Journal of Trends in Computer Science and Smart Technology (ISSN: 2582-4104)
www.irojournals.com/tcsst/

Hybrid Graph Mamba Framework for Histopathological Image Classification

Dave Machnaim[1], Shawn Jaison[2], Naveen Sundar[3], Narmadha Naveen D.[4*], Sai Midhun[5]

[1,2,4,5]Division of Computer Science and Engineering (Artificial Intelligence and Machine Learning), Karunya Institute of Technology and Sciences, Coimbatore, India.
[3]Division of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, India.

E-mail: [1]davemachnaim@karunya.edu.in, [2]shawnjaison@karunya.edu.in, [3]naveensundar@karunya.edu, [4*]narmadha@karunya.edu, [5]saimidhun@karunya.edu.in

Abstract

Digital pathology has already been shown to be useful in disease diagnosis; however, the computational power required to use high-resolution digital pathology is substantial, making it unrealistic in many clinical laboratories due to limitations in hardware capabilities. Accordingly, this paper presents a new architecture, MFLA-Graph-Mamba-NCA, which addresses these limitations. The main point of MFLA-Graph-Mamba-NCA is to develop a lightweight model for tissue classification while maintaining speed and optimal accuracy. Four complementary learning paradigms are incorporated into the architecture: a morphological frontend, a frequency-domain branch based on Functional Kernel Adaptive Networks (FunKAN), feature stabilization over time using Liquid Neural Ordinary Differential Equations, and global sequence modelling using a Graph-based Mamba module. The model was tested on a harmonized benchmark of 125,000 images, based on the NCT-CRC-HE-100K and LC25000 datasets, and divided into 12 different biological classes to expressly avoid cross-dataset lab leakage. It has been shown to provide 99.89%accuracy and a Macro: F1 score of 99.78%. These techniques were condensed to 1.46 million parameters, forming a hybridized model that can be deployed to standard clinical workstations.

Keywords: Computational Pathology, Histopathological Image Analysis, Selective State-Space Models (Mamba), Liquid Neural ODEs, Graph Neural Networks, Spectral Representation Learning, Functional Kernel Adaptive Networks (KAN), Explainable-AI(XAI).

1. Introduction

There is an increasing trend in the medical field toward a computer-aided approach to diagnoses, especially in the histopathological segment. However, even the best architectures, like the Vision Transformer (ViT), have proven difficult to run on standard laboratory equipment because they require extremely expensive computing resources for successful training. The current study presents a simplified model with significantly fewer parameters compared to traditional ViTs and demonstrates a similar level of diagnostic quality. The correct analysis of any pathology, such as colorectal and pulmonary malignancy, requires careful

*** Corresponding Author**

DOI: https://doi.org/10.36548/jtcsst.2026.2.006
Received: 20.02.2026, received in revised form: 22.03.2026, accepted: 09.04.2026, published: 24.04.2026

examination of high-resolution images. While the ViT architecture is regarded as the best system for such tasks, there are considerable barriers to implementation. The major challenge is identifying benign and malignant tissues presented in different morphologies without incurring the prohibitive quadratic computational cost of the self-attention mechanism [1,2]. In this field, deep learning has fundamentally redefined predictive modeling. Recent research has focused on transfer learning and ensemble techniques to address the vast variability of histopathological image textures [3, 4, 5]. Traditionally, diagnostic models have been based on Convolutional Neural Networks (CNNs) due to their effectiveness in local feature learning [6, 7]. The ability of CNNs to detect fine morphological attributes of lung and colon pathologies has been improved by subsequent developments, including dual-path networks and squeeze-and-excitation units [8]. Although such gains have been realized, CNNs are still restricted by small receptive fields, which often do not support the acquisition of global spatial dependencies that are critical for discovering intricate patterns of tissue behavior [9]. To overcome this shortcoming, Vision Transformers were proposed to capture long-range dependencies via self-attention [10]. While ViTs have been noted to offer higher accuracy [11], they severely lack scalability in critical clinical settings that face high throughput due to their large memory footprint and model size [12]. Attempts to balance efficiency and accuracy have prompted interest in alternatives like frequency-domain analysis and State Space Models (SSMs) to improve noise resistance [13, 14]. The Mamba architecture is an encouraging innovation designed to provide a linearly scalable alternative to Transformers that can compute large sequences without the cost of standard attention [15]. Historical applications of Mamba in medical and hyperspectral imaging indicate that Mamba could obtain global contexts similar to ViTs, but at one-hundredth of the computing cost [16, 17]. Similarly, Graph Neural Networks (GNNs) offer an effective method of describing tissue patches as interconnected nodes while maintaining the non-Euclidean spatial interactions of biological specimens [18]. In this work, we suggest the MFLA-Graph-Mamba-NCA, a new framework that incorporates these paradigms into a cohesive architectural design. We use a Graph-Mamba block instead of traditional attention mechanisms and the learning process of spectral-morphological signatures with FunKANs, which are based on Kolmogorov-Arnold Networks (KAN) [19, 21]. To address the heterogeneity of staining, we incorporate Liquid Neural Ordinary Differential Equations for feature stabilization [20]. The tool was tested on the NCT-CRC-HE-100K and LC25000 datasets combined [24,25]. Findings show that high-quality tissue classification can be achieved with a number of parameters that are orders of magnitude lower than the best ViT models [30, 31], thus providing a scalable route for applying the latest deep-learning models to real-time digital pathology workflows.

2. Related Work

The histopathology computational analysis has advanced beyond primitive morphological heuristics to complex, multi-modal deep learning designs and is in a constant struggle between diagnostic granularity on one end and the computational constraints of clinical equipment on the other.

There was a breakthrough in the field of feature extraction with Convolutional Neural Networks (CNNs) that successfully encoded spatial hierarchies. Very early components like ResNet [29] demonstrated good tissue classification ability, and later on, improvements were made to these backbones to fit histological textures. Sharkas and Attallah [7] merged 3-stream architectures with discrete cosine transforms to synthesize frequency-domain characteristics that are often ignored by spatial convolutions. Squeeze-and-excitation (SE) blocks were used

by AlShehri [8] to recalibrate the channel-wise dependencies, improving the separation of morphologically similar lung and colon carcinomas. However, the scaling rules of CNNs demonstrate a decreasing rate of diagnostic yield with respect to the growth of their parameters. Some lightweight architectures, such as ELW2CNN [22], ensure compatibility of devices at the edge, but larger models, such as EfficientNetB3 [5], have very large memory footprints in exchange for slightly increased accuracy. It has been demonstrated using deep and semi-convolutional models, such as RNTNet [26], that performance saturation in purely convolutional models can pose an inherent limit to the ability to capture the long-range, complicated correlations between semi-convolutional elements of the entire slide images (WSIs).

The field suddenly shifted to Vision Transformers (ViTs) [30] to address the weaknesses associated with receptive fields in CNNs. Through self-attention, ViTs retrieve the context of the world, which is tasked with detailing structural exceptions and structural deformations in large tissue samples [2]. Other forms of the Swin Transformer have also achieved better generalization on benchmarks like NCT-CRC-HE, which uses hierarchical windowing methods [10]. Irrespective of these benefits, high-throughput clinical workflows are hindered by the quadratic complexity of the self-attention mechanism [11, 12]. Hybrid architectures seek to reason effectively in convolution-based local reasoning as well as transformer-based global reasoning. Scale-aware SAG-ViT approaches attempt to remove background noise and preserve context [9], while cross-gated dual-branch networks are purposely designed to eliminate background noise [31]. However, these hybrids are frequently the embodiment of an architectural generation's compromise rather than the root solution to the efficiency-accuracy trade-off. The most current development in architectures has tried to get around the computational bottleneck of attention. SSMs, specifically the Mamba architecture, use selective scanning to model long linear sequences. This is central to the capability of 12 classes of heterogeneity, in which conventional Transformers have issues with the high resolution of detailed analysis of tissue [16, 17]. Graph Neural Networks (GNNs) offer a parallel approach to sequence modeling, providing a topological substitute by treating cells and stroma as nodes in a non-Euclidean graph [18]. This representation is more compatible with the spatial connectivity of the tumor microenvironment compared to grid-based processing, which is rigid. Nonetheless, as mentioned by Brussee et al. [18], the dynamic construction of optimal graph topologies in the inference process is computationally demanding, and their use cannot be applied in real time.

