left(DWConv_2 \left(ReLU6(PWConv_1(x)) \right) \right) \quad (4)$$

The CondMobileNetV2 variant adapts this baseline by replacing specific standard convolutional layers with CondConv. Rather than using a fixed kernel, each CondConv layer dynamically computes the convolutional kernel as a weighted combination of n expert filters. The routing weights $\alpha_i(x)$, modulated by the input instance via a compact routing network (global average pooling followed by a fully connected layer and sigmoid activation), determines the contribution of each expert. This mechanism enables adaptive feature extraction tailored to individual histopathology image variations.

CondConv layers are incorporated at key locations within MobileNetV2, maintaining architectural consistency while enhancing model flexibility. The overall architectural comparison is presented in Figure 2. Figure 2(a) presents the baseline MobileNetV2 architecture composed of standard depthwise separable convolutions and inverted residual blocks. Figure 2(b) illustrates the proposed CondMobileNetV2, where selected convolutional layers and the final classifier are equipped with CondConv modules incorporating dynamic routing. By dynamically combining multiple expert kernels through a routing mechanism, the architecture adapts feature extraction to the characteristics of each histopathological image while preserving the lightweight design of the original network.

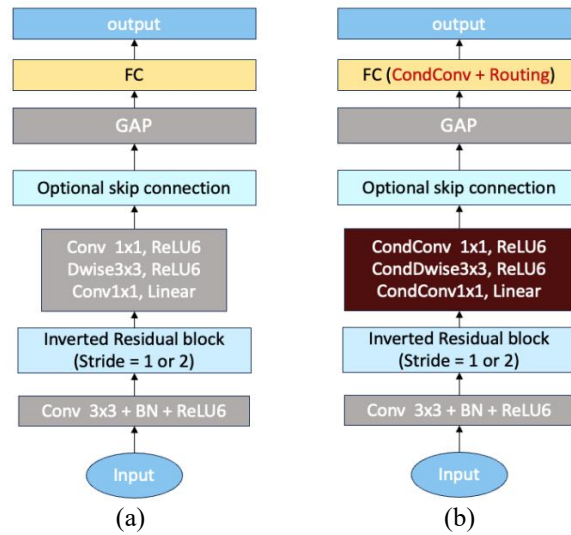


Figure 2. Structural comparison between (a) original MobileNetV2 and (b) CondMobileNetV2

2.2.2. CondDenseNet121

DenseNet121 [20] receives the feature maps of all preceding layers as input, promoting feature reuse, and alleviating the vanishing gradient problem. Mathematically, the input to l th layer is given by concatenation $x_l = H_l([x_0, x_1, \dots, x_{l-1}])$, where $H_l(\cdot)$ is a composite function of batch normalization, rectified linear unit (ReLU), and convolution operations. This architecture enables deeper networks with fewer parameters, as it avoids redundant feature computation.

This study adapts DenseNet121 into a dynamic variant—CondDenseNet121—by selectively integrating conditional convolutions into the architecture. This enhancement is motivated by the need to address the multiscale nature of histopathological images, where morphological features vary significantly across magnification levels. Since deeper layers encode more abstract, scale-sensitive semantics, this study inserts CondConv layers into early and final blocks to allow the network to dynamically adjust its filters based on input characteristics. The architectural comparison between the baseline and proposed models is presented in Figure 3. Figure 3(a) presents the baseline DenseNet121 architecture, composed of four densely connected blocks linked by transition layers. Figure 3(b) illustrates the proposed CondDenseNet121 architecture, in which the initial 7×7 convolution and the final dense block are replaced with CondConv-enhanced counterparts, while preserving the dense connectivity and output dimensionality of the original architecture. This preserves compatibility and enables direct comparison with the baseline model. The resulting hybrid design retains DenseNet's compactness and efficient gradient flow while introducing input-conditioned flexibility, improving the model's ability to generalize across varied histopathological textures.

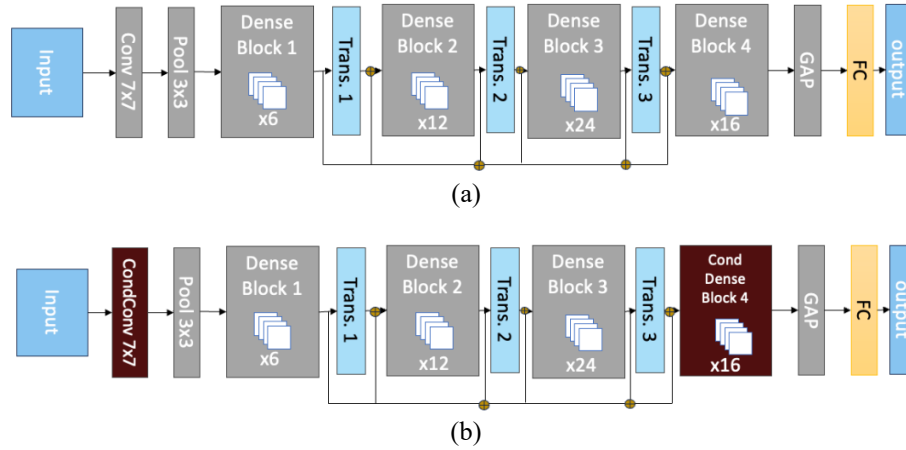


Figure 3. Architectural comparison between (a) standard DenseNet121 and (b) CondDenseNet121

