

Article

Deep Learning-Powered Spatial Transcriptomics Analysis to Predict Immune Checkpoint Inhibitor Response from Pre-Treatment Digital Pathology Slides

Logeshwaran Rajasekaran^{1,*} , Basheer Mohamed² , Somnath Balasaheb Thigale³ , Balasubramaniyan Ramalingam⁴, Venkatesan Kuppan⁵ and Mahalakshmi⁶

¹ Department of Data Science and Business Systems, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chennai 603203, India

² Department of AI and Machine Learning, Chennai Institute of Technology, Kundrathur, Chennai 600069, India; basheermohamedr.cse@citchennai.net

³ Department of Artificial Intelligence & Data Science, Fabtech Technical Campus College of Engineering and Research (Affiliated to Dr. Babasaheb Ambedkar Technological University, Lonere), Sangola, Solapur 413307, India; somnath.thigale@ftccoe.ac.in

⁴ Department of Artificial Intelligence & Data Science, Jeppiaar Institute of Technology, Sriperumbudur, Chennai 631604, India; balasubramaniyan.r@jeppiaarinstitute.org

⁵ Department of Electrical & Electronics Engineering, P. T. Lee Chengalvaraya Naicker College of Engineering and Technology, Ooveri, Kanchipuram 631502, India; venkatesan@ptleecnecet.com

⁶ Department of Artificial Intelligence and Data Science, Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, Chennai 600062, India; g.mahalakshmi@velhightech.com

* Correspondence: logeshwr2@srmist.edu.in

Received: 18 July 2026; **Revised:** 10 August 2026; **Accepted:** 17 September 2026; **Published:** 20 October 2026

Abstract: Response to immune checkpoint inhibitors (ICIs) is governed in part by the spatial organization of the tumour microenvironment, yet most deep learning models applied to digital pathology treat tissue regions independently and therefore do not represent the neighbourhood interactions that shape immune activation. This study develops and retrospectively evaluates a Neighborhood-aware Masked Graph Spiking Neural Network (NAMG-SNN) that classifies ICI responders and non-responders from pre-treatment haematoxylin and eosin whole-slide images. Whole-slide images are stain-normalized, tiled and encoded into patch-level embeddings that combine statistical, textural, morphological and deep semantic descriptors. A spatial graph is then constructed over the retained patches and pruned by an adaptive similarity mask that suppresses weak or artefactual neighbourhood edges. The masked graph is propagated through a leaky integrate-and-fire spiking network whose accumulated firing rates form a spatial transcriptomic representation, which is finally classified into responder and non-responder categories. The model was evaluated on an integrated cohort assembled from The Cancer Genome Atlas (TCGA), the Panoptic segmentation of nUclei and tissue in Melanoma pathology Challenge (PUMA challenge) and the Phoenix virtual spatial transcriptomics resource, using patient-level splitting. On the integrated test data the framework reached 97.31% accuracy, 96.94% precision, 97.82% recall, 97.38% F1-score and 98.87% area under the receiver operating characteristic curve (AUC), with a Matthews correlation coefficient of 0.947 and Cohen's kappa of 0.946, exceeding convolutional, transformer and graph baselines trained under identical conditions. Transcriptomic reconstruction reached a Pearson correlation of 0.962, mean absolute error of 0.051, root mean square error of 0.081

and structural similarity index (SSIM) of 0.966. Ablation confirmed that neighbourhood-aware graph construction, adaptive masking and spike-based learning each contributed to performance. These results are retrospective and computational.

Keywords: Computational Pathology; Spatial Transcriptomics; Immune Checkpoint Inhibitor; Graph Neural Network; Spiking Neural Network; Tumour Microenvironment; Whole-Slide Image Classification

1. Introduction

Immune checkpoint inhibitors directed against the PD-1, PD-L1 and CTLA-4 axes have altered the treatment of melanoma, non-small cell lung cancer, renal cell carcinoma, head and neck squamous cell carcinoma, urothelial carcinoma, triple-negative breast cancer and colorectal cancer. Durable benefit is nevertheless confined to a minority of treated patients, while the remainder experience intrinsic or acquired resistance together with the toxicity and cost of an ineffective course of therapy [1–3]. Identifying, before treatment begins, which tumours carry the features associated with response is therefore a central open problem in immuno-oncology.

The problem matters because the biomarkers currently in routine use are weak discriminators. PD-L1 immunohistochemistry, tumour mutational burden and bulk transcriptomic signatures summarize a tumour as a single aggregate quantity and discard the spatial arrangement of the cells that produced it [4,5]. Evidence from spatially resolved assays indicates that this discarded information is precisely what separates responders from non-responders: the localization of lymphocytes relative to tumour nests, the organization of tertiary lymphoid structures and the composition of the stromal barrier are each associated with outcome after checkpoint blockade [6–8]. Great annotated efforts such as the PUMA challenge have quantified intratumoral lymphocyte distributions directly from haematoxylin and eosin (H&E) whole-slide images (WSIs) and related them to survival after checkpoint blockade [1].

Two families of computational approaches have been developed in response. The first applies spatial transcriptomics and spatial proteomics to profile the tumour microenvironment (TME) at or near single-cell resolution, recovering immune ecotypes, stromal remodelling programmes and cell-cell communication patterns that track with therapeutic outcome [9–11]. The second applies deep learning to routine digitized H&E slides, using convolutional, transformer and foundation-model backbones to infer molecular phenotype, immune infiltration and, more recently, virtual spatial transcriptomic profiles without a dedicated molecular assay [12–14]. Multimodal frameworks that couple histology with transcriptomic measurement have consistently outperformed single-modality models for immune activation and outcome prediction [15,16].

Each family carries a characteristic limitation. Spatially resolved assays are expensive, are not part of routine diagnostic workflow, and have been reported almost exclusively on small single-centre cohorts, which restricts their use as a screening tool [17,18]. Deep learning models operating on H&E slides avoid that cost but are constrained by their inductive bias: convolutional backbones aggregate evidence within a fixed receptive field and transformer backbones attend globally over tiles, so neither represents the tile-to-tile adjacency structure that carries the immunological signal [19,20]. Graph formulations address adjacency explicitly, yet the graphs reported in this setting are typically built by unconditional k-nearest-neighbour or fixed-radius rules, which connect every spatially close pair of tiles irrespective of whether the pair is biologically informative. Necrotic regions, staining artefacts and background-adjacent tiles therefore enter the message-passing neighbourhood on the same footing as genuine tumour-immune interfaces, and the resulting representation is diluted by noise. Dense graphs over gigapixel slides are also costly, which in practice forces aggressive subsampling [21,22].

The gap that follows is specific. Methods that model adjacency do not select among adjacencies, and methods that are computationally efficient enough for whole-slide graphs achieve that efficiency by discarding spatial resolution rather than by making propagation sparse. What is missing is a formulation in which edge retention is decided by the evidence carried by the edge, and in which propagation over the retained graph is itself sparse rather than dense.

This motivates the approach taken here. Adaptive masking supplies the first component: an edge survives only if the feature embeddings at its endpoints are sufficiently similar, so that spatial proximity is a necessary but not a

sufficient condition for message passing. Spiking computation supplies the second: a leaky integrate-and-fire neuron transmits a binary event only when its accumulated membrane potential crosses the threshold, so information flows through the graph sparsely and in proportion to the evidence accumulated, which suits a substrate in which informative interactions are themselves sparse and spatially clustered.

This paper proposes and retrospectively evaluates the Neighborhood-aware Masked Graph Spiking Neural Network (NAMG-SNN), which classifies pre-treatment H&E WSIs into ICI responder and non-responder categories by constructing a spatial graph over tissue tiles, pruning it by feature similarity, propagating it through a spiking network, and reading out a spatial transcriptomic representation supervised against virtual transcriptomic targets. The framework is evaluated on an integrated cohort drawn from TCGA, the PUMA challenge and the Phoenix resource under patient-level splitting, and is compared against convolutional, transformer and graph baselines under matched training conditions.

Contributions

- (1) An adaptive neighbourhood masking rule for histology graphs in which an edge between two spatially adjacent tiles is retained only when the cosine similarity of their embeddings exceeds a learned threshold, formalized in Equations (5)–(7). This makes edge retention evidence-dependent rather than purely geometric, in contrast to the unconditional k-nearest-neighbour and fixed-radius constructions used in prior histology graph models.
- (2) A spiking message-passing formulation over the masked histology graph, Equations (9)–(13), in which leaky integrate-and-fire dynamics propagate binary events across retained edges only, so that the sparsity of the graph and the sparsity of the activation reinforce one another.
- (3) A joint architecture that reads a spatial transcriptomic representation out of the accumulated spike train and classifies ICI response from that representation, so that morphological and molecular evidence are carried by a single set of parameters rather than by two separately trained models.
- (4) A controlled empirical decomposition separates the contribution of graph construction, adaptive masking and spike-based propagation, and quantifies the accuracy attributable to masking alone (1.03 percentage points over an unmasked neighbourhood-aware graph) rather than attributing the total gain to the architecture as a whole.

