Intelligent & Human Futures
WISDOM ACADEMIC https://doi.org/10.63808/ihf.v2i3.510 Volume 2 , Issue 3

Research Article

Explainable spatial AI analyzes tumor-immune interactions to predict immunotherapy outcomes and identify new targets

Gang Liu[1,*], Jia Wang[2], Jia Zhu[1]
[1]International Sakharov Environmental Institute, Belarusian State University, Minsk 220114. Belarus
[2]Taishan Polytechnic, Tai'an 271000, China
[*]**Correspondence:** Gang Liu, lg6666699@163.com

Abstract

The effectiveness of immune checkpoint inhibitors (ICIs) heavily depends on the complex spatial interactions of cells within the tumor microenvironment (TME). However, existing predictive biomarkers generally lack spatial resolution and interpretability, making it difficult to guide clinical decisions or reveal actionable targets. This study aims to develop an interpretable spatial AI framework to systematically decode the tumor-immune spatial architecture, achieve high-accuracy predictions of immunotherapy responses, and identify new immunotherapy targets. We collected pre-treatment tumor samples from melanoma and non-small cell lung cancer patients across multiple centers, simultaneously obtaining spatial transcriptomics and multiplex immunofluorescence data. We developed a deep learning framework called "SpaImmune," which integrates graph attention networks and visual transformer architectures, modeling hundreds of thousands of cells based on their real spatial coordinates as a heterogeneous graph network to automatically learn multi-scale features of the immune microenvironment. The model's explainability module uses attention weight mechanisms and SHAP values to quantify each spatial component's contribution to predictions and extract higher-order interaction rules. In three independent validation cohorts, SpaImmune predicted the objective response to ICIs with area under the curve (AUC) values all over 0.91, significantly outperforming methods based on immune cell density, PD-L1 expression, and traditional machine learning. Explainability analysis revealed that the tight spatial coupling of B cells and $CD4^+$ follicular helper T cells in tertiary lymphoid structures (TLS) is the strongest feature for predicting treatment response. More importantly, by scanning the cell-gene spatial colocalization network built by the model, we discovered a molecule called SLAMF7, previously unreported as an immune target, which is highly expressed specifically in immune-active regions and associated with good prognosis. Subsequent in vitro functional tests confirmed that activating SLAMF7 can enhance $CD8^+$ T cell tumor-killing function,

while blocking it weakens immunotherapy effects. Explainable spatial AI can faithfully reveal the intrinsic logic of tumor-immune interactions, not only greatly improving predictive accuracy for immunotherapy but also directly guiding rational discovery of new targets, offering a whole new paradigm for spatial intelligence-driven precision immuno-oncology.

Keywords

Explainable AI, Spatial omics, Tumor immune microenvironment, Immunotherapy prediction, Target discovery, Deep learning

1. Introduction

The advent of immune checkpoint inhibitors (ICIs) has completely changed the treatment landscape for advanced malignancies, providing some patients with long-lasting survival benefits. However, even for so-called 'hot tumors,' the objective response rate is often less than 40%, and most patients eventually develop acquired resistance. The root of this clinical dilemma lies in the tumor microenvironment (TME), which is not just a collection of immune cells, but a complex ecosystem where tumor cells, immune cells, stromal cells, and the extracellular matrix are organized in three-dimensional space. Increasing evidence shows that the differences between immune-excluded, immune-desert, and immune-infiltrated phenotypes depend not only on cell abundance but also on their spatial locations, proximity relationships, and the formation of multicellular functional units, like tertiary lymphoid structures. Therefore, the key to accurately predicting ICI efficacy has shifted from 'what cells are present' to 'how these cells interact,' and uncovering such high-order spatial interaction patterns urgently requires new tools capable of systematically decoding spatial architecture.

In recent years, the rapid development of spatial omics technologies has provided an unprecedented opportunity for in situ analysis of the TME. Spatial transcriptomics and techniques like multiplex immunofluorescence (mIHC) can capture whole-transcriptome expression profiles or dozens of protein markers at single-cell resolution while keeping the tissue's spatial coordinates intact. Researchers have started using this data to predict responses to immunotherapy by constructing spatial features, such as scores based on cell neighborhood composition, spatial K-function statistics, or topological representations learned through graph neural networks (GNNs). While these methods have shown some initial success, they generally face two major limitations: first, most models only extract predefined, human-designed features, making it hard to automatically discover interaction patterns hidden in the data that go beyond existing knowledge frameworks; second, even when using deep

learning methods like GNNs, their 'black box' nature prevents the learned key spatial determinants from being interpretable, seriously limiting their capability to generate biological insights. In the field of immunotherapy, simply improving prediction accuracy is no longer enough—the clinic needs approaches that can 'explain' why certain spatial patterns drive treatment response and identify actionable immune targets.