2.3. Hypothesis and evaluation protocol

This study hypothesizes that replacing standard convolutional layers with conditional convolution blocks in MobileNetV2 and DenseNet121 leads to statistically significant improvements in classification performance across varying magnification levels in histopathological breast cancer images. This study posit that such gains arise from the enhanced adaptive capacity of CondConv models when confronted with the heterogeneous structures and scales intrinsic to histological patterns. To rigorously validate this hypothesis, this study adopts a two-phase evaluation protocol. The first phase involves a quantitative assessment of each model's performance using accuracy, macro-precision, macro-recall, and macro-F1 score, all computed at the final epoch. Let TP_i , FP_i , FN_i denote true positives, false positives, and false negatives for class i , and C the number of classes. The metrics are given by (5) to (8).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$Precision = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i+FP_i} \quad (6)$$

$$Recall = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i+FN_i} \quad (7)$$

$$F1 - score = \frac{1}{C} \sum_{i=1}^C \frac{2P_i R_i}{P_i + R_i} \quad (8)$$

Here, this study adopted a layered statistical evaluation framework to assess the significance of model performance differences across magnification levels. First, the Friedman test [19] is employed to detect overall differences among multiple classifiers over epochs. It is a non-parametric alternative to repeated-measures analysis of variance (ANOVA) and tests the null hypothesis that all models perform equally. The test statistic is given by (9).

$$\chi_F^2 = \frac{12N}{k(k+1)} \left[\sum_{j=1}^k R_j^2 \right] - 3N(k+1) \quad (9)$$

Where k is the number of models, N is the number of epochs, and R_j is the sum of ranks of the j -th model.

If the Friedman test reveals significant differences (i.e., $p < \alpha$), this study conduct the Nemenyi post-hoc test to identify where those differences lie [21]. The critical difference (CD) for a pairwise comparison is computed as (10).

$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}} \quad (10)$$

Where q_α is the critical value from the studentized range distribution based on the number of models and the chosen significance level α . To complement the global analysis with targeted comparisons, the Wilcoxon

signed-rank test [22] is applied between each CondConv model and its corresponding baseline. This test examines whether the median difference of paired performance measures is zero and is defined as (11).

$$W = \sum_{i=1}^n s \operatorname{sign}(d_i) \cdot R_i \quad (11)$$

Where d_i is the difference between paired observations, R_i is the rank of the absolute difference, and n is the number of epochs where $d_i \neq 0$. All statistical tests are executed independently for each magnification level with a significance threshold set at $\alpha = 0.05$, enabling resolution-specific hypothesis validation and minimizing type I errors in multi-resolution evaluation.

3. RESULTS AND DISCUSSION

Two experimental settings were considered: the original imbalanced BreakHis dataset and a balanced variant generated through class-wise sampling. This setup allows assessment of model behavior under both natural and controlled class distributions. All models were trained using identical hyperparameters, with a batch size of 32, the Adam optimizer, a fixed learning rate, and binary cross-entropy loss. Training was conducted for 50–60 epochs to ensure convergence. For CondConv-based models, the number of expert kernels was set to four. All experiments were run on a CUDA-enabled NVIDIA platform.

3.1. Quantitative evaluation across magnification levels and architectures

Figure 4 shows the training and validation behavior of all models across magnification levels. Each subplot in this figure compares the convergence behavior of baseline models (MobileNetV2 and DenseNet121) and their CondConv-enhanced counterparts (CondMobileNetV2 and CondDenseNet121). Solid lines represent training performance, while dashed lines correspond to validation exponential moving average (EMA) smoothing is applied to highlight convergence trends. Clear differences in generalization behavior are observed between baseline architectures and their CondConv variants.

At 40× magnification as shown in Figure 4(a), CondMobileNetV2 converges smoothly, with closely aligned training and validation curves and a final validation accuracy of approximately 91%. MobileNetV2 exhibits moderate overfitting, while DenseNet121 and CondDenseNet121 show larger train–validation gaps and higher variability. These results suggest that CondMobileNetV2 generalizes more consistently at low magnification, where coarse structural features dominate.

At 100× magnification as shown in Figure 4(b), all models show rapid loss reduction during early training. Generalization behavior diverges in later epochs. DenseNet121 achieves strong training accuracy but plateaus at a validation loss above 0.3, indicating overfitting. CondDenseNet121 shows a similar pattern, with validation loss remaining high (≈ 0.4) and a persistent train–validation gap. In contrast, CondMobileNetV2 maintains closer curve alignment and converges to a validation accuracy of approximately 91% with lower loss (≈ 0.28).