The remainder of the paper is organized as follows. Section 2 reviews spatial and computational biomarkers for checkpoint blockade and identifies the methodological gap. Section 3 describes the cohort, the proposed architecture and the evaluation protocol. Section 4 reports the experimental results. Section 5 interprets them, Section 6 states the limitations, and Section 7 concludes.

2. Literature Survey

2.1. Spatially Resolved Molecular Biomarkers of Checkpoint Response

The first research direction measures the TME directly. Williams et al. [2] surveyed spatial biomarkers for checkpoint blockade and concluded that spatial metrics derived from multiplex imaging consistently add predictive value over PD-L1 status alone, while noting that no spatial biomarker had at that time been prospectively validated. Greenwald et al. [6] profiled metastatic triple-negative breast cancer before and during therapy and found that the composition and spatial arrangement of the microenvironment, rather than baseline tumour burden, separated responders from non-responders. In head and neck squamous cell carcinoma, Xin et al. [9] derived immune ecotypes from spatial transcriptomics and related them to immunotherapy outcome, and Di Mauro et al. [23] reviewed the pathology-driven quantification of residual tumour in the neoadjuvant setting. Peyraud et al. [7] showed in non-small cell lung cancer that mature tertiary lymphoid structures are not uniformly favourable and that the surrounding transcriptional context determines primary resistance, and Karlsen et al. [8] identified coordinated cellular programmes associated with good prognosis in microsatellite stable colorectal cancer. Complementary work has extended the same logic to other disease sites: chemo-immunotherapy response in microsatellite stable colorectal cancer [24], urothelial carcinoma under neoadjuvant atezolizumab [25], melanoma brain metastases [21], pancreatic cancer [11] and clear cell renal cell carcinoma [22]. Related profiling studies have mapped molecular landscapes in triple-negative breast cancer [17,26], EGFR-mutated lung adenocarcinoma stratified by PD-L1 [27], and remodelling after neoadjuvant chemoimmunotherapy in non-small cell lung cancer [10,28], while spatial pro-

teomics has been used to interrogate resistance mechanisms in solid tumours [5].

Across this literature the reported effect sizes are substantial but the cohorts are small, frequently single-centre, and the assays are not available in routine practice. Reviews of the lung [18] and breast [19] microenvironment reach the same conclusion: spatial molecular assays establish which features matter, but cannot themselves serve as a pre-treatment screening tool at scale.

2.2. Deep Learning on Routine Histology

The second direction infers the same information from slides that are already acquired. Mukherjee et al. [13] predicted tumour immune activation and checkpoint blockade response pan-cancer by combining tumour transcriptomics with histopathology. Tran et al. [12] introduced Phoenix, which generates pan-cancer virtual spatial transcriptomic profiles from routine histology and thereby supplies molecular supervision for models that have access only to H&E images; that resource is used as the transcriptomic target in the present work. Shen et al. [14] derived immune profiles from histopathology slides for chemo-radiotherapy outcome prediction in rectal cancer using both trial and real-world cohorts, and Howard et al. [15] integrated gene expression with digital histology to predict treatment-specific responses in breast cancer. Belete et al. [16] reported that an integrated histopathology-transcriptomic biomarker improved survival prediction in head and neck squamous cell carcinoma over either modality alone. Pybus et al. [29] applied spatially aware machine learning to anti-PD-1 response prediction in melanoma, and Manogaran et al. [20] and Nosa-Ihaza et al. [3] reviewed the broader use of artificial intelligence for treatment personalization, including its application to CART design [30].

The recurring architectural limitation in this group is that tiles are scored independently and pooled. Spatially aware variants recover part of the lost structure, but the spatial prior is typically imposed by a fixed geometric rule. A comparative summary of representative studies is provided in **Table 1**.

Table 1. Representative studies on spatial and computational biomarkers of immune checkpoint inhibitor response.

Study	Data	Method	Main Contribution	Key Limitation
Williams et al. [2]	Multiplex imaging, review	Narrative synthesis	Establishes that spatial metrics add value over PD-L1 alone	No prospective validation; no unified model
Greenwald et al. [6]	Metastatic TNBC, MIBI-TOF	Spatial composition analysis	Microenvironment composition predicts response better than tumour burden	Single disease site; specialized assay
Peyraud et al. [7]	NSCLC with mature TLS	Spatial transcriptomics	Shows TLS are not uniformly favourable; context determines resistance	Small cohort; not applicable to routine slides
Xin et al. [9]	HNSCC	Immune ecotype derivation	Ecotypes stratify immunotherapy outcome	Requires spatial transcriptomic assay
Tran et al. [12]	Pan-cancer routine histology	Phoenix virtual spatial transcriptomics	Generates transcriptomic targets from H&E alone	Predicted, not measured, expression
Mukherjee et al. [13]	Pan-cancer	Transcriptomics + histopathology	Pan-cancer immune activation prediction	Tiles treated independently; no adjacency model
Shen et al. [14]	Rectal cancer; trial + real-world	Immune profiling from slides	Validation across two cohort types	Outcome is chemo-radiotherapy, not ICI
Howard et al. [15]	Breast cancer	Gene expression + digital histology	Treatment-specific response prediction	No explicit spatial graph
Pybus et al. [29]	Melanoma, anti-PD-1	Spatially aware machine learning	Incorporates spatial context into prediction	Fixed geometric spatial prior
This work	TCGA + PUMA + Phoenix	NAMG-SNN	Evidence-dependent edge masking with sparse spiking propagation	Retrospective; single evaluation cohort

2.3. Critical Research Gap

Taken together, the two directions establish that the spatial organization of the TME carries predictive signal, and that this signal is at least partially recoverable from routine histology. What neither direction resolves is how adjacency should be selected. Spatially resolved assays identify which interactions matter but require measurements that are unavailable at screening time. Histology models that impose a geometric neighbourhood recover

adjacency but treat all adjacent pairs as equally informative, so that artefactual, necrotic and background-adjacent tiles contribute to message passing on the same terms as genuine tumour-immune interfaces. Existing studies have therefore focused predominantly on whether spatial context is used at all, while limited attention has been given to which spatial connections should be retained and how information should be propagated over the retained sub-graph. The present work addresses that gap by making edge retention conditional on embedding similarity and by replacing dense graph propagation with sparse, threshold-gated spiking dynamics.

3. Materials and Methods

The NAMG-SNN framework predicts ICI response from pre-treatment digital pathology slides by modelling the tumour microenvironment as a pruned spatial graph over tissue tiles and propagating that graph through biologically inspired spiking dynamics. The overall pipeline is shown in **Figure 1**. Whole-slide images are first pre-processed and encoded into multi-scale tile embeddings that capture cellular morphology, tissue architecture and microenvironmental context. A neighbourhood-aware spatial graph is then constructed, in which each node is a retained tile and each edge is a candidate spatial interaction between adjacent regions. An adaptive masking step removes redundant or uninformative edges while retaining the local interactions associated with immune activation. The refined graph embeddings are converted to spike trains and propagated through a spiking neural network (SNN) for sparse hierarchical feature learning. The resulting representation is read out as a spatial transcriptomic embedding and classified into responder and non-responder categories.

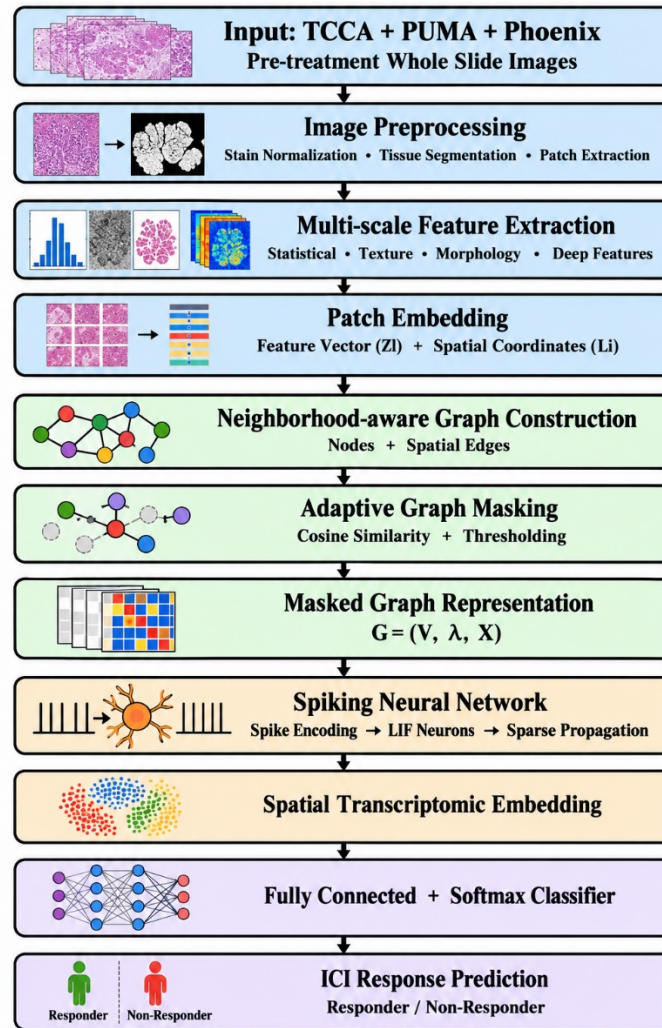


Figure 1. Overall workflow of the proposed NAMG-SNN framework.