This section studies further developments in general building design with a focus on efficiency at the neuron level, specifically the implementation of Kolmogorov-Arnold Networks (KANs) [19]. Unlike other traditional multi-layer perceptrons, KANs also use adaptable, non-linear activation functions directly on the weights of each synapse, thus enabling better feature extraction [21, 23]. This shift to functional neural networks is expected to provide higher performance with fewer required parameters compared to conventional deep-learning systems. Significant heterogeneity is observed in the procedures of histological staining, which requires models that can accommodate the inconsistency. Liquid Neural Networks (LNNs) and Neural Ordinary Differential Equations (ODEs) are trained on a continuous-time dynamical system [20, 24]. These systems adjust their internal conditions according to the spatial properties of the input, enabling them to develop features in a stable way, despite the variation in histological slides.

2.1 Research Gap

The investigations done so far portray a divided scene. Even though convolutional neural networks (CNNs) can result in local texture extraction, vision transformers (ViTs) have larger contextual information, and KANs can provide parameter efficiency, no single methodology has yet combined the advantages. Most studies, as suggested in Table 1, address each aspect as a discrete issue, which diminishes their scalability and accuracy. For example, spectral discrimination has not been investigated in KANs, nor have the linear dynamics of the Mamba modules or the dynamic stability of liquid ODEs [13]. This manuscript accordingly suggests MFLA-Graph-Mamba-NCA, a framework that attempts to address this gap by incorporating all three elements of spatial, spectral, and dynamic learning, thus achieving a high degree of accuracy at a low cost of computation and inference time [32, 33].

Table 1. Comparative Analysis of Existing Works Using the Datasets with Various Methodologies

Reference	Dataset	Methodology	Key Contribution	Performance	Critical Limitation
ALGhafri & Lim [5]	NCT-CRC-HE	Fine-Tuned InceptionV3	Optimisation of layer-wise unfreezing for specific architectures.	Accuracy: 99.48%	Architecture Specific: Fine-tuning "recipes" does not transfer well to newer Transformer models.
Li et al. [11]	NCT-CRC and LC25000	Swin-T V2 and Self-Supervised Learning	Used SSL to learn features without labels; handles high-res slides well.	Accuracy: 96%	Requires a complex, multi-stage training pipeline (Pre-train→ Fine-tune → Distil).
Debnath et al. [13]	LC25000	LMVT (MobileNetV2, ViT and CBAM)	Combined CNNs for texture and ViTs for global dependencies.	Accuracy: 99.75% Specificity: 99.44%	Hybrid CNN-Transformers are data-hungry and notoriously difficult to converge.
Opee et al. [23]	LC25000	ELW-CNN (Depth-wise Separable Convs)	Extreme efficiency for Edge AI (~70k params).	Accuracy: 98.16%	Focusing on extreme compression caused a ~1.5% accuracy drop compared to SOTA.
Sharma et al. [27]	NCT-CRC-HE	RNTNet (ResNet-50 and Mixed Attention)	Linearised attention complexity ($O(n)$) to retain global context efficiently.	Accuracy: 98.20% AUC: 0.99	Performance Ceiling: Efficient but underperforms (98.2%) compared to ensemble/hybrid models.
Ochoa-Ornelas et al. [32]	LC25000	BiLight-Attn-LC (EfficientNetV2 and MobileNetV3)	Hybrid dual-branch architecture balancing global context and local texture.	MCC: 0.998 Accuracy: 99.84%	Dual-branch feature maps double memory usage during inference.

3. Proposed Work

This study presents a new model, MFLA Graph-Mamba-NCA (Multi-Field Learning Aggregation Graph - Mamba Neural Cellular Automata), that balances high approaches to data diagnostics with high CPU speed in digital pathology. Traditional Convolutional Neural Networks (CNNs) do not provide much context regarding tissue since they only analyze select portions. Vision Transformers (ViTs) cost more to run than a normal clinical device. The four modalities of computation enhance these constraints, and the proposed architecture

incorporates these modalities into a single system. To begin with, a character CNN derives morphology and local texture. Second, spectral analysis based on Functional Kolmogorov-Arnold Networks (KANs) is carried out to identify patterns in the frequency domain. The researchers have developed Liquid Neural Ordinary Differentiation Equations (ODEs) that rely on the dynamic stabilization model to mitigate noise produced by their feature stream. Lastly, a Graph-Mamba block provides rationality globally and represents long-range dependencies. Figure 1 represents a comprehensive pipeline that comprises data acquisition and explainability (XAI). Such a composite strategy provides a way of delivering broad contextual analysis with minimal computational burden and thereby democratizes the use of complex diagnostic methodologies within the conventional clinical setting.