At 200× magnification as shown in Figure 4(c), both DenseNet121 and CondMobileNetV2 exceed 92% validation accuracy. CondDenseNet121 continues to exhibit higher validation loss (≈ 0.35). At 400× as shown in Figure 4(d), performance differences narrow across models. DenseNet121 and CondMobileNetV2 achieve the lowest validation losses (≈ 0.22 and ≈ 0.24) and highest accuracies ($\approx 94\%$ and $\approx 93\%$), while CondDenseNet121 converges more smoothly but with a higher final loss (≈ 0.30). These results indicate that the benefit of conditional parameterization decreases at very high magnification.

Test-set results are summarized in Figure 5 using macro-averaged metrics computed from confusion matrices. At 40× magnification as shown in Figure 5(a), CondMobileNetV2 achieves an accuracy of 0.8978 and a macro-F1-score of 0.8807, comparable to DenseNet121 (accuracy 0.89, macro-recall 0.87) while showing a more balanced precision–recall trade-off. At 100× magnification as shown in Figure 5(b), DenseNet121 attains the highest performance (accuracy 0.92, macro-F1 0.91), with CondMobileNetV2 remaining competitive (accuracy 0.9108, macro-F1 0.8902). At 200× magnification as shown in Figure 5(c), CondDenseNet121 achieves the strongest results, reaching an accuracy of 0.9108 and a macro-F1-score of 0.8902, surpassing its baseline. At 400× magnification as shown in Figure 5(d), performance differences narrow across models, with macro-F1-scores around 0.85 and MobileNetV2 slightly outperforming others (accuracy 0.90, macro-F1 0.88).

Under the original imbalanced setting, baseline models—particularly DenseNet121—show notable instability. At 100×, validation accuracy fluctuates between 0.70 and 0.88, while at 200×, validation loss exceeds 0.60 despite low training loss. In contrast, CondMobileNetV2 exhibits smoother convergence and stable macro-F1-scores across magnifications, and CondDenseNet121 shows similar stability at 200×. The smaller performance gap between balanced and imbalanced settings for CondConv-based models indicates reduced sensitivity to class imbalance.