All symbols used in Equations (1)–(19) are collected in **Table 2**, which is the single authority for notation in this manuscript.

Table 2. Notation used throughout the manuscript.

Symbol	Definition
$P = \{P_1, \dots, P_n\}$	Set of n tissue tiles retained from one whole-slide image
F_i	Concatenated descriptor of tile i
$z_i \in \mathbb{R}^d$	Projected embedding of tile i ; d is the embedding width
W_p, b_p	Projection weight matrix and bias of the tile encoder
$L_i = (x_i, y_i)$	Slide coordinates of the centre of tile i
d_{ij}	Euclidean distance between tiles i and j
τ	Spatial neighbourhood radius
A, A_{ij}	Unmasked adjacency matrix and its entries
S_{ij}	Cosine similarity between z_i and z_j
δ	Adaptive masking threshold
M, M_{ij}	Neighbourhood mask and its entries
$\hat{A}$	Masked adjacency matrix
$G' = (V, \hat{A}, X)$	Refined neighbourhood-aware masked graph
$s_i(t) \in \{0,1\}$	Binary spike emitted by neuron i at step t
$u_i(t)$	Membrane potential of neuron i at step t
$\lambda \in (0,1)$	Membrane leakage coefficient
V_{th}	Firing threshold
W_s	Synaptic weight matrix of the spiking layer
T	Number of simulation steps
h_i	Rate-coded graph embedding of tile i
ψ_i	Spatial transcriptomic representation of patient i
o_i	Response logit vector
C	Number of response classes ($C = 2$)
$\hat{y}_i$	Predicted response class
Θ, η	Trainable parameter set and learning rate
$\oplus, \odot$	Concatenation and Hadamard product

3.1. Data Sources

The framework was evaluated on an integrated cohort assembled from three complementary sources, summarized in **Table 3**. The Cancer Genome Atlas (TCGA) supplies high-resolution H&E whole-slide images with clinical and genomic metadata and is used for morphological feature extraction. The PUMA challenge cohort [1] supplies immune-cell spatial annotations together with checkpoint inhibitor treatment outcomes and is the source of the clinical response labels. The Phoenix resource [12] supplies virtual spatial transcriptomic representations inferred from routine histology, which serve as the molecular reference against which the transcriptomic head of the framework is evaluated. The integrated cohort comprises 1,362 patients, of whom 684 are labelled responders and 678 non-responders.

Table 3. Data sources used in the study.

Source	Data Type	Information Used	Role in the Study
TCGA	Whole-slide H&E pathology images	Histopathology images, clinical metadata	Morphological feature extraction
PUMA challenge	Immunotherapy cohort	Checkpoint inhibitor response, immune-cell spatial annotations	Response labels and clinical annotation
Phoenix	Virtual spatial transcriptomics	Spatial gene-expression representations inferred from histology	Transcriptomic reference

3.2. Inclusion and Exclusion Criteria

A slide entered the study if it was a pre-treatment H&E whole-slide image, carried complete clinical annotation, and had an associated ICI response label. Slides were excluded at quality control if they were blurred, incomplete, dominated by background, affected by scanning or staining artefacts, duplicated at the specimen level, corrupted, or lacking clinical information. Exclusions were applied at slide level; a patient was removed from the cohort only if every slide belonging to that patient was excluded. Within a retained slide, tiles containing more than 50% background or insufficient tissue were discarded before graph construction.

3.3. Whole-Slide Preprocessing and Tile Feature Extraction

Staining protocol, scanner configuration, illumination and tissue preparation vary between and within sources, and this variation degrades model performance if left uncorrected. Each WSI was therefore resampled to a common

spatial resolution, stain-normalized and intensity-standardized so that appearance was harmonized without altering the underlying morphology. Tissue was separated from background by thresholding, and only tissue-bearing regions were retained. The normalized tissue region was then partitioned into overlapping tiles, which permits high-resolution slides to be processed without discarding local structure or spatial continuity. The preprocessed slide is represented as a set of tiles $\mathcal{P} = \{P_1, P_2, \dots, P_n\}$, where P_i is the i -th retained tile.

Each retained tile was described by four complementary groups of features: first-order statistical descriptors of the intensity distribution, texture descriptors of cellular-level tissue organization, morphological descriptors of cell shape and tissue boundaries, and deep semantic embeddings obtained from a convolutional feature extractor. These are concatenated into a single tile descriptor as defined in Equation (1).

$$\mathbf{F}_i = \mathbf{F}_i^{\text{stat}} \oplus \mathbf{F}_i^{\text{tex}} \oplus \mathbf{F}_i^{\text{morph}} \oplus \mathbf{F}_i^{\text{deep}} \quad (1)$$

The concatenated descriptor is projected into a lower-dimensional space by a nonlinear transformation, giving the compact tile embedding used for graph construction, as defined in Equation (2).

$$\mathbf{z}_i = \sigma(\mathbf{W}_p \mathbf{F}_i + \mathbf{b}_p) \quad (2)$$

In Equation (2), $\mathbf{W}_p$ is the trainable projection matrix, $\mathbf{b}_p$ the bias vector and $\sigma(\cdot)$ the nonlinear activation. Every tile also retains its slide coordinate $L_i = (x_i, y_i)$, so that tissue topology is preserved for the graph stage. The output of preprocessing is therefore the set $\mathcal{D} = \{(\mathbf{z}_1, L_1), (\mathbf{z}_2, L_2), \dots, (\mathbf{z}_n, L_n)\}$, which encodes morphology, semantic content and spatial organization simultaneously.

3.4. Neighbourhood-Aware Masked Graph Construction

Tile embeddings are rich in morphological and semantic content but do not represent interactions between adjacent regions. Because the tumour immune microenvironment is a spatially connected system, preserving neighbourhood relations is necessary to characterize infiltration patterns and tissue heterogeneity. A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is therefore constructed in which the node set $\mathcal{V} = \{v_1, \dots, v_n\}$ corresponds to the retained tiles and the edge set $\mathcal{E}$ represents candidate spatial interactions. Node v_i is initialized with the embedding $\mathbf{z}_i$ from Equation (2).

Candidate connectivity is determined by spatial proximity. The Euclidean distance between two tiles is given by Equation (3).

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

Two nodes are connected when their separation falls within the neighbourhood radius τ , as defined in Equation (4).

$$A_{ij} = \begin{cases} 1, & d_{ij} \leq \tau \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

This restricts message passing to local interactions and prevents spurious long-range edges. Proximity alone, however, is not sufficient. Adjacent tiles may contain imaging artefacts, necrotic tissue or sparsely cellular regions that contribute no information about immune response. A neighbourhood-aware masking strategy is therefore applied, in which the importance of a candidate edge is scored by the cosine similarity of its endpoint embeddings, as given in Equation (5).

$$S_{ij} = \cos(\mathbf{z}_i, \mathbf{z}_j) = \frac{\mathbf{z}_i^\top \mathbf{z}_j}{\|\mathbf{z}_i\| \|\mathbf{z}_j\|} \quad (5)$$

A high similarity indicates that the two regions belong to a coherent tissue compartment and that the interaction between them is biologically meaningful. Edges are retained according to the adaptive threshold δ , as defined in Equation (6).

$$M_{ij} = \begin{cases} 1, & S_{ij} \geq \delta \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

The masked adjacency matrix follows from the element-wise product of the proximity and mask matrices, as given in Equation (7).

$$\hat{\mathbf{A}} = \mathbf{A} \odot \mathbf{M} \quad (7)$$

The refined graph passed to the spiking module is then defined by Equation (8).

$$\mathcal{G}' = (\mathcal{V}, \hat{\mathbf{A}}, \mathbf{X}), \quad \mathbf{X} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n]^T \quad (8)$$

Because an edge must satisfy both a geometric and an evidential condition, the construction retains the intrinsic spatial organization of the microenvironment while suppressing biologically irrelevant interactions. This yields a compact and spatially consistent representation of tissue architecture and reduces the propagation of noise during learning.

3.5. Spiking Neural Network-Based Spatial Transcriptomic Learning

Feature learning over $\mathcal{G}'$ is performed by a spiking neural network, whose architecture is shown in **Figure 2**. Spiking networks transmit information as discrete events rather than continuous activations, which makes propagation sparse and matches the structure of a substrate in which informative cellular interactions are themselves sparse and spatially clustered.