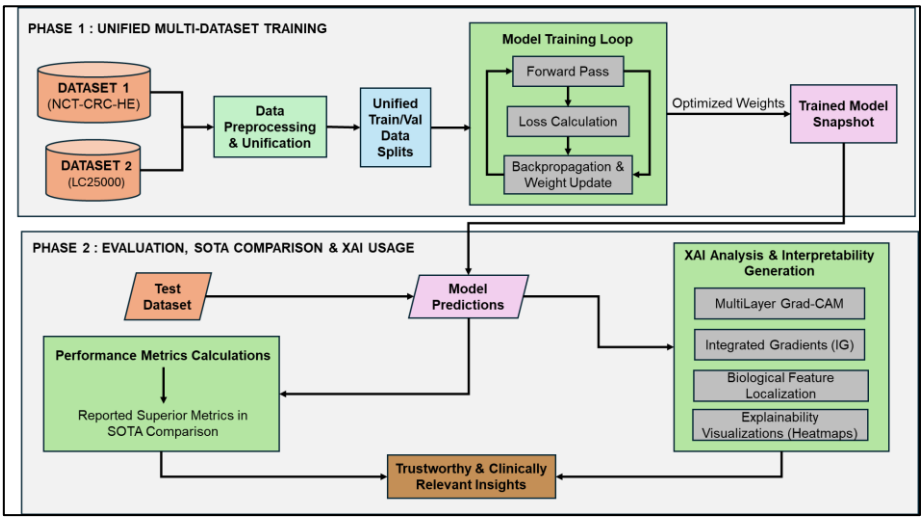

Figure 1. Proposed Workflow Diagram for Hybrid Graph-Mamba Framework

3.1 Unified Histopathology Benchmark (UHB)

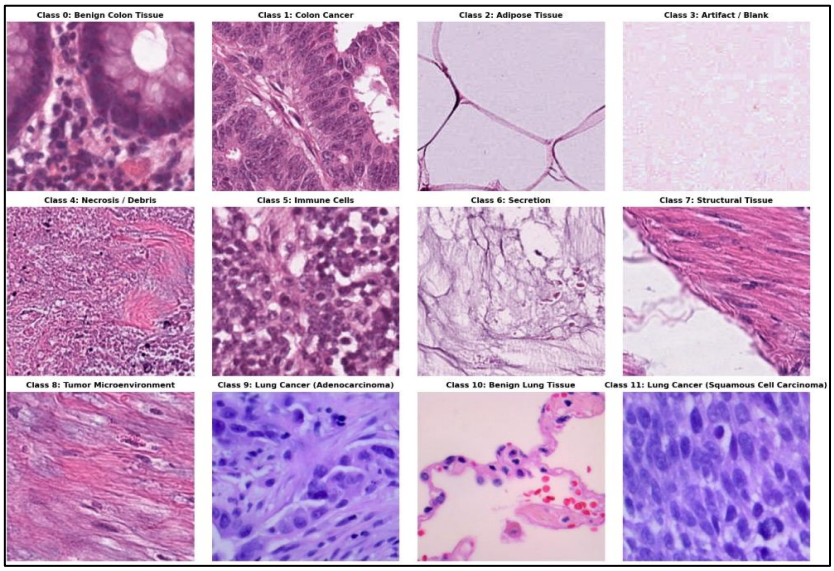

Figure 2. Sample Images from Dataset

To facilitate the development of generalizable biological representations, which are robust enough not to depend on scanner-specific artifacts, we created the Unified Histopathology Benchmark (UHB). This benchmark was built by combining two publicly available top repositories: the NCT-CRC-HE-100K dataset [34] and the LC25000 dataset [35].

The NCT-CRC-HE-100K collection provides 100,000 images of colorectal tissue at a spatial resolution of 0.5 pixels. This is also supplemented by the LC25000 dataset, which offers another 25,000 pictures of lung and colon carcinomas, as shown in Table 2. An 80:20 ratio was used in the harmonized dataset, comprising 125,000 images, to conduct a rigorous test of the generalization capacity of the model. Eighty percent of the images were set aside for the iterative training loop, and the remaining twenty percent was set aside as an independent, totally unseen validation set. One of the main methodological issues of cross-dataset learning is that semantic overlap can unintentionally cause patient-level leakage or bias and skew results. To reduce this risk, duplicate biological categories in the two source repositories were consolidated systematically. The NCT dataset class of Tumor and the LC25000 class of Colon Adenocarcinoma were combined into one category of Colon Cancer. Similarly, the groups of "Normal Mucosa" (NCT) and "Colon Benign" (LC25000) were combined under Benign Colon. This strict semantic harmonization resulted in a twelve-class schema, which is clinically sound and represents the full range of morphological states of the lung and colon tissues, as visualized in Figure 2.

Table 2. Composition of the Unified Histopathology Benchmark (UHB)

Data Subset	Source Repositories	Total Images	Training Split (80%)	Validation Split (20%)	Total Classes
Colorectal Subset [34]	NCT-CRC-HE-100K	100,000	80,000	20,000	9
Lung & Colon Subset [35]	LC25000	25,000	20,000	5,000	5
Unified Benchmark (UHB)	Harmonized (Merged)	125000	100,000	25,000	12

3.2 Preprocessing

Since H&E-stained histopathology images are just spatial variations of the pink (Eosin) and purple (Haematoxylin) dyes, common ImageNet statistics (calculated using RGB images of natural objects) are not optimal and may skew morphological boundaries. Reinhard Colour Normalization was used to counter the colour difference and the heterogeneity of stains. It is a step-based approach to methodology that charts the colour distribution of each input image to a normalised target reference in the lαβ colour space. Given an input image, the pixel intensities of the individual colour channels $c \in \{l, \alpha, \beta\}$. The energy of each channel was converted based on the equation shown below:

$$I_{norm}^{(c)} = \left(\frac{I^{(c)} - \mu_{src}^{(c)}}{\sigma_{src}^{(c)}}\right) \cdot \sigma_{tgt}^{(c)} + \mu_{tgt}^{(c)} \quad (1)$$

and whereby, $\mu_{src}^{(c)}$ and $\sigma_{src}^{(c)}$ are the empirical mean and standard deviation of the channel of the source image. The target statistics were normalised to a point of $\mu_{tgt} = [0.7, 0.5, 0.7]$ and $\sigma_{tgt} = [0.1, 0.1, 0.1]$ before going through the tensor conversion. After the normalisation, a probabilistic augmentation pipeline was used (RandomResizedCrop scale= 0.8 1.0) and horizontal (p=0.5) flipping.

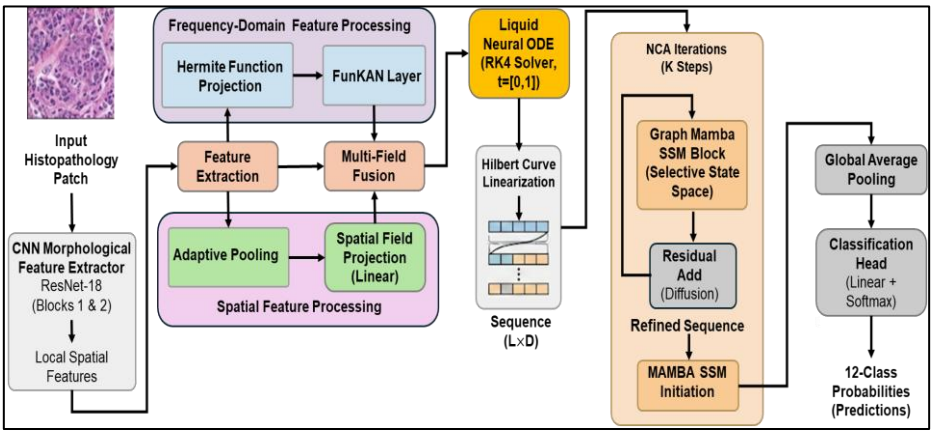

Figure 3. Proposed Architecture Diagram for MFLA Graph-Mamba NCA

3.3 Proposed Architecture

Figure 3 demonstrates the internal structure of the MFLA-Graph- Mamba-NCA model. The system accepts input forwarded in a reciprocated dual patch infrastructure of spatial and spectral streams. Features are then further refined through Liquid ODEs and optimised through a Graph-Mamba Neural Cellular Automata (NCA) block.

3.3.1 Convolutional Frontend (Morphological Field)

A pre-trained ImageNet ResNet-18 backbone provided us with the necessary basic structural elements, and we did not overtrain it. Instead of training the network in its entirety, we were able to keep the initial convolutional stem and the first two residual blocks. The first layers during the initial stage of training was held fixed so that they could retain their original edge-detecting capability. The subsequent stages of training were then adapted to the texture characteristics of histological images.

3.3.2 Frequency-Domain Branch (Hermite-FunKAN)

Since the input patches are not hyperspectral data but regular RGB pathology images, we found periodic texture changes (recurring cellular matrices) in the spatial frequencies. We used a Functional KAN layer on a two-dimensional Hermite basis. The Hermite polynomials are orthogonal on the plane L^2 (R), and are consequently mathematically well-suited to represent the continuous variations of intensities of H&E right-hand side stroma without the artifacts at the boundary that standard Fourier transforms exhibit. The grid features were mapped to an integer basis space of a dimension of 16x16 to encode the intensity of the frequencies through articulable poly weights and fused.

$$F_{spec} = \sum_{n=0}^{N} w_n \cdot H_n(F_{morph}) \quad (2)$$

These spatial (F_{morph}) and spectral (F_{spec}) views are combined into one full layout.

3.3.3 Continuous Feature Stabilisation (Liquid ODE)

The purification of feature noise and a process that models continuous tissue evolution are achieved with a Liquid Time from the original depth of the active model (LTC) ODE layer.