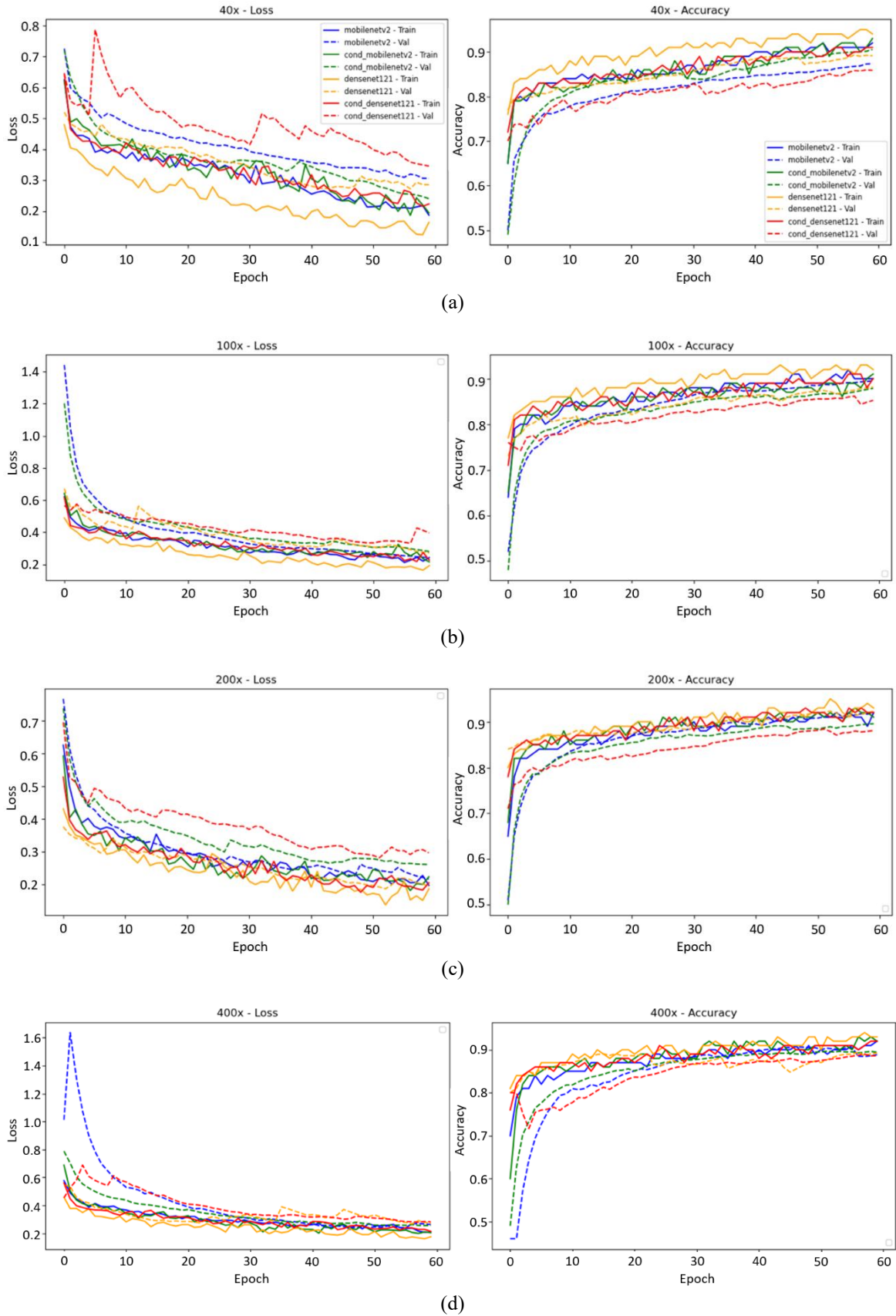


Figure 4. Training and validation loss and accuracy curves across four magnification levels: (a) 40×, (b) 100×, (c) 200×, and (d) 400×

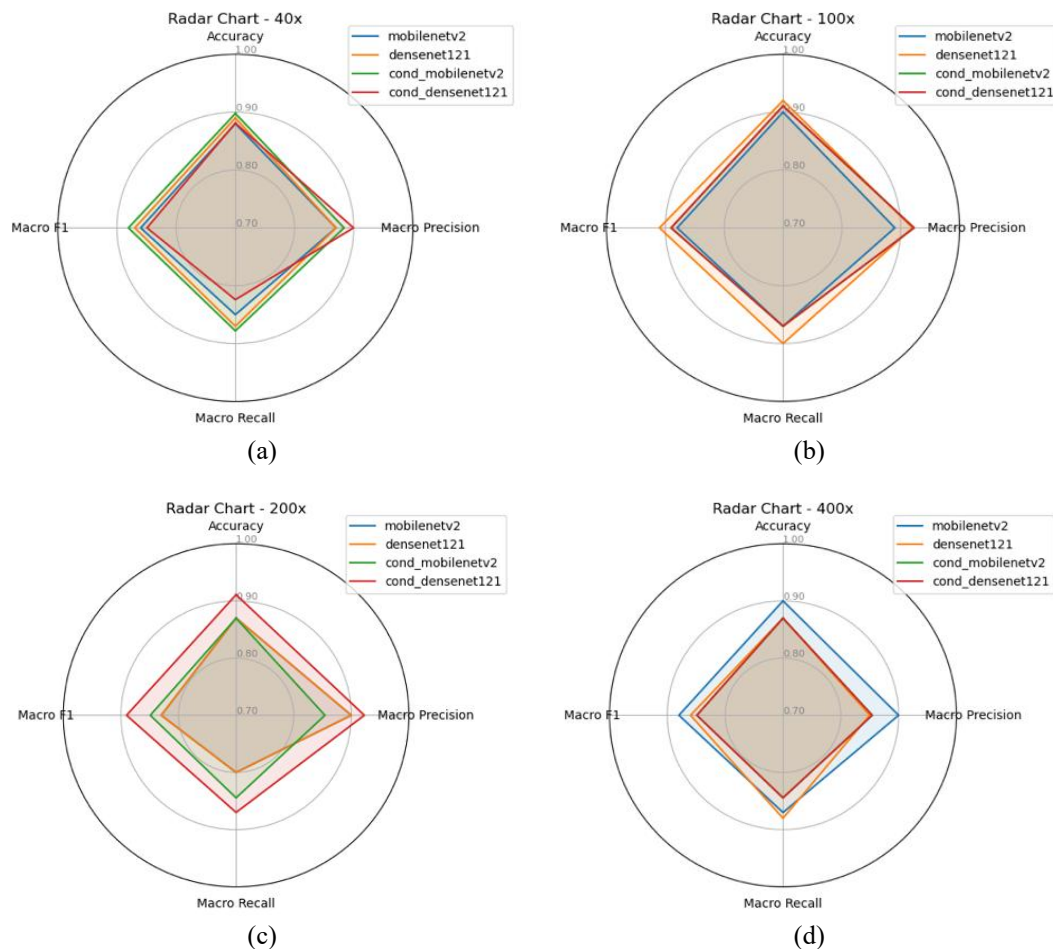


Figure 5. Comparative radar plots of test-set performance metrics across magnification levels: (a) 40× magnification, (b) 100× magnification, (c) 200× magnification, and (d) 400× magnification

3.2. Statistical validation

Statistical significance across models and magnification levels was assessed using non-parametric tests. For each magnification, the Friedman test was applied to compare model rankings across training epochs under the null hypothesis of equivalent performance. Significant differences were observed at 40× ($p < 0.0001$), 200× ($p = 0.0021$), and 400× ($p < 0.0001$). No significant difference was detected at 100× ($p = 0.1948$), indicating similar model behavior at this scale.

For magnifications with significant Friedman results, Nemenyi post-hoc tests were conducted. Pairwise adjusted p-values are shown as matrix heatmaps in Figure 6. Each heatmap in this figure reports the adjusted p-values from the Nemenyi test, comparing all model pairs at 40× as shown in Figure 6(a), 100× as shown in Figure 6(b), 200× as shown in Figure 6(c), and 400× as shown in Figure 6(d). Statistically significant differences ($p < 0.05$) are highlighted in darker shades. At 40×, both CondConv-based models significantly outperformed their corresponding baselines ($p < 0.01$). At 200×, CondDenseNet121 achieved significantly higher performance than both MobileNetV2 and DenseNet121. At 400×, only CondDenseNet121 showed a significant improvement over its baseline ($p = 0.0138$), indicating reduced benefits of conditional parameterization at very high magnification.