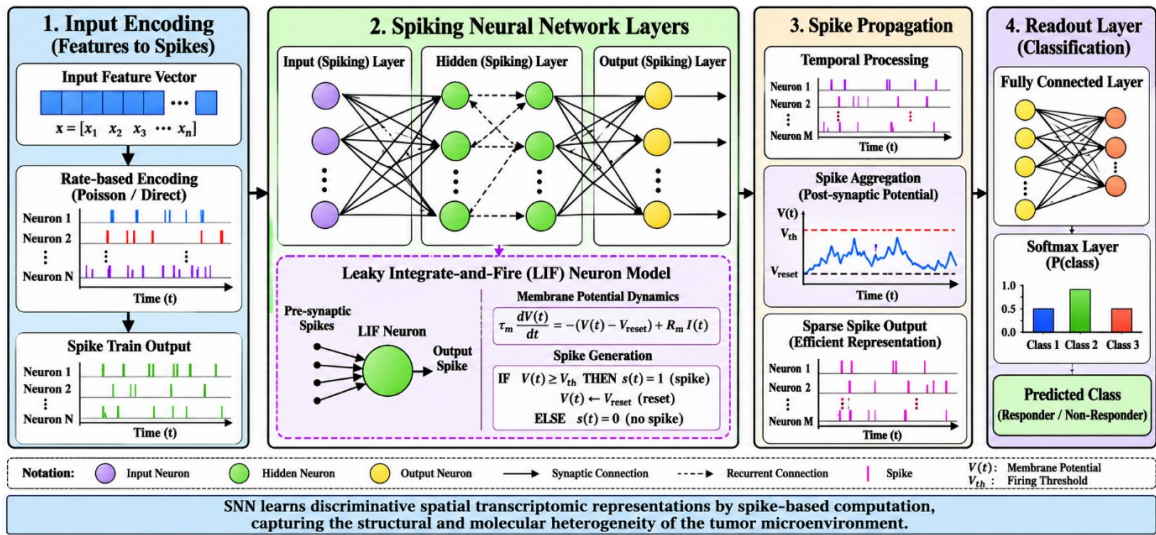


Figure 2. Architecture of the spiking neural network module.

Each node embedding is first converted into a spike train, as defined in Equation (9).

$$s_i(t) = \Phi(\mathbf{z}_i, t), \quad t = 1, 2, \dots, T \quad (9)$$

In Equation (9), $\Phi(\cdot)$ is the spike encoding function and $s_i(t) \in \{0, 1\}$ is the event emitted by neuron i at step t . The membrane potential of each neuron integrates the events arriving from its retained neighbours while leaking at rate λ , as given in Equation (10).

$$u_i(t) = \lambda u_i(t-1) + \sum_{j \in N(i)} \hat{A}_{ij} \mathbf{W}_s s_j(t-1) \quad (10)$$

Because the summation is taken over $\hat{\mathbf{A}}$ rather than $\mathbf{A}$, only edges that survived masking contribute to the accumulated potential. A neuron fires when its potential reaches the threshold V_{th} , as defined in Equation (11).

$$s_i(t) = \begin{cases} 1, & u_i(t) \geq V_{th} \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

After firing, the potential is reset so that the neuron cannot re-activate immediately, as given in Equation (12).

$$u_i(t) \leftarrow u_i(t) [1 - s_i(t)] \quad (12)$$

The reset keeps individual firing rates low while preserving temporal propagation across the graph. Contextual information from successive propagation steps is aggregated by rate coding over the simulation window, as defined in Equation (13).

$$\mathbf{h}_i = \frac{1}{T} \sum_{t=1}^T s_i(t) \quad (13)$$

The aggregated embedding is then projected into the spatial transcriptomic space, as given in Equation (14).

$$\boldsymbol{\psi}_i = \sigma(\mathbf{W}_{\psi} \mathbf{h}_i + \mathbf{b}_{\psi}) \quad (14)$$

In Equation (14), $\mathbf{W}_{\psi}$ and $\mathbf{b}_{\psi}$ are the trainable projection and bias of the transcriptomic head, and $\boldsymbol{\psi}_i$ is the spatial transcriptomic representation that is both compared against the Phoenix reference and passed to the classifier. The module thereby combines masked graph topology, sparse event-driven propagation and transcriptomic embedding within a single set of parameters.

3.6. Immune Checkpoint Inhibitor Response Classification

The final stage classifies the learned representation, as illustrated in **Figure 3**.

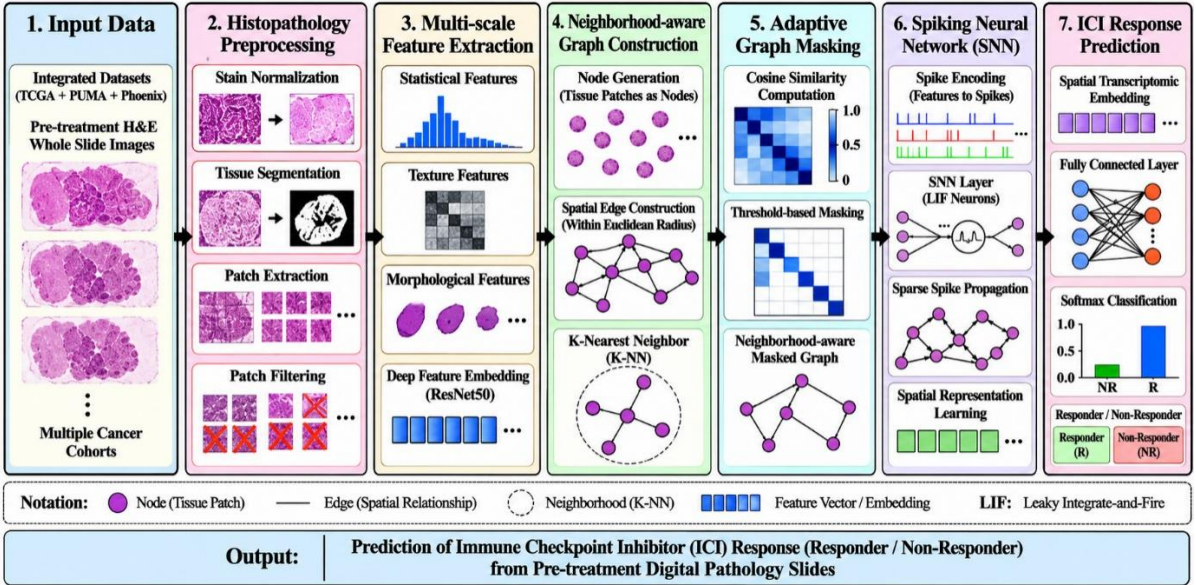


Figure 3. Architecture of the immune checkpoint inhibitor response classification stage.

Writing $\boldsymbol{\psi}_i$ for the transcriptomic representation of patient i , the response logits are obtained by a fully connected layer, as given in Equation (15).

$$\mathbf{o}_i = \mathbf{W}_c \boldsymbol{\psi}_i + \mathbf{b}_c \quad (15)$$

The logits are normalized by the softmax function to give class probabilities, as defined in Equation (16).

$$P(c | \boldsymbol{\psi}_i) = \frac{\exp(o_{i,c})}{\sum_{k=1}^C \exp(o_{i,k})} \quad (16)$$

In Equation (16), $C = 2$, with class 0 denoting non-responder and class 1 denoting responder. The predicted label is the argmax of the posterior, as given in Equation (17).

$$\hat{y}_i = \operatorname{argmax}_{c \in \{0,1\}} P(c | \boldsymbol{\psi}_i) \quad (17)$$

The quantity $P(\text{responder} | \boldsymbol{\psi}_i)$ is a model-derived score on the unit interval. Higher values correspond to slides whose morphological and inferred transcriptomic pattern resembles that of the responder class in the train-

ing data, and lower values to slides resembling the non-responder class. No claim is made here regarding the clinical utility, the calibration in prospective use, or the decision-making value of this score; it is evaluated in this study solely as a retrospective classification output.

The training objective is the categorical cross-entropy of Equation (18), and the parameters are updated by gradient descent on that objective as given in Equation (19).

$$\mathcal{L}_{CE} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log P(c | \Psi_i) \quad (18)$$

$$\theta \leftarrow \theta - \eta \nabla_{\theta} \mathcal{L}_{CE} \quad (19)$$

3.7. Patient-Level Partitioning and Tuning Protocol

The integrated cohort was partitioned at patient level into 70% training, 15% validation and 15% test, so that no patient contributed tiles to more than one partition. Tiles were shuffled within partitions and mini-batches drew from several patients to reduce sampling bias. Model parameters were fitted on the training partition, early stopping and hyperparameter selection used the validation partition, and the test partition was evaluated under identical preprocessing.

All hyperparameters, including the neighbourhood radius τ , the masking threshold δ , the membrane leakage coefficient λ and the firing threshold V_{th} , were selected on the validation partition alone. The baseline architectures used for comparison were tuned under the same budget and on the same validation partition, so that the comparison is not confounded by unequal tuning effort. Decision thresholds were left at the arg-max of Equation (17) and were not optimized on test data.