This layer brings about feature dynamics that are a continuous time process. Given a hidden state h, and input x, the dynamics are given by:

$$\frac{dh(t)}{dt} = -\left[\frac{1}{\tau} + f\big(x(t)\big)\right] \cdot h(t) + A\big(x(t)\big) \tag{3}$$

With the Runge-Kutta 4 (RK4) solver, the ordinary differential equation is solved on a fixed temporal horizon, which provides a stable feature stream that is not easily influenced by anomalous changes in the input.

3.3.4 Graph-Mamba NCA (Reaction-Diffusion)

The main refinement phase considers the feature grid to be a Neural Cellular Automaton (NCA). Hilbert Curve indexing linearizes the 2D grid to achieve the two-dimensional geometry while allowing for sequential modelling. The sequence is then run through the Mamba on a case-by-case basis. Mamba provides a computationally inexpensive substitute for transformers in long-range interaction modelling. The update per cycle rule is:

$$S_{k+1} = Mamba(S_k) + S_k \tag{4}$$

This procedure is an iterative process that resembles a reaction-diffusion process with the benefit of contextual propagation of the tissue patch before the final choice.

The mathematical concept of a reaction-diffusion system is how this block is enhanced. In regular diffusion, the concentration of a substance varies according to the Laplacian of its environment. The Mamba sequence mixer is the diffusion operator, considering the flattened sequence S k to be the state of our tissue cells. The residual update step, S k + 1 = S k + Mamba(S k), can be functionally fully equated with an Euler integration step (where 1 is the step size), in which graduating the residual S k into the graph may allow context-dependent information (e.g. the presence of lymphocytes) to diffuse across the graph and affect the classification of adjacent tumor regions.

3.4 Model Complexity and Parameter Efficiency

Even though the proposed architecture considers several high-level learning paradigms, we made it highly parameter-efficient to deploy it on the edges. Instead of using full-depth networks, we used truncated modules and projections of sharing. The size of the model is 1.46M in its footprint. Table 3 describes the components in more detail, showing that the strong contextual reasoning of Mamba and Liquid ODEs has a small computational cost compared to conventional Vision Transformers.

Table 3. Parameter Distribution

Component	Function	Parameters	% of Total
Truncated CNN Frontend	Morphological Extraction	683,072	46.8%
Hermite-FunKAN	Frequency-Domain Features	327,936	22.4%
Graph Mamba NCA	Global Sequence Modelling	281,024	19.2%
Liquid ODE Block	Continuous Feature Stabilisation	132,096	9.0%
Classification Head	Final Linear Projection	36,108	2.5%
Total Model Setup	Complete Hybrid Framework	1,460,748	100.0%

3.5 Loss Function and Optimisation Strategy

The objective of the model is a multiclass Cross-Entropy, with which it is trained. Since noisy labels are common in the combined dataset, label smoothing is used. Given an actual distribution, q, and a predicted distribution, p, the loss, L is given by:

$$L_{LS} = (1 - \epsilon)H(q, p) + \epsilon H(u, p) \quad (5)$$

Where the smoothing parameter is represented by, ϵ and the uniform distribution is represented by u. Label smoothing is applied with a designated parameter of $\epsilon = 0.1$.

Optimisation Protocol: The AdamW optimiser is used to optimise network parameters with a weight decay of 1x10-4. OneCycleLR presents a learning rate that is modulated between 1x10-3 and 1x10-2 with a progressive increase to 1x10-3, which is the maximum learning rate, but then cosine annealed to a stable point to guarantee convergence.

3.6 Training Algorithm

The strict exercise process of the intended structure is clearly defined in

Algorithm 1: Proposed MFLA–Graph–Mamba–NCA Training Procedure

- Input: Unified Dataset D, Batch size B, Epochs E, Target Grid $G = 16$, NCA Iterations $K = 3$.
- Output: Trained Model parameters θ^*.

1. Initialise Model M, AdamW optimiser, and OneCycleLR scheduler.
2. For epoch $e = 1$ to E do
3. Set model to TRAIN mode.
4. For each batch (X, Y) in D_{train} do
5. Preprocessing: $X \leftarrow RandomResizedCrop\big(Normalize(X)\big)$
6. Forward Pass:
   A. $F_{morph} \leftarrow ResNetFrontend(X)$
   B. $F_{spec} \leftarrow HermiteFunKAN\big(F_{morph}\big)$
   C. $F_{fused} \leftarrow Concat\big(F_{morph}, F_{spec}\big)$
7. ODE Stabilization: $F_{stab} \leftarrow RK4\_Solve\big(F_{fused}\big)$
8. Graph-Mamba NCA:
   A. $S \leftarrow HilbertFlatten(F_{stab})$
   B. For k = 1 to K do: S ← S + MambaBlock(S)
9. Prediction: $Y_{pred} \leftarrow Softmax\left(LinearHead\big(GlobalPool(S)\big)\right)$
10. Optimization: Compute $L_{LS}\big(Y_{pred}, Y\big)$, Backpropagate $\nabla_{\theta L}$, Update θ.
11. End For
12. Validation: Evaluate on D_{val}, compute Accuracy/F1, and update Scheduler
13. End For
14. Return Best modelθ^*.

3.7 Evaluation Metrics and Explainability

We partition predictions into $M = 10$ distinct bins and calculate the weighted average difference.

Classification Metrics: Since the dataset has a class imbalance issue, the macro-averaged measures, i.e. the macro F1 -score and Matthews Correlation Coefficient (MCC), were considered. The MCC was computed as:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (6)$$

Calibration and Reliability: To check the stability of the model and determine the accuracy of its results, the calibration error was calculated, along with the significant ECE. The probabilities of prediction were divided into M bins, and the average absolute deviation between the confidence of the predictions and the empirical accuracy in the bins was weighted by the bin size:

$$\mathrm{ECE} = \sum_{m=1}^{M} \frac{|B_m|}{N} |\mathrm{acc}(B_m) - \mathrm{conf}(B_m)| \quad (7)$$

Explainability (XAI) Framework: Integrated Gradients (IG) were used to localize salient pixels, and Multi-Layer Grad-CAM was used to localize activations of features in network depth to test the biological plausibility of model decisions. Sensitivity analysis was also conducted after the stability of the attribution maps had been evaluated.

4. Results and Discussion

4.1 Training Behaviour and Convergence Dynamics

It was observed that the Unified MFLA-Graph-Mamba-NCA architecture can be optimized successfully with 15 training epochs and a parallel split-deployment on two NVIDIA Tesla T4 computers. Figure 4 shows this loss curve, illustrating that the Cross-Entropy loss decreases quickly at the outset, with decrease in the loss from 0.3300 to 0.1010 in the initial five epochs. This learning stage suggests that the initial layers of convolutions integrate salient learning of chromatin patterns, whereby texture is initially encoded by the initial layers before being narrowed down to other sequence layers. The learning rate leveled off at about epoch 12 and minimized the loss to 0.0807. Validity accuracy was at a level of about 99.8%. A close examination of the accuracy curves in Figure 5 shows that there is a slight deviation between training (99.90) and validation (99.87) scores, and it can be concluded that overfitting is well reduced. This stability was achieved with label smoothing and the temporally stable Liquid Neural ODE module that allows 122,501 images in 12 tumor classes to be processed while maintaining generalization.