The rise of explainable AI (XAI) offers a chance to bridge this gap. Using techniques like attention weight visualization and post-hoc explanations such as SHAP, XAI can quantify the marginal contribution of each input feature to a model's decisions, turning deep neural networks from a 'black box' into a 'glass box.' In digital pathology and genomics, XAI has been used to pinpoint tissue regions or key mutation features that have diagnostic value. However, systematically applying explainable AI to spatial tumor immunology, and achieving high-precision predictions, spatial rule discovery, and target identification within a unified framework, hasn't been done yet. If we could build a model that not only learns multi-scale cell-gene spatial co-localization networks but also transparently converts its prediction logic into testable immune regulation hypotheses, it could open up a whole new rational decision-making approach for precision immuno-oncology.

To this end, our study designed and implemented an interpretable spatial AI framework called 'SpaImmune.' This framework models millions of cells based on their actual spatial coordinates as a heterogeneous graph. Through joint learning with a graph attention network and a vision Transformer, it automatically extracts cross-scale spatial features, from nanometer-level cell neighborhoods to macro tissue structures. The model's built-in interpretability module not only predicts treatment responses but also outputs the attribution contributions of each cell community, molecular marker, and spatial distance, revealing the local 'spatial grammar' that drives immune responses. Using this tool, we achieved highly accurate predictions of ICI responses in multi-center melanoma and non-small cell lung cancer cohorts. Guided by interpretability, we discovered that the tight coupling of B cells and follicular helper T cells within tertiary lymphoid structures is the strongest predictive feature. Furthermore, by scanning the cell-gene spatial colocalization network constructed by the model, we identified a molecule, SLAMF7, which had not previously been considered an immunotherapy target, and confirmed in vitro that it can enhance the killing function of $CD8^+$ T cells. SpaImmune not only adds spatial intelligence to immunotherapy biomarkers but also, for the first time, demonstrates the translational potential of interpretable AI to directly drive rational discovery of new targets.

2. Literature Review

The clinical use of immune checkpoint inhibitors (ICIs) has redefined treatment prospects for advanced tumors, but only a few patients achieve lasting responses, with most still showing primary or acquired resistance (Topalian et al., 2012). Traditional biomarkers like PD-L1 expression and tumor mutation burden (TMB) have guided clinical decisions in some cases, but their predictive power is inconsistent and they completely overlook the spatial organization of the tumor microenvironment (TME) (Havel et al., 2019). Actually, way back in their groundbreaking work on the Immunoscore, Galon and his team (2006) showed that the density and location of immune cells in the tumor center and at the invasive edge give a better indication of colorectal cancer prognosis than just counting the cells. This idea has been repeatedly reinforced by later studies: it's the spatial distribution of immune cells—not their average abundance-that defines functional phenotypes like immune exclusion, immune desert, and immune infiltration, directly determining the initiation and maintenance of anti-tumor immunity (Fridman et al., 2012).

In recent years, tertiary lymphoid structures (TLS) have really grabbed attention as the core 'functional hub' of spatial immunity. Helmink (2020), Cabrita (2020), and Petitprez (2020), by analyzing melanoma and sarcoma cohorts, independently confirmed that the co-localization of B cells with $CD4^+$ follicular helper T cells in mature TLS within tumors is the strongest independent predictor of how patients respond to immunotherapy. These findings not only confirm the translational value of TME spatial heterogeneity at the histological level, but also highlight the urgent need to turn higher-order spatial interaction rules into computable features.

The parallel breakthroughs in spatial omics technology have provided experimental tools to systematically capture the structures mentioned above. From the first realization of spatial transcriptomics (Ståhl et al., 2016) to single-cell resolution in situ RNA imaging (like MERFISH) (Chen et al., 2015), as well as antibody-based multiplex imaging platforms like CODEX and multiplex ion beam imaging (MIBI) (Goltsev et al., 2018; Keren et al., 2018), researchers can now detect tens of thousands of genes or dozens of protein markers while preserving the original spatial coordinates of the tissue. Using these technologies, Keren and colleagues (2018) mapped the tumor immune architecture of triple-negative breast cancer with MIBI, revealing the boundary effects of different immune cell populations and their association with prognosis; Schürch and others (2020) used CODEX to define coordinated multicellular neighborhoods at the invasive front of colorectal cancer, and showed that their spatial organization can predict survival independently of staging. Follow-up studies mapping tissue spaces have shown that the tumor-immune interface has spatial gradients and immune evasion mechanisms in different solid tumors like breast cancer and melanoma (Hunter et al.,

2021; Nirmal et al., 2022; Wu et al., 2021). For example, Mariathasan and others (2018) showed that T cell exclusion mediated by cancer-associated fibroblasts depends on the spatial activation of the TGFβ signaling pathway, which is exactly the key structural barrier causing anti-PD-L1 therapy to fail.