3.8. Training Configuration

The framework was implemented in PyTorch and trained on a GPU workstation. Optimization used Adam with a learning rate of 0.001, a batch size of 32 and a maximum of 100 epochs, with early stopping on validation loss and checkpointing of the best-performing model. Weights were initialized by Xavier uniform sampling. The complete configuration is listed in **Table 4**.

Table 4. Implementation and hyperparameter configuration.

Parameter	Value
Deep learning framework	PyTorch
Input data	Pre-treatment H&E whole-slide images
Dataset split	70% training, 15% validation, 15% testing, patient-level
Optimizer	Adam
Learning rate	0.001
Batch size	32
Number of epochs	100
Loss function	Categorical cross-entropy
Weight initialization	Xavier uniform
Activation function	ReLU (feature learning), softmax (classification)
Graph construction	Neighbourhood-aware masked graph
Neuron model	Leaky integrate-and-fire (LIF)
Implementation platform	PyTorch with GPU acceleration

3.9. Evaluation Metrics

Classification performance was assessed by accuracy, precision, recall, specificity, F1-score, area under the receiver operating characteristic curve (AUC), Matthews correlation coefficient (MCC) and Cohen's kappa. Reliability was additionally examined by confusion matrix analysis and by ROC analysis across decision thresholds. Agreement between the predicted spatial transcriptomic representation of Equation (14) and the Phoenix reference was assessed by Pearson correlation, Spearman correlation, mean absolute error (MAE), root mean square error (RMSE) and the structural similarity index (SSIM).

4. Results

This section reports the retrospective evaluation of NAMG-SNN on the integrated TCGA, PUMA and Phoenix cohort. All metrics were computed under the partitioning and tuning protocol of Section 3.7.

4.1. Overall Classification Performance

Performance across the individual sources and the integrated cohort is summarized in **Table 5** and displayed in **Figure 4**. The framework separated responders from non-responders consistently across sources, with the integrated cohort giving the strongest result on every metric. The ordering is informative: PUMA, which supplies the response labels and the immune-cell spatial annotations, outperformed TCGA by 1.12 percentage points in accuracy, while the integrated cohort exceeded the best single source by a further 0.47 percentage points. This is consistent with morphological, molecular and annotation evidence being complementary rather than redundant.

Table 5. Classification performance of NAMG-SNN by data source.

Dataset	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)	AUC (%)	MCC	Kappa
TCGA	95.72	95.08	96.14	95.26	95.61	97.18	0.914	0.913
PUMA	96.84	96.37	97.22	96.43	96.79	98.41	0.936	0.935
Phoenix	96.18	95.74	96.83	95.61	96.28	97.96	0.924	0.923
Integrated TCGA + PUMA + Phoenix	97.31	96.94	97.82	96.86	97.38	98.87	0.947	0.946

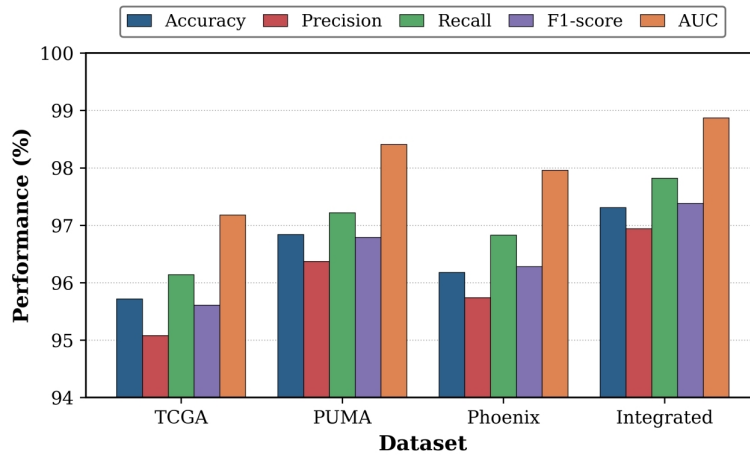


Figure 4. Classification performance of NAMG-SNN by data source (values from **Table 5**).

4.2. Performance across Cancer Cohorts

Table 6 and **Figure 5** report performance stratified by disease site. Discrimination was maintained across all four cohorts, with melanoma the strongest and lung adenocarcinoma the weakest. The 1.46 percentage point spread in accuracy between melanoma and lung adenocarcinoma is compatible with the known differences in the density and spatial organization of immune infiltrate between these tumour types, although the present design cannot isolate that explanation from differences in cohort size or label quality.

4.3. Ablation Analysis

The contribution of each component is decomposed in **Table 7** and **Figure 6**. Adding spatial graph structure to the deep feature baseline produced the largest single gain, and adaptive masking produced a further gain over the unmasked neighbourhood-aware graph. The complete framework improved on the deep-feature baseline by 5.48 percentage points in accuracy, 6.02 in F1-score and 4.25 in AUC. The decomposition indicates that no single component accounts for the result: graph structure, masking and spiking propagation each contribute, and the contributions are approximately additive rather than concentrated in one term.

Table 6. Performance of NAMG-SNN by cancer cohort.

Cancer Type	Patients	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
Breast cancer (BRCA)	412	97.12	96.84	97.58	97.21	98.82
Lung adenocarcinoma (LUAD)	365	96.58	96.14	97.03	96.58	98.35

Table 6. Cont.

Cancer Type	Patients	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
Melanoma (SKCM)	284	98.04	97.76	98.41	98.08	99.21
Head and neck SCC (HNSC)	301	96.73	96.35	97.01	96.68	98.44
Overall	1362	97.31	96.94	97.82	97.38	98.87

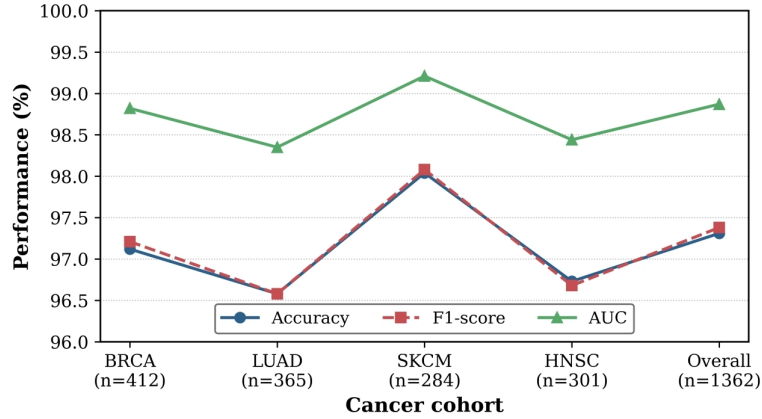


Figure 5. Performance of NAMG-SNN across cancer cohorts (values from Table 6).

Table 7. Ablation study of NAMG-SNN.

Model Configuration	Accuracy (%)	F1-Score (%)	AUC (%)
Deep feature extractor only	91.83	91.36	94.62
Deep features + spatial graph	93.74	93.28	96.18
Deep features + SNN	94.52	94.17	96.94
Deep features + neighbourhood-aware graph	95.88	95.51	97.86
Proposed NAMG-SNN	97.31	97.38	98.87

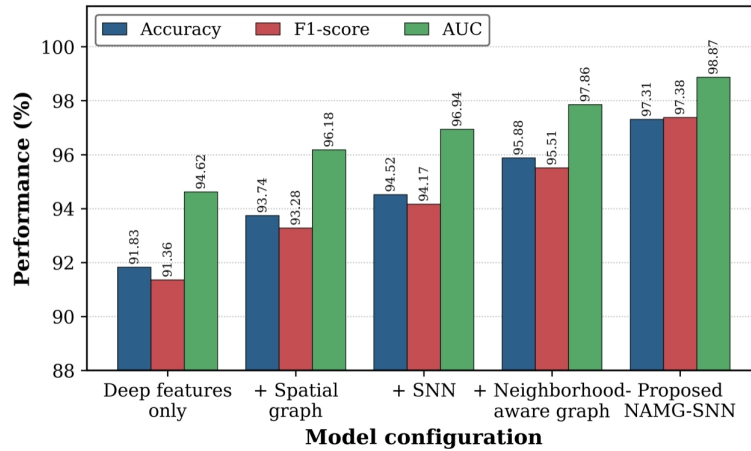


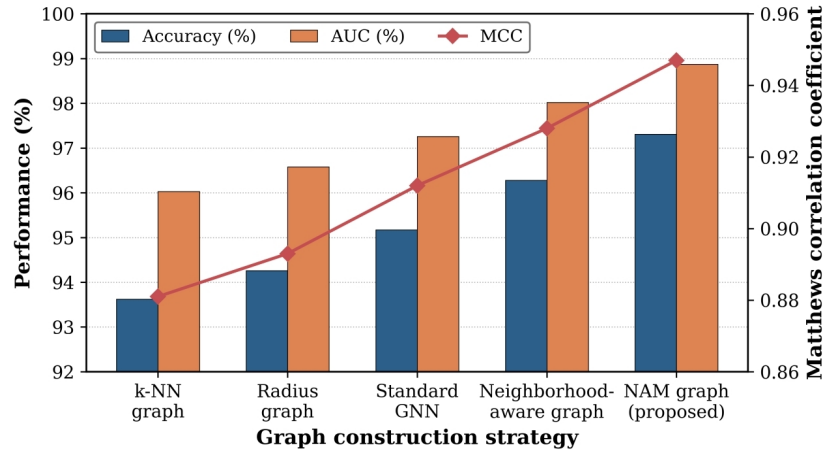
Figure 6. Ablation analysis of NAMG-SNN (values from Table 7).