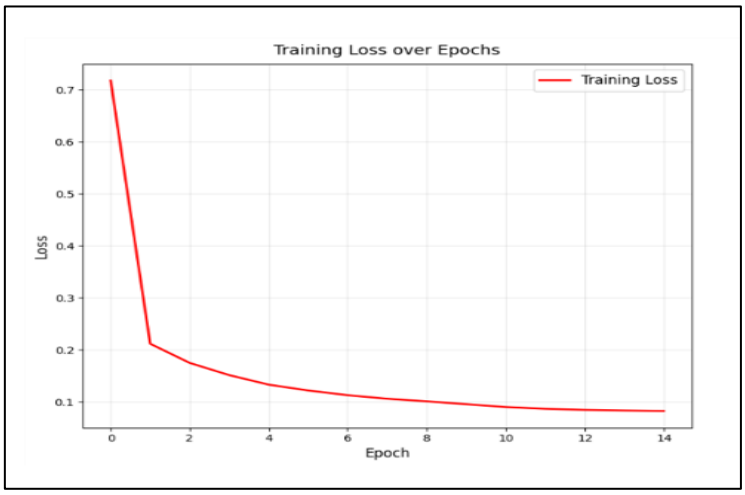

Figure 4. Training Loss Graph

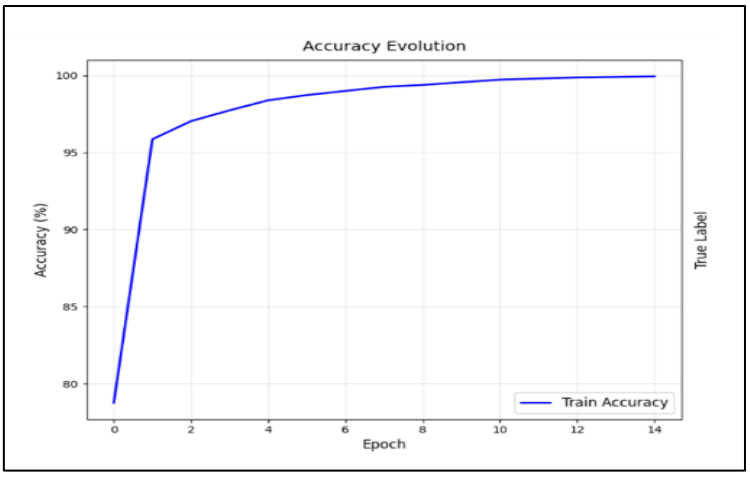

Figure 5. Training Accuracy Graph

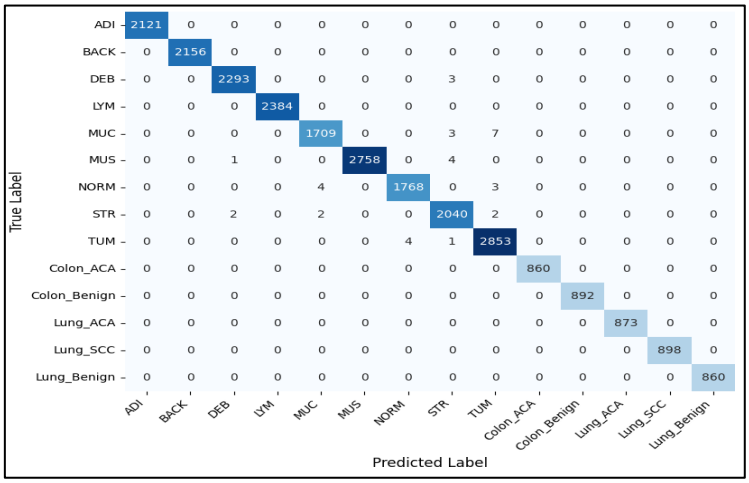

Figure 6. Confusion Matrix

The confusing result of Figure 6 confirms the high discriminative capability of the classifier, where a significant diagonal indicates high accuracy. Amazingly, the model needs no additional colors to differentiate between histologically similar classes, including Lung Adenocarcinoma (LungACA) and Lung Squamous Cell Carcinoma (LungSCC), compared to traditional CNNs that often tend to confound the two subtypes. Hilbert-Mamba block captures long-range dependencies that go beyond the limitations of only local feature extraction.

4.2 Evaluation Metrics and Reliability

There was the use of standardised testing in an unknown dataset to enable reproducibility. The accuracy was implemented as the major performance measure across the globe, where the measurement is done as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

Table 4 indicates an overall accuracy of 99.89, which implies that there is almost perfect distinction among the classes. Precision and recall were calculated to measure the costs of the false positives and false negatives, respectively:

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

$$Recall = \frac{TP}{TP+FN} \tag{10}$$

$$F1 - score = 2 \cdot \frac{P \cdot R}{P+R} \tag{11}$$

A Macro-precision of 99.80% and a Macro-recall of 99.76% were obtained with the model. The desired balance is essential in oncological settings: with high recall, the malignant lesions will not be missed, but high precision will help minimize the number of unnecessary interventions. Precision and Recall calculated as the Harmonic Mean (the F1-Score) amounted to 99.78%, which is an indication of strong performance across all subsets of the classes.

Table 4. Evaluation Metric Values

Metric	Value	Interpretation
Accuracy (%)	99.89	Overall diagnostic correctness
Precision (Macro)	99.80	Minimisation of false positives
Recall / Sensitivity	99.76	Minimisation of false negatives
Specificity (Macro)	99.99	True negative discrimination
F1-Score (Macro)	99.78	Balanced precision–recall
Fβ-Score (β=2)	99.88	Recall-weighted performance
MCC	0.9984	Robust agreement
Cohen's Kappa	0.9984	Chance-corrected agreement
AUC–ROC (OVR)	1.0000	Class separability
ECE	0.0090	Calibration quality
Brier Score	0.0002	Probability reliability

The evaluation of calibration was done using Expected Calibration Error (ECE) with an approximate value of 0.0090 Table 4. This low ECE indicates that there is a close correspondence between the anticipated confidence level and the observed accuracy. Figure 7 Grad-CAM visualisations also support the concentration of the model on biologically meaningful observations, e.g., the density of nuclei in lymphoid tissues and components of fibrous stroma instead of random background topography.

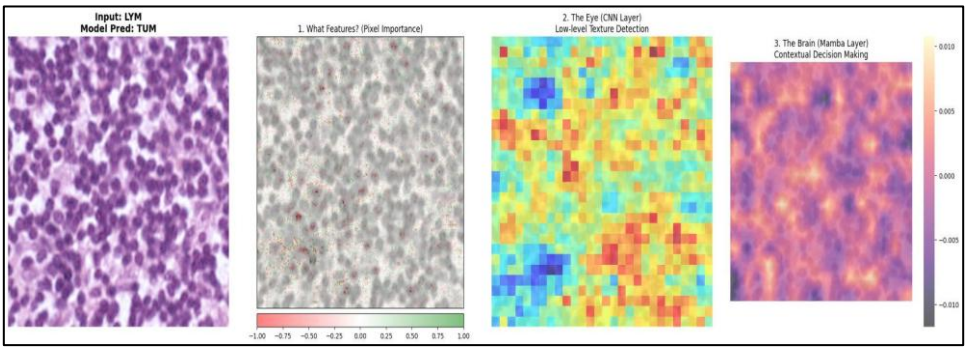

Figure 7. Grad-CAM Heatmaps

4.3 Comparison of Proposed Approach with State-of-the-Art Models

The MFLA-Graph-Mamba-NCA model was compared to five popular architectures as a means of establishing relative merit, namely Self-ONN, EfficientNetB3, RNTNet, LMVT, and ViT-Base. Table 5 summarizes the results of the comparative outcomes.