To complement these global analyses with targeted comparisons, Wilcoxon signed-rank tests were applied between each CondConv model and its corresponding baseline (Table 1). CondMobileNetV2 showed statistically significant improvements over MobileNetV2 at 40×, 100×, and 200× ($p < 0.001$), but not at 400×. CondDenseNet121 significantly outperformed DenseNet121 across all magnifications, with weaker significance at 400×. Overall, these results support the hypothesis that conditional convolution enhances model adaptability and generalization, particularly at low-to-intermediate magnifications where structural variability is most pronounced.

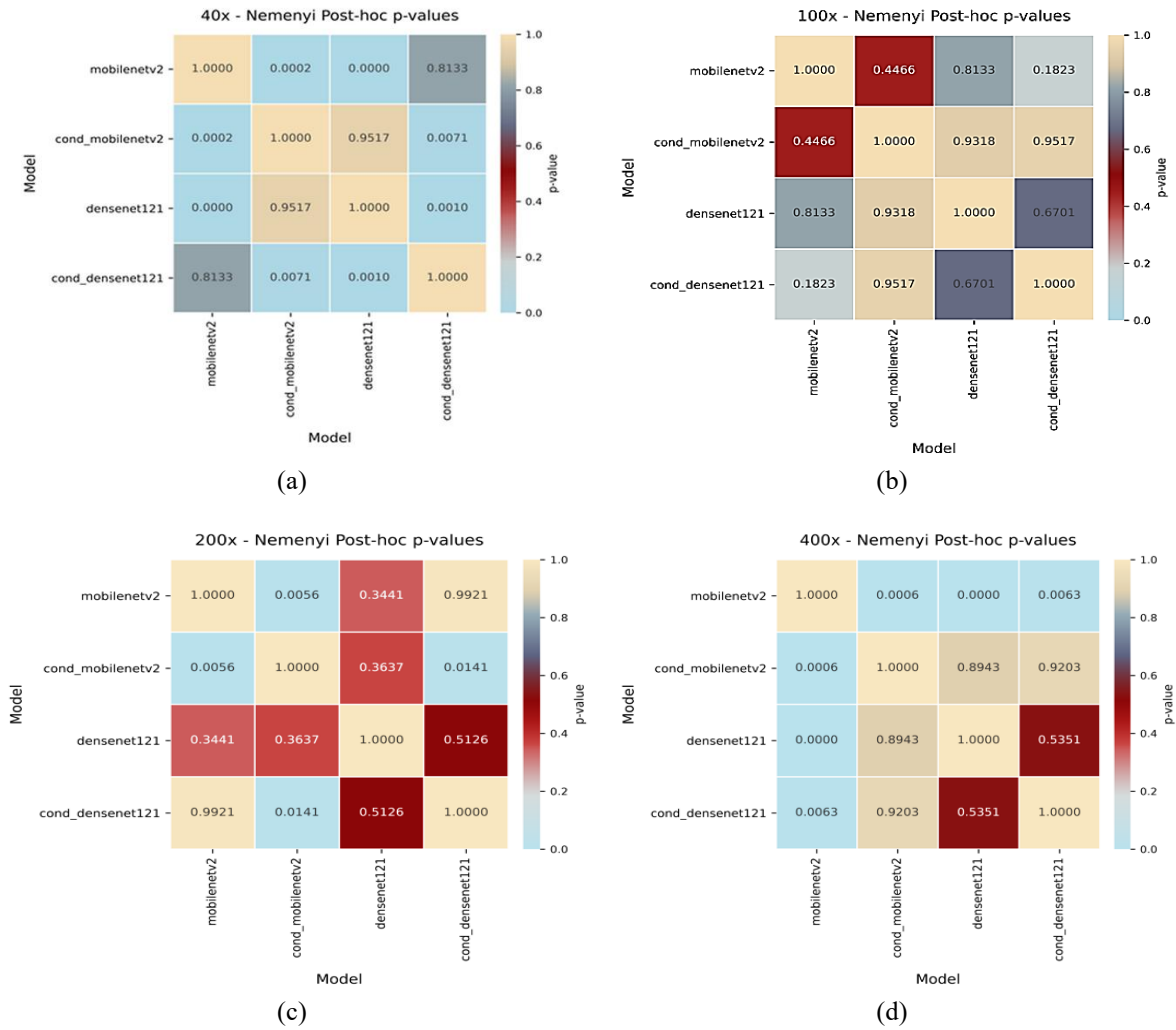


Figure 6. Nemenyi post-hoc pairwise p-value matrices across magnification levels: (a) 40×, (b) 100×, (c) 200×, and (d) 400×