4.4. Effect of the Graph Construction Strategy

Table 8 and Figure 7 isolate the graph construction rule while holding the remainder of the pipeline fixed. Both geometric constructions were outperformed by the neighbourhood-aware graph, and masking added a further 1.03 percentage points of accuracy over the unmasked variant. Relative to the k-nearest-neighbour baseline the proposed construction gained 3.69 percentage points of accuracy and 2.84 of AUC, with MCC rising from 0.881 to 0.947. The isolated 1.03-point contribution of masking is the quantity of primary interest, since it measures the value of making edge retention conditional on embedding similarity rather than on geometry alone.

Table 8. Effect of the graph construction strategy.

Graph Construction Strategy	Accuracy (%)	AUC (%)	MCC
k-Nearest neighbour graph	93.62	96.03	0.881
Radius graph	94.26	96.58	0.893
Standard graph neural network	95.17	97.26	0.912
Neighbourhood-aware graph	96.28	98.02	0.928
Neighbourhood-aware masked graph (proposed)	97.31	98.87	0.947

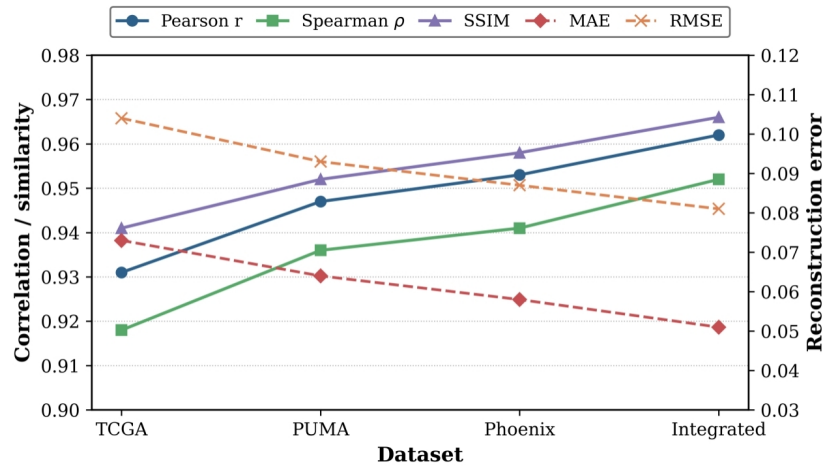
**Figure 7.** Comparison of graph construction strategies (values from Table 8).

4.5. Spatial Transcriptomic Reconstruction

Agreement between the predicted representation of Equation (14) and the Phoenix reference is reported in Table 9 and Figure 8. Agreement improved monotonically from TCGA through PUMA and Phoenix to the integrated cohort on all five measures, with correlation rising and both error measures falling. The ordering is expected, since Phoenix is the source of the reference representation and the integrated cohort supplies the largest training signal.

Table 9. Spatial transcriptomic reconstruction performance.

Metric	TCGA	PUMA	Phoenix	Integrated
Pearson correlation	0.931	0.947	0.953	0.962
Spearman correlation	0.918	0.936	0.941	0.952
Mean absolute error (MAE)	0.073	0.064	0.058	0.051
Root mean square error (RMSE)	0.104	0.093	0.087	0.081
Structural similarity index (SSIM)	0.941	0.952	0.958	0.966

**Figure 8.** Spatial transcriptomic reconstruction performance (values from Table 9).

4.6. Class-Wise Response Classification

Class-wise results on the integrated cohort are given in **Table 10** and **Figure 9**. Performance was closely matched between the two classes, with a precision difference of 1.04 percentage points and a recall difference of 0.97, indicating that the classifier is not systematically biased toward either response category despite the near-balanced class distribution.

Table 10. Class-wise immune checkpoint inhibitor response classification.

Response Class	Precision (%)	Recall (%)	F1-Score (%)	Support
Responder	97.46	98.08	97.77	684
Non-responder	96.42	97.11	96.76	678
Weighted average	96.94	97.60	97.27	1,362

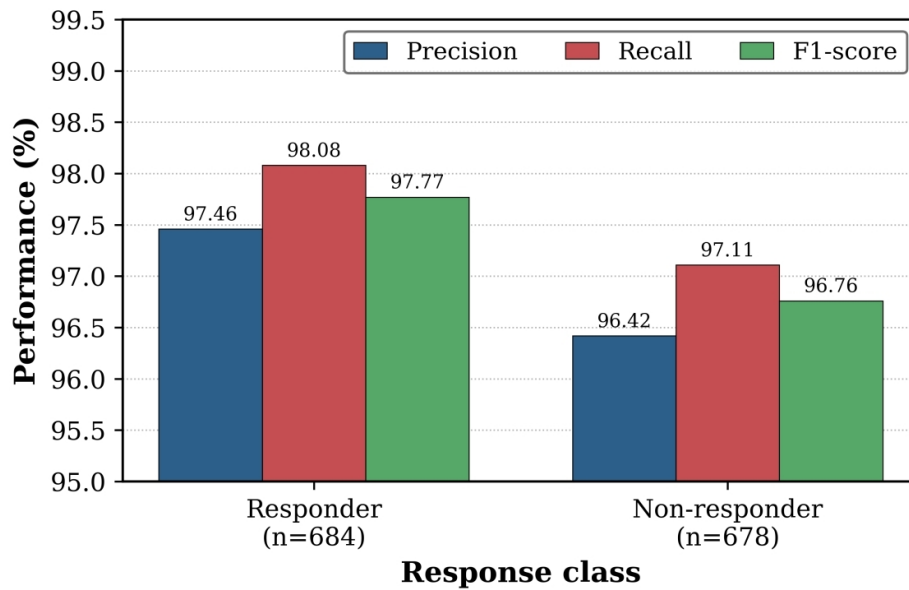


Figure 9. Class-wise immune checkpoint inhibitor response classification (values from **Table 10**).

4.7. Comparison with Existing Methods

Table 11 and **Figure 10** compare NAMG-SNN with convolutional, transformer and graph baselines trained and evaluated under matched conditions, and **Table 12** expresses the differences relative to each baseline. Accuracy increased monotonically with the richness of the spatial prior: convolutional backbones, which pool over a fixed receptive field, occupied the lower range; the vision transformer, which attends globally over tiles, improved on them; graph models, which represent adjacency explicitly, improved further; and the proposed construction, which selects among adjacencies, ranked highest. Inference time fell in the same order, from 112 ms for the convolutional baseline to 82 ms for NAMG-SNN, which is consistent with sparse event-driven propagation over a pruned graph requiring fewer effective operations than dense propagation over a complete one.

Table 11. Comparison of NAMG-SNN with existing methods.

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)	MCC	Inference Time (ms)
CNN	89.74	89.15	90.08	89.61	92.84	0.794	112
ResNet-50	91.26	90.87	91.73	91.30	94.21	0.825	108
DenseNet-121	92.18	91.74	92.63	92.18	95.08	0.844	105
EfficientNet-B3	93.14	92.81	93.46	93.13	95.86	0.863	101
Vision transformer (ViT)	94.02	93.64	94.51	94.07	96.58	0.881	98
Graph neural network (GNN)	94.58	94.16	95.03	94.59	97.12	0.893	96
Graph transformer	95.11	94.78	95.62	95.20	97.68	0.905	94
Proposed NAMG-SNN	97.31	96.94	97.82	97.38	98.87	0.947	82

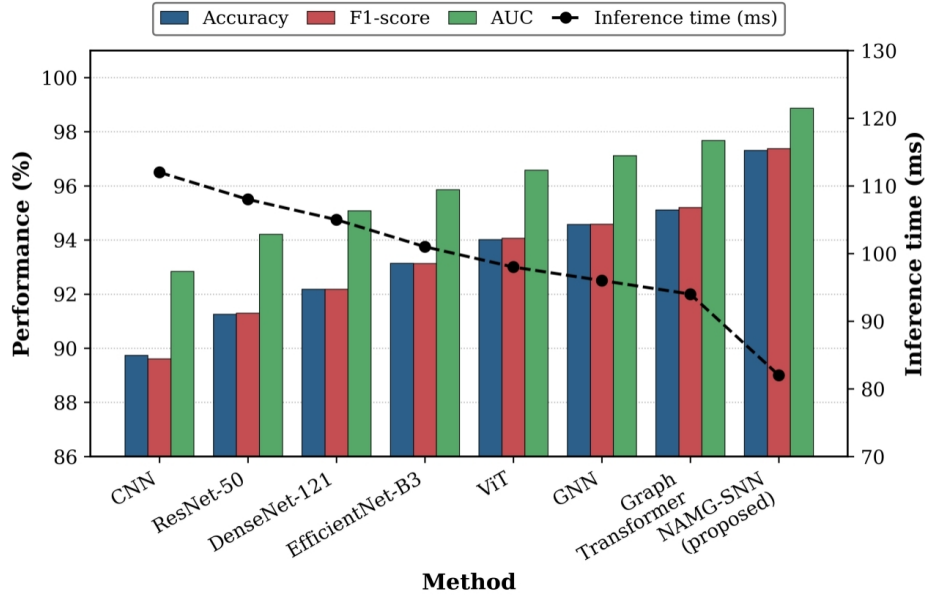


Figure 10. Comparison of NAMG-SNN with existing methods (values from **Table 11**).