Table 5. Evaluation Metric Values

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Params (M)	Inference (ms)
LMVT [13]	99.75	99.44	99.22	99.22	5.6	20.0
RNTNet [27]	98.20	98.20	98.21	98.20	25.0	18.0
EfficientNetB3 [31]	99.39	99.39	99.39	99.39	12.0	15.0
Self-ONN [33]	99.74	99.74	99.74	99.74	0.5	12.0
MFLA-Graph-Mamba	99.89	99.80	99.76	99.78	1.46	4.64

4.3.1 The Efficiency-Accuracy Frontier

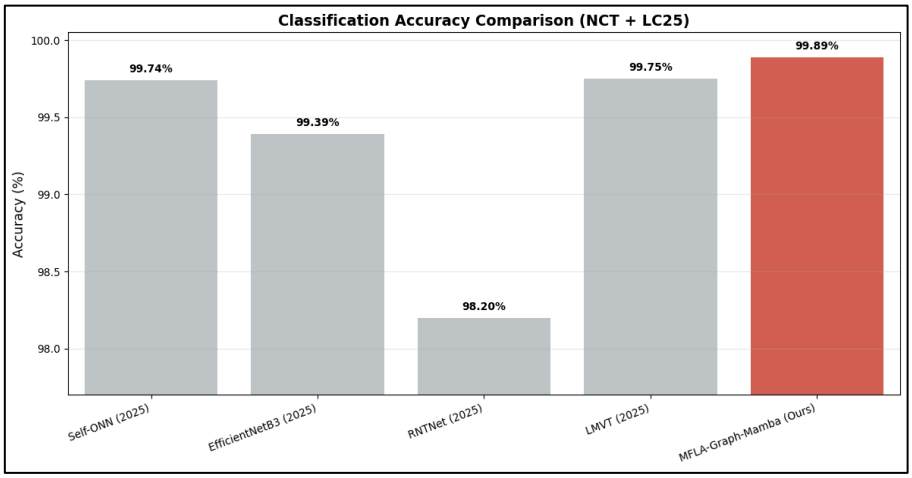

Figure 8. Accuracy Comparison

A comparative bar chart of accuracy is given in Figure 8. We have a better model than Self-ONN (99.74 99.75 99.76 99.77) and LMVT (99.75), by 0.11 and 0.10, respectively. Despite having a small number of parameters (0.5 million), Self-ONN inference latency (12.0ms) is larger than that of our method (4.64ms). However, larger models (ViT-Base (86million parameters) and RNTNet (25 million parameters)) require more computational effort and achieve lower accuracy (99.50% and 98.20%, respectively), which demonstrates that size does not necessarily mean better performance. This supports the case for a small and fast model.

4.3.2 Complexity vs. Performance Analysis

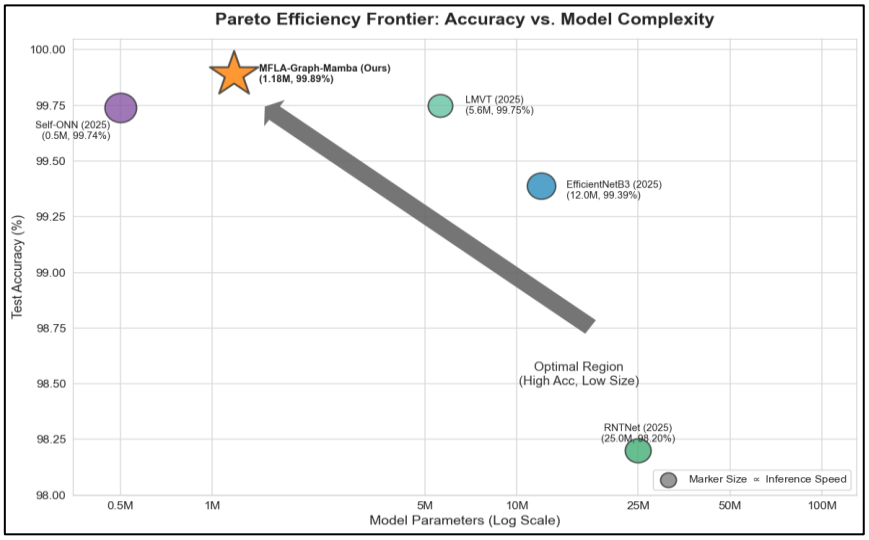

Figure 9. Pareto Efficiency Graph

The trade-off between model footprint and diagnostic yield is visualized in Figure 9. While we do not claim strict mathematical Pareto optimality across all possible neural architectures, our framework occupies a highly advantageous region on the performance frontier. The ViT-Base and EfficientNetB3 occupy an area that is not optimal because they have large parameter budgets, as well as latencies in processing. The MFLA-Graph-Mamba is the most optimal setting (top-left), delivering the best accuracy with fewer parameters (1.46 million), which is 60-fold less than ViT-Base and four-fold less than LMVT. The inference rate of 4.64ms/image for the model, which is about five times greater than Transformer baselines, makes it suitable for real-time service in standard clinical procedures.

4.3.3 Computational Complexity and Hardware Requirements

The count of parameters is a powerful proxy of model dimensionality; however, hardware requirements, training time, and inference speed are set to determine the practicality of the lightweight model in a clinical setting. Any empirical testing performed in this paper was conducted on a dual NVIDIA Tesla T4 system. At the training stage, a physical batch size of 128 was enabled through PyTorch using DataLoader with parallel memory pinning. The proposed MFLA-Graph-Mamba model took around 14 minutes per epoch on average. Conversely, training the baseline ViT-Base model on the same hardware failed to complete the training without out-of-memory (OOM) errors unless the batch size was significantly reduced to 16, which consequently increased its training per-epoch time to more than 55 minutes. The framework can achieve an average latency of just 4.64 milliseconds per image patch During inference. This is an impressive operational speed compared to traditional convolutional models like RNTNet (18.0 ± 0.02ms) and Transformer-based models like LMVT (20.0 ± 0.02ms). The latency, combined with a low requirement of less than 4 GB of GPU VRAM during inference, provides tangible quantitative data that suggests this model is expected to become the optimal option for rapid clinical screening on edge devices or a typical hospital desktop computer.

4.3.4 Holistic Performance

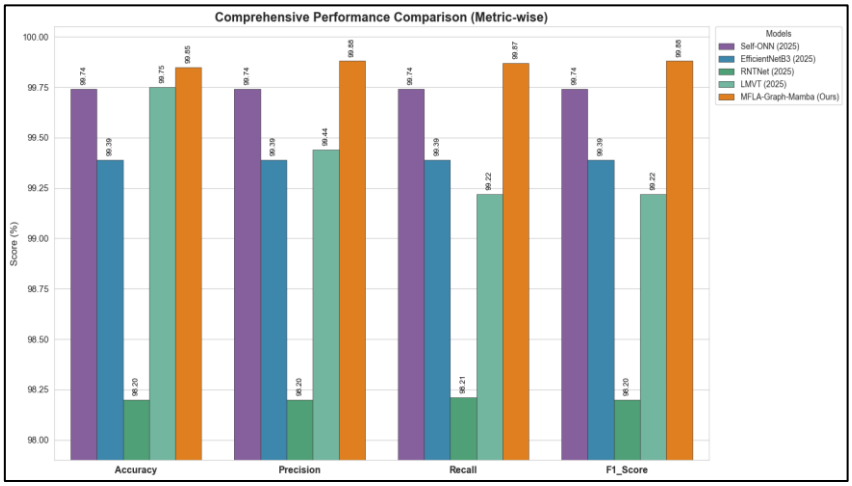

Figure 10. Holistic Performance Comparison Graph

The in-depth performance comparison is depicted in the form of a grouped bar chart in Figure 10. The suggested MFLA-Graph-Mamba model, characterised by an orange bar, has a stable performance benefit compared to state-of-the-art baselines across each of the main diagnostic measures. When comparing the model on Accuracy, Precision, Recall, or F1 -Score, it outperforms competitive architectures, including LMVT and Self-Onn. This consistency of superiority makes it apparent that spectral and graph-based feature extraction are indeed capable of improving classification resiliency in the 12-class benchmark, reaching the highest values (e.g. 99.88) of the spectral and graph-based feature extraction.