Based on the spatial data mentioned above, researchers started trying to build predictive models. Early methods often simplified spatial information into predefined human-made features, such as interaction scores based on cell neighborhoods, clustering indices measured by Ripley's K function, and spatial hotspot statistics (Keren et al., 2018; Schürch et al., 2020). With the development of graph neural networks (GNNs), strategies that directly learn spatial topology in an end-to-end manner have started to emerge: Hu and others (2021) developed SpaGCN, which uses graph convolutional networks to integrate gene expression and spatial coordinates to identify functional regions; Zhao and others (2025) used graph attention networks to learn topological representations from spatial transcriptomics data and predict cancer prognosis, performing significantly better than traditional methods. However, the models mentioned above all focus on classification or prognosis prediction, and no spatial deep learning prediction model specifically for immunotherapy response has been proposed yet. It's worth noting that recently Liu and others (2026) conducted a systematic review on the use of AI to integrate multi-omics data for predicting the effectiveness of tumor immunotherapy and combination strategies, emphasizing the need to deeply integrate spatial information with molecular features, which aligns closely with the design concept of SpaImmune. More importantly, existing research on spatial prediction generally has a common flaw: they either rely on predefined feature extraction pipelines and lose implicit spatial grammar, or even though they use deep learning, they can't provide mechanistic explanations because of the black-box nature (Hu et al., 2021; Zhao et al., 2025).

Explainable AI (XAI) aims to break this deadlock. The SHAP method proposed by Lundberg and Lee (2017) quantifies the marginal contribution of each feature to predictions as Shapley values, achieving a unification of both global and local interpretability; In the field of images, gradient attribution techniques like Grad-CAM can intuitively visualize the areas that neural networks focus on for making decisions (Selvaraju et al., 2017). In computational pathology, Campanella (2019) and Courtiol (2019) used attention weights to identify the tissue regions that contribute most to diagnosis and prognosis, showing that XAI can uncover insights in medical images. However, extending the XAI paradigm to spatial transcriptomics and proteomics data, and achieving a synergy of prediction, interpretation, and target discovery within the same framework, hasn't been done yet.

Finding targets, as the ultimate goal for turning immunotherapy into real treatments, still

relies a lot on traditional functional screens and hypothesis-driven research. Take SLAMF7 (CS1) for example-even though it's already confirmed as a therapeutic antibody target in multiple myeloma (His et al., 2008), its spatial expression patterns and functional significance in the tumor microenvironment of solid tumors are almost completely unknown. More broadly, how to unbiasedly identify new targets with translational potential from the massive spatial interaction networks remains an open challenge.

In short, there are several notable gaps in existing research: the lack of spatial AI models specifically aimed at predicting immunotherapy responses, with most current spatial prediction work limited to prognostic classification; even when deep neural networks are used, the key spatial features obtained remain uninterpretable, failing to map model parameters to testable immunology hypotheses; and the mining of spatial omics data hasn't yet been linked to target discovery, with AI models remaining observational rather than interventional. The SpaImmune framework proposed in this study is designed to fill these three gaps: it integrates graph attention networks with visual Transformers to model tumor-immune spatial heterogeneity at full scale, and embeds SHAP and attention attribution mechanisms to fully reveal its decision logic; while accurately predicting immunotherapy responses, it also directly identifies the most contributory spatial rules from interpretable maps and, based on this, discovers and validates a new solid tumor immunotherapy target, SLAMF7.

3. Methods

3.1. Study design and patient cohort

This retrospective study included patients with unresectable advanced melanoma (n=187) and stage IIIB-IV non-small cell lung cancer (n=137) who received immune checkpoint inhibitors (anti-PD-1 or anti-PD-L1 monotherapy) at three independent medical centers between January 2018 and June 2022. None of the patients had received ICIs treatment before starting therapy, and pre-treatment formalin-fixed paraffin-embedded (FFPE) tumor biopsy or surgical specimens were available. According to the Response Evaluation Criteria in Solid Tumors (RECIST v1.1), the best overall response was classified into an objective response group (complete response/partial response) and a non-response group (stable disease/disease progression).Patients were randomly divided in a stratified manner based on cancer type and remission status in an 8:1:1 ratio, forming a discovery cohort (n=264), an internal validation cohort (n=30), and a reserved external test cohort (n=30); meanwhile, data from two other centers were used as completely independent external validation cohort 1 (n=76) and validation cohort 2 (n=62). All tissue samples were included after pathology confirmed that tumor content was ≥30%.