Table 1. Wilcoxon signed-rank test results comparing CondConv models to their baselines across magnifications

Magnification	Pair	W-stat	p-value
40×	cond_mobilenetv2 vs mobilenetv2	246.5	< 1e-4
	cond_densenet121 vs densenet121	251.5	< 1e-4
100×	cond_mobilenetv2 vs mobilenetv2	275.0	0.0004
	cond_densenet121 vs densenet121	405.0	0.0013
200×	cond_mobilenetv2 vs mobilenetv2	359.0	0.0003
	cond_densenet121 vs densenet121	95.0	< 1e-4
400×	cond_mobilenetv2 vs mobilenetv2	689.5	0.5472
	cond_densenet121 vs densenet121	496.5	0.0138

3.3. Discussion

The findings indicate that kernel adaptivity introduced through CondConv influences model behavior in a resolution-dependent manner. Rather than uniformly increasing peak accuracy, CondConv improves generalization stability under magnification variability and class imbalance. At low magnification (40×), where discriminative cues are dominated by global tissue organization and coarse morphology, CondMobileNetV2 exhibits more stable learning dynamics than both baseline architectures and CondDenseNet121. This suggests that input-adaptive kernel selection is particularly beneficial when static filters struggle to generalize across heterogeneous low-resolution appearances. This behavior complements earlier patch-based and fusion approaches that addressed this variability externally [23]. At intermediate magnification (100×), where local cellular detail and emerging architectural context jointly inform

classification, baseline DenseNet121 shows clear overfitting when trained from scratch. In contrast, CondMobileNetV2 demonstrates improved alignment between training and validation behavior, indicating better regulation of feature specialization. This behavior contrasts with transfer learning–based DenseNet121 approaches that report higher peak accuracy at this scale by leveraging pretrained representations [24]. In the absence of transfer learning, CondConv contributes primarily to robustness rather than absolute performance maximization. At higher magnification ($200\times$), deeper architectures begin to benefit from conditional parameterization, as reflected by improved test-set performance in CondDenseNet121. This transition suggests that adaptive filtering becomes more effective once sufficient spatial detail is present. At very high magnification ($400\times$), performance differences narrow across architectures, indicating diminishing returns for kernel adaptivity when discriminative cellular patterns are already explicit and well localized.

Class imbalance further differentiates model behavior. Baseline architectures exhibit pronounced instability under the original BreakHis distribution, consistent with prior studies that relied on data augmentation or ensemble strategies to mitigate imbalance [25]–[27]. In contrast, CondConv-based models demonstrate reduced sensitivity to class skew through smoother learning dynamics rather than inflated accuracy. This indicates that kernel adaptivity provides an intrinsic form of robustness that complements data-level interventions.

Recent attention-based, generative adversarial network (GAN)-driven, and CNN-transformer hybrid frameworks have reported strong performance on BreakHis by explicitly modeling long-range dependencies, selecting discriminative regions, or fusing multi-scale representations [28]–[31]. While effective, these approaches rely on architectural augmentation, ensemble design, or pretraining. In contrast, the present study isolates kernel adaptivity as a single architectural mechanism, clarifying its specific contribution to robustness under magnification variability without relying on external aggregation or attention modules.

3.4. Limitations and future works

This study is limited to a single dataset (BreakHis), which may restrict generalizability across institutions and acquisition settings. Another limitation is that the number of CondConv experts was fixed to four for all architectures and magnification levels due to computational constraints and was not treated as a tunable hyperparameter. Prior work on CondConv indicates that increasing expert capacity, when appropriately regularized, can further improve performance, particularly in lightweight architectures [13]. Consequently, the present results may underestimate the full potential of kernel adaptivity, especially at higher resolutions or in deeper models. Future work may integrate conditional parameterization with attention mechanisms to jointly address kernel adaptivity and feature saliency. Incorporating CondConv into attention-based or CNN-transformer architectures could enable adaptive feature extraction at the kernel level while complementing higher-level modeling of long-range dependencies. In addition, extending the evaluation to multiple histopathological cancer types and multi-institutional datasets will be necessary to assess the generalizability of conditional parameterization across diverse tissue morphologies and imaging conditions.

4. CONCLUSION

This study investigated the impact of CondConv on histopathological breast cancer image classification across multiple magnification levels and under class imbalance. CondConv-enhanced MobileNetV2 and DenseNet121 consistently achieved comparable or improved performance relative to their static counterparts, with the most pronounced gains at $100\times$ and $200\times$ magnifications. These improvements were maintained under imbalanced conditions, suggesting that adaptive kernel parameterization enhances robustness to data heterogeneity and class imbalance. Statistical analyses further supported the significance of these performance gains in several experimental settings. Although the advantages diminished at $400\times$ magnification, likely due to representational saturation, the findings overall validate the effectiveness of conditional convolution for improving feature adaptability and generalization in computational pathology. Nevertheless, this study is limited by the use of a single dataset, two baseline architectures, and a fixed number of CondConv experts constrained by computational resources. Future work will investigate CondConv within transformer-based and hybrid architectures, evaluate larger numbers of experts, and validate the approach on multi-organ, multi-class, and whole-slide image datasets to further assess its scalability and clinical applicability.