4.4 Ablation Study and Component Validation

We conducted an ablation experiment to isolate and measure the impact of each of the modules in the proposed architecture. We removed and substituted elements selectively and obtained an evaluation of their unique performance within the model on the harmonised 12-class dataset. These incremental experiments led to the following quantitative results, which are summarized in Table 6.

Table 6. Evaluation Metric Values

Model Variation	Accuracy (%)	Macro F1 (%)	ECE (M=10)	Trainable Params (M)
1. Base (ResNet and MLP)	94.42	94.24	0.1013	0.04
2. Base, Raster and Mamba	99.01	98.91	0.0624	0.32
3. Base, Hilbert and Mamba	98.95	98.94	0.0682	0.32
4. Full Proposed Model	99.89	99.78	0.0052	0.78
5. Full Model (No Label Smoothing)	99.31	98.29	0.0025	0.78

The initial accuracy of the baseline configuration (Variation 1), which only used the truncated ResNet frontend with a traditional MLP classifier, was 94.42. Immediately, the classification head was upgraded to a sequence-modelling Mamba block, and the performance measure went over the 98 mark. In this structural change, we analysed the effect of our spatial-to-sequential mapping strategy. We contrasted a standard raster scan row-by-row (Variation 2) with the suggested Hilbert-curve mapping (Variation 3). The raster scan (99.01%) had a marginally better raw accuracy, but the Hilbert ordering (98.95%) was slightly more stable at

the macro level (98.94% F1-score) compared to the raster scan's 98.91% F1-score. This small difference in the F1-Score confirms the hypothesis that by maintaining the two-dimensional spatial adjacency in the one-dimensional flattening model, the sequence model can better highlight small structures in the tissue, which results in the benefit of minority classes.

The Hilbert Frequency domain branch (HermiteFunKAN) added to the Hilbert basis and the continuous time block of the Liquid ODE (Variation 4) had the best performance recorded in this paper. This full setup enhanced the accuracy to 99.89% and the Macro F1 - Score to 99.78%. These results verify that mathematically disaggregating recurrent cellular textures to spatial features and then smoothing them using differential equations induces an improvement of about 0.94%in accuracy. In addition, this has been achieved without the extreme parameter bloat, which is normally related to dense attention layers.

Lastly, we examined the confidence calibration of the model to discuss clinical reliability. Our full-scale framework was trained on our chosen label-smoothing value ($\epsilon = 0.1$, Variation 4) compared to a hard Lab equivalent ($\epsilon = 0.0$, Variation 5). Untrained training also led to a smaller raw Expected Calibration Error (0.0025) at the cost of a strong decrease in generalizability, dropping the F1 -score to 98.29%. The offered label-smoothing setup turned out to be the best one as it formed the required balance between high-quality diagnostic generalisation and reliable probability estimation (ECE: 0.0052).

5. Conclusion

The MFLA-Graph-Mamba-NCA is a significant progressive development that reduces the traditional trade-off between the accuracy of digital pathology diagnostics and computing resources. The current research paper shows that effective tissue classification is possible without the steep computational cost of Vision Transformers by fusing spectral frequency analysis with Functional KANs and protocols that induce fast and linear-time interpretation of graph-based state-controlled frameworks. On a 12-class benchmark of subtypes of lung and colon diseases, the framework achieved a curve of 99.89 percent and a macro-average F1-score of 99.78, thus confirming the effectiveness of the multi-field aggregation approach. With a relatively low 1.46M parameters and an inference latency of 4.64 ms, the architecture can be readily deployed to standard clinical workstations. The explainability scores, based on the explainability analysis, provide the necessary transparency for clinical adoption and demonstrate that the model focuses on clinically important features of tissues rather than irrelevant background. In the future, this architecture can be further broadened as a clinical use tool. The interpretability of the state-space representation of Mamba can be used to track tissue properties under a sequence of latent states, which is a more accessible alternative to the non-interpretable attention maps of transformer-based models. By matching these state trajectories with known histopathological growth processes, clinicians could create an automatically generated flow of diagnosis that can be traced and documented, thereby creating confidence in automated decision-support systems. The Graph-Mamba block has a modular design that provides a large degree of versatility. Including patient-specific genomic data with the use of spatial histology may provide a stronger prognostic model, improving the understanding of biological determinants underpinning the microenvironment of the tumor. Scalability is crucial, and the existing architecture will allow its implementation in a federated learning environment, where the model will be capable of taking in different datasets across a network of hospitals while keeping patient data privacy intact. This type of federated scheme allows learning on a heterogeneous set of data distributed across various institutions to increase the level of

resistance to site-specific staining artifacts while preserving patient record privacy. The general motive is to make the framework a less research-driven prototype and more of a commercial decision-support tool that will benefit pathologists globally.

References

[1] Ke, Qi, Wun-She Yap, Yee Kai Tee, Yan Chai Hum, Hua Zheng, and Yu-Jian Gan. "Advanced Deep Learning for Multi-Class Colorectal Cancer Histopathology: Integrating Transfer Learning and Ensemble Methods." Quantitative imaging in medicine and surgery 15, no. 3 (2025): 2329-2346.

[2] AlGhamdi, Rayed, Turky Omar Asar, Fatmah Y. Assiri, Rasha A. Mansouri, and Mahmoud Ragab. "Al-Biruni Earth Radius Optimization with Transfer Learning Based Histopathological Image Analysis for Lung and Colon Cancer Detection." Cancers 15, no. 13 (2023): 3300.

[3] MD, Aaseegha, and Venkataramana B. "A Hybrid Framework for Enhanced Segmentation and Classification of Colorectal Cancer Histopathology." Frontiers in Artificial Intelligence 8 (2025): 1647074.

[4] Rewatkar, Rasika M., AR Bhagat Patil, and Ashutosh Bagde. "A Comprehensive Evaluation of Data Analysis Approaches for Predicting Colorectal Cancer." In AIP Conference Proceedings, vol. 3188, no. 1, AIP Publishing LLC, 2024, 100094.

[5] ALGhafri, Houda Saif, and Chia S. Lim. "Fine-Tuning Models for Histopathological Classification of Colorectal Cancer." Diagnostics 15, no. 15 (2025): 1947.

[6] Ke, Qi, Yan Chai Hum, Wun-She Yap, Tian Swee Tan, Humaira Nisar, Hamam Mokayed, AiQuan Li, Rong Gao, and YuJian Gan. "Histopathological Classification of Colorectal Cancer Based on Domain-Specific Transfer Learning and Multi-Model Feature Fusion." Scientific Reports 15, no. 1 (2025): 35155.

[7] Prakash, U. M., S. Iniyan, Ashit Kumar Dutta, Shtwai Alsubai, Janjhyam Venkata Naga Ramesh, Sachi Nandan Mohanty, and Khasim Vali Dudekula. "Multi-Scale Feature Fusion of Deep Convolutional Neural Networks on Cancerous Tumor Detection and Classification Using Biomedical Images." Scientific Reports 15, no. 1 (2025): 1105.

[8] Sharkas, Maha, and Omneya Attallah. "Color-CADx: A Deep Learning Approach for Colorectal Cancer Classification Through Triple Convolutional Neural Networks and Discrete Cosine Transform." Scientific Reports 14, no. 1 (2024): 6914.

[9] AlShehri, Helala. "Dual-Path Convolutional Neural Network with Squeeze-and-Excitation Attention for Lung and Colon Histopathology Classification." Journal of Imaging 11, no. 12 (2025): 448.