Table 12. Difference between NAMG-SNN and each baseline. Classification differences are in percentage points; the inference-time column is a relative reduction.

Compared Method	Accuracy (pp)	F1-Score (pp)	AUC (pp)	Inference Time Reduction (%)
CNN	+7.57	+7.77	+6.03	26.79
ResNet-50	+6.05	+6.08	+4.66	24.07
DenseNet-121	+5.13	+5.20	+3.79	21.90
EfficientNet-B3	+4.17	+4.25	+3.01	18.81
Vision transformer	+3.29	+3.31	+2.29	16.33
Graph neural network	+2.73	+2.79	+1.75	14.58
Graph transformer	+2.20	+2.18	+1.19	12.77

5. Discussion

The principal finding of this study is that selecting among spatial adjacencies, rather than merely representing them, accounts for a measurable share of the classification performance obtainable from pre-treatment H&E slides. The comparison in **Table 8** isolates this effect: holding the encoder, the spiking module and the classifier fixed, replacing an unmasked neighbourhood-aware graph with the masked variant raised accuracy by 1.03 percentage points and MCC from 0.928 to 0.947. Since the two configurations differ only in which edges survive, the difference is attributable to edge selection rather than to model capacity, which increases in neither direction.

This result is consistent with the biology reported in the spatial literature. Studies that profile the microenvironment directly have repeatedly found that the informative structure is local and compartmentalized: the interface between tumour nests and lymphocyte aggregates, the organization of tertiary lymphoid structures, and the composition of the stromal barrier [6,7]. A geometric neighbourhood rule spans such interfaces indiscriminately, so a tile of necrotic tissue adjacent to a tumour-immune interface enters the receptive field on the same terms as the interface itself. Conditioning retention on embedding similarity removes a subset of these edges before propagation, which offers a mechanistic account of the observed gain, though the present experiments do not verify that the removed edges are the artefactual ones. A direct test would require annotating edges by tissue compartment and measuring the composition of the retained subgraph, which was not performed here.

The second observation concerns cost. Inference time fell monotonically with the spatial prior, reaching 82 ms against 112 ms for the convolutional baseline, a reduction of 26.79%. Two mechanisms plausibly contribute: masking reduces the number of edges over which messages are passed, and threshold-gated firing means that a node transmits only when its accumulated potential crosses V_{th} , so activations are binary and sparse rather than dense and real-valued. This combination is unusual among the methods compared, in which accuracy and cost

typically trade against one another.

The ordering across cancer cohorts in **Table 6** merits comment. Melanoma gave the strongest result and lung adenocarcinoma the weakest, a 1.46 percentage point spread. Melanoma is among the most heavily infiltrated tumour types and its immune architecture is comparatively well organized, which would favour a method that depends on the coherence of local neighbourhoods. The present design cannot separate that explanation from the alternatives, namely differing cohort sizes, differing label quality across sources, or differing staining characteristics, and the cohorts are too small for the differences to be regarded as established.

The transcriptomic results in **Table 9** show agreement with the Phoenix reference rising from TCGA to the integrated cohort. The direction is expected but its interpretation is limited in an important respect: the Phoenix reference is itself inferred from histology rather than measured by a spatial assay [12]. Agreement therefore establishes that the proposed representation reproduces what a second histology-derived model predicts, not that it recovers measured gene expression. Establishing the latter would require evaluation against a matched, directly measured spatial transcriptomic cohort, and no such comparison is reported here.

Relative to the published literature, the level of discrimination reported in **Table 11** is at the upper end of what has been reported for ICI response prediction from histology alone, where AUC values in the range of roughly 0.70 to 0.85 are more typical for spatially aware models on independent cohorts [14,29]. That gap is large enough to warrant caution rather than confidence. Two explanations are consistent with the present evidence, and the study as designed does not distinguish between them: the architecture may genuinely improve on prior formulations, or the assembled cohort may be more separable than those used elsewhere, for example through label definitions that exclude ambiguous cases. Evaluation on an external cohort would be required to separate the two, and the reported values should be read as internal-validation estimates until such an evaluation is available.

The practical significance of this work is therefore methodological rather than clinical. The contribution is a demonstration that evidence-dependent edge retention and sparse spiking propagation are compatible with, and appear to benefit, graph learning on gigapixel histology. Whether the resulting score carries information that would be useful in patient management is a separate question that this retrospective design does not address and that would require prospective evaluation with pre-specified endpoints.

6. Limitations

The study is subject to the following limitations, several of which bear directly on how the reported values should be interpreted.

- (1) **Retrospective design.** All results are retrospective and computational. No prospective evaluation, no clinical endpoint and no reader study were performed, and the reported classification metrics do not constitute evidence of clinical utility.
- (2) **Single evaluation cohort.** Evaluation used one integrated cohort with an internal patient-level partition. No external, geographically or temporally independent cohort was used, so generalization beyond the assembled data is untested and the reported values are internal-validation estimates.
- (3) **Transcriptomic reference is inferred, not measured.** The Phoenix targets are model-derived from histology [12], so the agreement reported in **Table 9** measures concordance with another computational prediction rather than with assayed gene expression.
- (4) **Point estimates without uncertainty intervals.** Performance is reported as single point estimates from one partition. Repeated partitioning and interval estimation would be required to characterize sampling variability, and the calibration of the predicted probability has not been examined.
- (5) **Label heterogeneity across sources.** Response labels were drawn from more than one source, and residual differences in how response was assessed across those sources cannot be excluded.
- (6) **Unverified masking mechanism.** The interpretation that masking removes artefactual and necrotic edges is a plausible account of the observed gain but was not verified against compartment annotations.
- (7) **Interpretability.** The framework produces a scalar score without spatial attribution. It does not indicate which tissue regions drove a given prediction, which limits its use for biological hypothesis generation.
- (8) **Class balance and cohort size.** The cohort is nearly balanced by construction, which may not reflect the response rates observed in unselected populations, and the per-cohort sample sizes of 284 to 412 patients

are modest for the differences discussed in Section 5.

7. Conclusion

This paper developed the Neighborhood-aware Masked Graph Spiking Neural Network for retrospective classification of immune checkpoint inhibitor response from pre-treatment H&E whole-slide images. The framework represents a slide as a spatial graph over tissue tiles, retains an edge only when the embeddings at its endpoints are sufficiently similar, propagates the pruned graph through leaky integrate-and-fire dynamics, and classifies the resulting representation.

On an integrated cohort assembled from TCGA, the PUMA challenge and the Phoenix resource, the framework reached 97.31% accuracy, 96.94% precision, 97.82% recall, 97.38% F1-score and 98.87% AUC, exceeding convolutional, transformer and graph baselines trained under matched conditions while requiring 82 ms per inference against 94 to 112 ms for those baselines. Controlled ablation attributed 1.03 percentage points of accuracy specifically to adaptive edge masking, holding the remainder of the pipeline fixed, and showed that graph construction, masking and spiking propagation contribute approximately additively.

The methodological conclusion supported by these experiments is that edge selection is a useful design variable in histology graph learning, and that sparse event-driven propagation is compatible with competitive accuracy at reduced inference cost. The results are retrospective and were obtained on a single integrated cohort with an internal partition; they establish computational classification performance and not clinical utility.

Three directions follow. First, the framework should be evaluated on an external cohort that is independent in both institution and time, with repeated partitioning and interval estimation to characterize sampling variability. Second, the transcriptomic head should be validated against directly measured spatial transcriptomic data rather than against inferred targets. Third, spatial attribution over the retained subgraph should be developed, so that a prediction can be traced to the tissue regions that produced it, which is a precondition for any subsequent investigation of biological validity.

Author Contributions

Conceptualization, L.R. and B.M.; methodology, L.R. and B.M.; software, L.R. and B.R.; validation, B.M., S.B.T. and V.K.; formal analysis, L.R. and M.; investigation, L.R., B.M. and S.B.T.; resources, B.M., S.B.T. and M.; data curation, L.R. and B.R.; writing—original draft preparation, L.R. and B.M.; writing—review and editing, L.R., B.M., S.B.T., B.R., V.K., and M.; visualization, L.R. and B.R.; supervision, B.M. and S.B.T.; project administration, L.R. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Ethical review and approval were not required for this study because the research involved the secondary computational analysis of previously available digital pathology datasets and did not involve the recruitment of human participants or the collection of new human-subject data.