FUNDING INFORMATION

This research was funded by the National Center for Scientific and Technical Research of Morocco under the ‘Ph.D.-associate scholarship –PASS’ program.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Wissal El Habti	✓	✓		✓	✓	✓			✓	✓	✓		✓	
Abdellah Azmani	✓			✓			✓			✓		✓		

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

ETHICAL APPROVAL

This study utilized the publicly available BreakHis dataset, which contains anonymized histopathological images. No human participants, animals, or personally identifiable data were involved in this research. As such, ethical approval and informed consent were not required.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [WEH], upon reasonable request.

REFERENCES

- [1] F. Bray *et al.*, "Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 74, no. 3, pp. 229–263, May 2024, doi: 10.3322/caac.21834.
- [2] C. Sitaula and S. Aryal, "Fusion of whole and part features for the classification of histopathological image of breast tissue," *Health Information Science and Systems*, vol. 8, no. 1, Dec. 2020, doi: 10.1007/s13755-020-00131-7.
- [3] Q. Qi *et al.*, "Label-efficient breast cancer histopathological image classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 5, pp. 2108–2116, Sep. 2019, doi: 10.1109/JBHI.2018.2885134.
- [4] V. Bhise, S. S. Rajan, D. F. Sittig, R. O. Morgan, P. Chaudhary, and H. Singh, "Defining and measuring diagnostic uncertainty in medicine: a systematic review," *Journal of General Internal Medicine*, vol. 33, no. 1, pp. 103–115, Jan. 2018, doi: 10.1007/s11606-017-4164-1.
- [5] W. Alyami, A. Kyme, and R. Bourne, "Histological validation of MRI: a review of challenges in registration of imaging and whole-mount histopathology," *Journal of Magnetic Resonance Imaging*, vol. 55, no. 1, pp. 11–22, Jan. 2022, doi: 10.1002/jmri.27409.
- [6] A. S. Panayides *et al.*, "AI in medical imaging informatics: current challenges and future directions," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 1837–1857, Jul. 2020, doi: 10.1109/JBHI.2020.2991043.
- [7] W. El Habti and A. Azmani, "Harnessing multi-source data for AI-driven oncology insights: productivity, trend, and sentiment analysis," *Applied Computer Science*, vol. 21, no. 1, pp. 70–82, Mar. 2025, doi: 10.35784/acs_6670.
- [8] M. H. Wilkerson, K. Lanouette, and R. L. Shareff, "Exploring variability during data preparation: a way to connect data, chance, and context when working with complex public datasets," *Mathematical Thinking and Learning*, vol. 24, no. 4, pp. 312–330, Oct. 2022, doi: 10.1080/10986065.2021.1922838.
- [9] M. Alruwaili and W. Gouda, "Automated breast cancer detection models based on transfer learning," *Sensors*, vol. 22, no. 3, Jan. 2022, doi: 10.3390/s22030876.
- [10] P. Wang, J. Wang, Y. Li, P. Li, L. Li, and M. Jiang, "Automatic classification of breast cancer histopathological images based on deep feature fusion and enhanced routing," *Biomedical Signal Processing and Control*, vol. 65, Mar. 2021, doi: 10.1016/j.bspc.2020.102341.
- [11] M. Toğaçar, K. B. Özkurt, B. Ergen, and Z. Cömert, "BreastNet: a novel convolutional neural network model through histopathological images for the diagnosis of breast cancer," *Physica A: Statistical Mechanics and its Applications*, vol. 545, May 2020, doi: 10.1016/j.physa.2019.123592.
- [12] S. Chattopadhyay, A. Dey, P. K. Singh, and R. Sarkar, "DRDA-Net: dense residual dual-shuffle attention network for breast cancer classification using histopathological images," *Computers in Biology and Medicine*, vol. 145, Jun. 2022, doi: 10.1016/j.combiomed.2022.105437.
- [13] B. Yang, Q. V. Le, G. Bender, and J. Ngiam, "CondConv: conditionally parameterized convolutions for efficient inference," in *33rd International Conference on Neural Information Processing Systems*, 2019, pp. 1307–1318, doi: 10.5555/3454287.3454404.
- [14] J. Zhang, Y. Xie, Y. Xia, and C. Shen, "DoDNet: learning to segment multi-organ and tumors from multiple partially labeled datasets," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2021, pp. 1195–1204, doi: 10.1109/CVPR46437.2021.00125.