[10] Venkatraman, Shravan, Jaskaran Singh Walia, and P. R. Joe Dhanith. "SAG-ViT: A Scale-Aware, High-Fidelity Patching Approach with Graph Attention for Vision Transformers." Complex & Intelligent Systems 11, no. 10 (2025): 428.

[11] Li, Meng. "Transformer-Based Self-Supervised Learning and Distillation for Medical Image Classification: Improving Colorectal Cancer Detection on nct-crc-he-100k With

Swin-T V2." In 2024 3rd International Conference on Cloud Computing, Big Data Application and Software Engineering (CBASE), IEEE, 2024, 644-648.

[12] Touvron, Hugo, Matthieu Cord, Matthijs Douze, Francisco Massa, Alexandre Sablayrolles, and Hervé Jégou. "Training Data-Efficient Image Transformers & Distillation Through Attention." In International conference on machine learning, PMLR, 2021, 10347-10357.

[13] Debnath, Jesika, Amira Hossain, Anamul Sakib, Hamdadur Rahman, Rezaul Haque, Md Redwan Ahmed, Ahmed Wasif Reza, SM Masfequier Rahman Swapno, and Abhishek Appaji. "LMVT: A Hybrid Vision Transformer with Attention Mechanisms for Efficient and Explainable Lung Cancer Diagnosis." Informatics in Medicine Unlocked (2025): 101669.

[14] Liu, Qifeng, Tao Zhou, Chi Cheng, Jin Ma, and Marzia Hoque Tania. "Hybrid Generative Adversarial Network Based on Frequency and Spatial Domain for Histopathological Image Synthesis." BMC bioinformatics 26, no. 1 (2025): 29.

[15] Karn, Prakash Kumar, Iman Ardekani, and Waleed H. Abdulla. "Generalized Framework for Liquid Neural Network Upon Sequential and Non-Sequential Tasks." Mathematics 12, no. 16 (2024): 2525.

[16] Gu, Albert, and Tri Dao. "Mamba: Linear-Time Sequence Modeling with Selective State Spaces." arXiv preprint arXiv:2312.00752 (2023).

[17] Liao, Jianshang, and Liguo Wang. "HyperspectralMamba: A Novel State Space Model Architecture for Hyperspectral Image Classification." Remote Sensing 17, no. 15 (2025): 2577.

[18] Sun, Yongheng, Yueh Z. Lee, Genevieve A. Woodard, Hongtu Zhu, Chunfeng Lian, and Mingxia Liu. "R2gen-mamba: A Selective State Space Model for Radiology Report Generation." In 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI), IEEE, 2025, 1-4.

[19] Brussee, Siemen, Giorgio Buzzanca, Anne MR Schrader, and Jesper Kers. "Graph Neural Networks in Histopathology: Emerging Trends and Future Directions." Medical Image Analysis 101 (2025): 103444.

[20] Liu, Ziming, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljačić, Thomas Y. Hou, and Max Tegmark. "Kan: Kolmogorov-Arnold Networks." arXiv preprint arXiv:2404.19756 (2024).

[21] Hasani, Ramin, Mathias Lechner, Alexander Amini, Daniela Rus, and Radu Grosu. "Liquid Time-Constant Networks." In Proceedings of the AAAI conference on artificial intelligence, vol. 35, no. 9, 2021, 7657-7666.

[22] Islam, Naeem Ul, Chung Tien Dat, and Muhammad Khalid. "Kolmogorov-Arnold Inspired Convolutional Networks for Enhancing PPO-based Online Reinforcement Learning." Journal of Experimental & Theoretical Artificial Intelligence 38, no. 2 (2026): 155-171.

[23] Opee, Shaiful Ajam, Arifa Akter Eva, Ahmed Taj Noor, Sayem Mustak Hasan, and M. F. Mridha. "ELW-CNN: An Extremely Lightweight Convolutional Neural Network for

Enhancing Interoperability in Colon and Lung Cancer Identification Using Explainable AI." Healthcare Technology Letters 12, no. 1 (2025): e12122.

[24] Riechie, Riechie, Vira Jessica, Matthew Kurniawan, and Feliks Victor Parningotan Samosir. "Convolutional Kolmogorov-Arnold Network for Pneumonia Detection in Medical Image Analysis." Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics 7, no. 3 (2025): 475-487.

[25] Ignatov, Andrey, and Grigory Malivenko. "Nct-crc-he: Not All Histopathological Datasets are Equally Useful." In European Conference on Computer Vision, Cham: Springer Nature Switzerland, 2024, 300-317.

[26] Maheshwari, Ullagadi, B. V. Kiranmayee, and Chalumuru Suresh. "Diagnose Colon and Lung Cancer Histopathological Images Using Pre-Trained Machine Learning Model." In 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), IEEE, 2022, 1078-1082.

[27] Sharma, Poonam, Bhisham Sharma, Ajit Noonia, Dhirendra Prasad Yadav, and Panos Liatsis. "Diagnosis of Colorectal Cancer Using Residual Transformer with Mixed Attention and Explainable AI." PLoS One 20, no. 11 (2025): e0335418.

[28] Amponsah, Anthony Afriyie. "Explainable AI for Histopathological Diagnosis: Comparison of Attention-Based Deep Learning Models with Expert Pathologists in Rare Tumor Classification." Journal of Integrated Research (2025): 102-131.

[29] Mothkur, Rashmi, and Pullagura Soubhagyalakshmi. "Grad-CAM Based Visualization for Interpretable Lung Cancer Categorization Using Deep CNN Models." Journal of Electronics, Electromedical Engineering, and Medical Informatics 7, no. 3 (2025): 567-580.

[30] Yan, Hongjie, Vivens Mubonanyikuzo, Temitope Emmanuel Komolafe, Liang Zhou, Tao Wu, and Nizhuan Wang. "Hybrid-RViT: Hybridizing ResNet-50 and Vision Transformer for Enhanced Alzheimer's Disease Detection." PloS one 20, no. 2 (2025): e0318998.

[31] Dosovitskiy, Alexey, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani et al. "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." arXiv preprint arXiv:2010.11929 (2020).

[32] Ochoa-Ornelas, Raquel, Alberto Gudiño-Ochoa, Sergio Octavio Rosales-Aguayo, Jesús Ezequiel Molinar-Solís, Sonia Espinoza-Morales, and René Gudiño-Venegas. "A Lightweight Cross-Gated Dual-Branch Attention Network for Colon and Lung Cancer Diagnosis from Histopathological Images." Medical Sciences 13, no. 4 (2025): 286.

[33] Said, Menatalla MR, Md Sakib Bin Islam, Md Shaheenur Islam Sumon, Semir Vranic, Rafif Mahmood Al Saady, Abdulrahman Alqahtani, Muhammad EH Chowdhury, and Shona Pedersen. "Innovative Deep Learning Architecture for the Classification of Lung and Colon Cancer from Histopathology Images." Applied Computational Intelligence and Soft Computing 2024, no. 1 (2024): 5562890.

[34] Kather, Jakob Nikolas, Niels Halama, and Alexander Marx. "100,000 Histological Images of Human Colorectal Cancer and Healthy Tissue." " https://www.kaggle.com/datasets/imrankhan77/nct-crc-he-100k (2018).

[35] Borkowski, Andrew A., Marilyn M. Bui, L. Brannon Thomas, Catherine P. Wilson, Lauren A. DeLand, and Stephen M. Mastorides. "Lung and Colon Cancer Histopathological Image Dataset (lc25000)." arXiv preprint arXiv:1912.12142 (2019).