Informed Consent Statement

Not applicable, as this study involved secondary analysis of previously available, de-identified digital pathology datasets and did not involve direct interaction with human participants or collection of new participant data.

Data Availability Statement

The data analyzed in this study were obtained from previously established digital pathology datasets and publicly available or institutionally accessible repositories. The study did not generate a new clinical dataset. The source datasets are subject to the access, registration, and/or data-use requirements of their respective repositories. All data supporting the findings of this study are available from the corresponding public repositories or

through the appropriate data-access procedures. The code and analytical procedures used in this study can be made available by the corresponding author upon reasonable request, subject to applicable dataset and repository restrictions.

Conflicts of Interest

The authors declare no conflict of interest.

AI Use Statement

During the preparation of this work, the authors used ChatGPT for assistance with language refinement, manuscript organization, and figure presentation. The authors subsequently reviewed and edited the content as necessary and take full responsibility for the final content of the published article.

References

1. Schuveling, M.; Liu, H.; Eek, D.; et al. Associations between spatial distribution of immune cell subsets and clinical outcomes in patients with advanced melanoma treated with immune checkpoint inhibitors: Results from the PUMA challenge. *Med. Image Anal.* **2026**, *115*, 104334.
2. Williams, H.L.; Frei, A.L.; Koessler, T.; et al. The current landscape of spatial biomarkers for prediction of response to immune checkpoint inhibition. *Npj Precis. Oncol.* **2024**, *8*, 178.
3. Nosa-Ihaza, E.A.; Edeh, E.C.; Geng, W.B.; et al. Artificial intelligence in cancer immunotherapy: Current trends in predicting response and personalizing treatment. *J. Egypt. Natl. Cancer Inst.* **2026**, *38*, 30.
4. Liu, W.; Puri, A.; Fu, D.; et al. Dissecting the tumor microenvironment in response to immune checkpoint inhibitors via single-cell and spatial transcriptomics. *Clin. Exp. Metastasis* **2024**, *41*, 313–332.
5. Zhou, X.M.M.; D'Amiano, A.J.; Lu, C.; et al. Spatial proteomics for investigating solid tumor resistance mechanisms. *Cancer Metastasis Rev.* **2025**, *44*, 76.
6. Greenwald, N.F.; Nederlof, I.; Sowers, C.; et al. Temporal and spatial composition of the tumor microenvironment predicts response to immune checkpoint inhibition in metastatic TNBC. *Nat. Cancer* **2026**, *7*, 435–450.
7. Peyraud, F.; Guegan, J.P.; Rey, C.; et al. Spatially resolved transcriptomics reveal the determinants of primary resistance to immunotherapy in NSCLC with mature tertiary lymphoid structures. *Cell Rep. Med.* **2025**, *6*, 101915.
8. Karlsen, V.T.; Majid, U.; Lindholm, H.T.; et al. Single cell spatial transcriptomics identifies coordinated cellular programs that are associated with good prognosis in microsatellite stable colorectal cancer. *bioRxiv* **2025**. [[CrossRef](#)]
9. Xin, Y.; Yang, L.; Liu, C.; et al. Spatial transcriptomics of immune ecotypes for predicting immunotherapy outcomes in head and neck squamous cell carcinoma. *Oral Oncol.* **2026**, *175*, 107887.
10. Cui, X.; Liu, S.; Song, H.; et al. Single-cell and spatial transcriptomic analyses revealing tumor microenvironment remodeling after neoadjuvant chemoimmunotherapy in nonsmall cell lung cancer. *Mol. Cancer* **2025**, *24*, 111.
11. Han, Y.B.; Lee, S.; Lee, J.O.; et al. Spatial transcriptomics reveal high T cell and monocyte status as predictive and prognostic markers in pancreatic cancer. *J. Transl. Med.* **2025**, *23*, 576.
12. Tran, M.; Gindra, R.H.; Putze, P.; et al. Pan-cancer virtual spatial transcriptomics from routine histology with Phoenix. *bioRxiv* **2026**. [[CrossRef](#)]
13. Mukherjee, S.; Patiyal, S.; Pal, L.R.; et al. Pan-cancer prediction of tumor immune activation and response to immune checkpoint blockade from tumor transcriptomics and histopathology. *bioRxiv* **2025**. [[CrossRef](#)]
14. Shen, Z.; Brand, D.; Simard, M.; et al. AI-powered immune profiling from histopathology slides for chemoradiotherapy outcome prediction in rectal cancer: A study using clinical trial and real-world cohorts. *EBioMedicine* **2025**, *122*, 105993.
15. Howard, F.M.; Dolezal, J.; Hieromnimon, H.; et al. Integration of gene expression and digital histology to predict treatment-specific responses in breast cancer. *medRxiv* **2025**. [[CrossRef](#)]
16. Belete, M.; Thakkar, N.; Khatri, I.; et al. Integrated histopathology-transcriptomic biomarker enhances survival prediction in HNSCC patients treated with immunotherapy. *Transl. Oncol.* **2026**, *70*, 102822.
17. Al-Mansour, F.S.H.; Almasoudi, H.H.; Albarrati, A. Mapping molecular landscapes in triple-negative breast cancer: Insights from spatial transcriptomics. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2025**, *398*, 11125–11143.

18. Lubo, I.; Hernandez, S.; Wistuba, I.I.; et al. Novel spatial approaches to dissect the lung cancer immune microenvironment. *Cancers* **2024**, *16*, 4145.
19. Ma, A.; Xiang, L.; Yuan, J.; et al. Spatial transcriptomics decodes breast cancer microenvironment heterogeneity: From multidimensional dynamic profiling to precision therapy blueprint construction. *Biomolecules* **2025**, *15*, 1067.
20. Manogaran, S.; Ramadoss, R.; Selvam, S.P.; et al. Artificial intelligence-driven spatial transcriptomics in OSCC: Mapping the tumor microenvironment and personalizing therapy. *J. Oral Biol. Craniofac. Res.* **2025**, *15*, 1862–1873.
21. Voglis, S.; Schulz, D.; Eling, N.; et al. Spatially-resolved single-cell imaging of melanoma brain metastases identifies localized immune patterns predictive of immune checkpoint blockade response. *Neuro-Oncol.* **2026**, *8*, 2211–2223.
22. Ran, L.; Hu, G.; Fan, C.; et al. Deciphering the mechanistic landscape of immune checkpoint blockade in ccRCC: From molecular drivers to therapeutic responses. *Front. Immunol.* **2026**, *17*, 1774959.
23. Di Mauro, A.; De Cecio, R.; Simonelli, S.; et al. Modern pathology-driven strategies in neoadjuvant immunotherapy for head and neck squamous cell carcinoma: From residual tumor quantification to spatial and AI-based biomarkers. *Cancers* **2026**, *18*, 1020.
24. Choo, J.; Zhao, J.J.; Lau, M.C.; et al. Spatial predictors of response to chemo-immunotherapy in microsatellite stable metastatic colorectal cancer. *Nat. Commun.* **2026**, *17*.
25. Nameki, R.; Kinong, J.; Huang, C.H.; et al. Spatial remodeling of the urothelial carcinoma tumor microenvironment shapes response to neoadjuvant atezolizumab. *medRxiv* **2026**. [[CrossRef](#)]
26. Shamis, S.A.; Savioli, F.; Ammar, A.; et al. Spatial transcriptomic analysis of tumour with high and low CAIX expression in TNBC tissue samples using GeoMx™ RNA assay. *Histol. Histopathol.* **2024**, *39*, 177–200.
27. Yu, C.; Shi, C.; Zhou, H.; et al. Spatially resolved whole-transcriptomic profiling of EGFR-mutated lung adenocarcinomas stratified by PD-L1 expression. *Cell. Oncol.* **2026**, *49*, 2.
28. Zheng, Y.; Sadee, C.; Ozawa, M.; et al. Single-cell multimodal analysis reveals tumor microenvironment predictive of treatment response in non-small cell lung cancer. *Sci. Adv.* **2025**, *11*, eadu2151.
29. Pybus, A.; Kirchgaessner, R.; Nguyen, J.; et al. Predicting anti-PD-1 immune checkpoint blockade response in melanoma patients with spatially aware machine learning models. *Npj Precis. Oncol.* **2026**, *10*.
30. Baena, J.C.; Victoria, J.S.; Toro-Pedroza, A.; et al. Smart CAR-T nanosymbionts: Archetypes and proto-models. *Front. Immunol.* **2025**, *16*, 1635159.



Copyright © 2026 by the author(s). Published by UK Scientific Publishing Limited. This is an open access article under the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Publisher's Note: The views, opinions, and information presented in all publications are the sole responsibility of the respective authors and contributors, and do not necessarily reflect the views of UK Scientific Publishing Limited and/or its editors. UK Scientific Publishing Limited and/or its editors hereby disclaim any liability for any harm or damage to individuals or property arising from the implementation of ideas, methods, instructions, or products mentioned in the content.