- [15] T. Chen *et al.*, “A task decomposing and cell comparing method for cervical lesion cell detection,” *IEEE Transactions on Medical Imaging*, vol. 41, no. 9, pp. 2432–2442, Sep. 2022, doi: 10.1109/TMI.2022.3163171.
- [16] R. Li *et al.*, “DHT-Net: dynamic hierarchical transformer network for liver and tumor segmentation,” *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 7, pp. 3443–3454, Jul. 2023, doi: 10.1109/JBHI.2023.3268218.
- [17] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, “A dataset for breast cancer histopathological image classification,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016, doi: 10.1109/TBME.2015.2496264.
- [18] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: inverted residuals and linear bottlenecks,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.
- [19] M. Friedman, “The use of ranks to avoid the assumption of normality implicit in the analysis of variance,” *Journal of the American Statistical Association*, vol. 32, no. 200, pp. 675–701, Dec. 1937, doi: 10.1080/01621459.1937.10503522.
- [20] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
- [21] A. Benavoli, G. Corani, and F. Mangili, “Should we really use post-hoc tests based on mean-ranks?,” *Journal of Machine Learning Research*, vol. 17, 2016, doi: 10.5555/2946645.2946650.
- [22] F. Wilcoxon, “Individual comparisons by ranking methods,” *Biometrics Bulletin*, vol. 1, no. 6, Dec. 1945, doi: 10.2307/3001968.
- [23] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, “Breast cancer histopathological image classification using convolutional neural networks,” in *2016 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2016, pp. 2560–2567, doi: 10.1109/IJCNN.2016.7727519.
- [24] J. Potsangbam and S. S. Devi, “Classification of breast cancer histopathological images using transfer learning with DenseNet121,” *Procedia Computer Science*, vol. 235, pp. 1990–1997, 2024, doi: 10.1016/j.procs.2024.04.188.
- [25] T. S. Ibrahim, M. S. Saraya, A. I. Saleh, and A. H. Rabie, “Accurate breast cancer diagnosis strategy (BCDS) based on deep learning techniques,” *Neural Computing and Applications*, vol. 37, no. 6, pp. 4617–4650, Feb. 2025, doi: 10.1007/s00521-024-10849-0.
- [26] A. A. Joseph, M. Abdullahi, S. B. Junaide, H. H. Ibrahim, and H. Chiroma, “Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer),” *Intelligent Systems with Applications*, vol. 14, May 2022, doi: 10.1016/j.iswa.2022.200066.
- [27] Z. Gandomkar, P. C. Brennan, and C. M. -Thoms, “MuDeRN: multi-category classification of breast histopathological image using deep residual networks,” *Artificial Intelligence in Medicine*, vol. 88, pp. 14–24, Jun. 2018, doi: 10.1016/j.artmed.2018.04.005.
- [28] S. R. S. Chakravarthy and H. Rajaguru, “MobileDANet integrating transfer learning and dynamic attention for classifying multi-target histopathology images with explainable AI,” *Scientific Reports*, vol. 15, no. 1, Oct. 2025, doi: 10.1038/s41598-025-21360-4.
- [29] R. Man, P. Yang, and B. Xu, “Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks,” *IEEE Access*, vol. 8, pp. 155362–155377, 2020, doi: 10.1109/ACCESS.2020.3019327.
- [30] X. Li, X. Shen, Y. Zhou, X. Wang, and T.-Q. Li, “Classification of breast cancer histopathological images using interleaved DenseNet with SENet (IDSNet),” *PLOS ONE*, vol. 15, no. 5, May 2020, doi: 10.1371/journal.pone.0232127.
- [31] C. C. Ukwuoma, M. A. Hossain, J. K. Jackson, G. U. Nneji, H. N. Monday, and Z. Qin, “Multi-classification of breast cancer lesions in histopathological images using DEEP_pachi: multiple self-attention head,” *Diagnostics*, vol. 12, no. 5, May 2022, doi: 10.3390/diagnostics12051152.

BIOGRAPHIES OF AUTHORS



Wissal El Habti    earned a bachelor's degree in Statistical Engineering, followed by a master's degree in Applied Analysis and Statistical Engineering. She is currently pursuing a Ph.D. in Data Science and Artificial Intelligence at Abdelmalek Essaadi University, within the Intelligent Automation and BioMed Genomics Laboratory, Faculty of Science and Technology, Abdelmalek Essaadi University, Tangier, Morocco. Her research focuses on the application of artificial intelligence in oncology, with an emphasis on deep learning techniques such as convolutional neural networks (CNNs), vision transformers (ViTs), and generative adversarial networks (GANs). By leveraging these advanced models alongside statistical methodologies, she aims to develop innovative tools to improve cancer diagnosis, prediction, and treatment, contributing to the advancement of precision medicine. She can be contacted at email: wissal.elhabti@etu.uae.ac.ma.



Dr. Abdellah Azmani    obtained his doctorate in Industrial Computing, Dynamic Systems Modeling, and Artificial Intelligence at the University of Sciences and Technologies of Lille in 1991. He worked as a professor at the Ecole Centrale de Lille, France and at the Institute of Computer and Industrial Engineering from Lens, France. He is a professor in the Department of Computer Science at Faculty of Sciences and Technologies of Tangier, Abdelmalek Essaadi University, Morocco. He is a founding member and director of the Intelligent Automation and BioMed Genomics Laboratory. He has directed numerous doctoral theses and scientific research projects. Its work focuses particularly on the development of generic and disruptive concepts promoting the integration of artificial intelligence and digital transformation, through the creation of decision-making IT solutions for public and private organizations. He can be contacted at email: a.azmani@uae.ac.ma.