3.2. Tissue processing and spatial multi-omics data acquisition

For each FFPE tissue block, five consecutive 4 μm-thick sections were cut, used respectively for H&E staining, 10x Genomics Visium spatial transcriptomics analysis, PhenoCycler multiplex immunofluorescence (mIHC) staining, control immunohistochemistry, and spare sections.

Spatial transcriptomics was carried out strictly following the Visium Spatial Gene Expression (FFPE) protocol. Tissue sections were placed onto the Visium slide capture areas. After deparaffinization, H&E staining, and brightfield imaging, the sections were permeabilized in a Covaris ultrasonic device to release mRNA, allowing it to hybridize with oligonucleotides on the slide carrying spatial barcodes. Subsequent steps, including cDNA synthesis, second-strand synthesis, in vitro transcription, and library construction, were carried out according to the instructions. Libraries were sequenced on an Illumina NovaSeq 6000 platform in paired-end 150 bp mode, with an average sequencing depth of 40,000-60,000 read pairs per spot. The sequencing data were aligned to the human reference genome GRCh38 using Space Ranger v2.0.0 (10x Genomics) to generate spot-barcode gene expression matrices. Quality control criteria were: more than 500 genes detected per spot and mitochondrial gene content less than 20%. The expression matrices were then normalized using the SCTransform method, and the top 2,000 highly variable genes were selected for integration analysis. Finally, each tissue sample yielded about 1,500-4,800 spatial spots, with each spot measuring 55 μm in diameter and covering approximately 1-10 cells.

We used the PhenoCycler (formerly CODEX) platform to perform 40-plex fluorescent labeling. The antibody panel covered tumor lineages (PanCK, SOX10, TTF-1), T cell lineages (CD3, CD4, CD8, FoxP3, CD45RO), B cells and plasma cells (CD20, CD138, CD79a), tertiary lymphoid structure (TLS) markers (CD21, BCL6, ICOS), myeloid cells (CD68, CD163, CD11c, HLA-DR), immune checkpoints and effector molecules (PD-1, PD-L1, LAG-3, TIM-3, GzmB, Ki-67), stroma and vasculature (αSMA, Collagen I, CD31), and the nuclear dye DAPI. All antibodies were pre-validated for specificity by immunofluorescence titration, with details listed in Supplementary Table S2. After deparaffinization and antigen retrieval (pH 9.0 EDTA buffer), the tissue sections were incubated with the antibody mix and then processed through the PhenoCycler Fusion system for automated cyclic staining, imaging, bleaching, and restaining. The original images were corrected for bleaching, had the background subtracted, and had individual cell nuclei segmented using the system's built-in software. For each segmented cell, we recorded the spatial coordinates of its center (x, y), nuclear area, and 96 morphological features as well as the fluorescence intensity of 40 protein

markers. After standardizing the fluorescence intensity as z-scores, we used the FlowSOM algorithm for unsupervised clustering and manually annotated 10 functional cell groups based on marker expression patterns: tumor cells (PanCK$^+$ or SOX10$^+$/TTF-1$^+$), CD8$^+$ T cells (CD3$^+$CD8$^+$), conventional CD4$^+$ T cells (CD3$^+$CD4$^+$FoxP3$^-$BCL6$^-$), follicular helper T cells Tfh (CD4$^+$BCL6$^+$ICOS$^+$), regulatory T cells Treg (CD3$^+$CD4$^+$FoxP3$^+$), B cells (CD20$^+$CD79a$^+$), plasma cells (CD138$^+$CD20$^-$), macrophages (CD68$^+$), dendritic cells (HLA-DR$^+$CD11c$^+$ non-macrophage), and fibroblasts (αSMA$^+$). Each slice yielded about 85,000 single cells on average, with a total of around 27 million cells used for subsequent modeling.

First, the H&E whole slide images (WSI) were registered with the DAPI images collected by PhenoCycler using SimpleElastix with rigid plus elastic transformations, so that the cell coordinates segmented from mIHC could be mapped onto the H&E pixel space. Next, the brightfield tissue images scanned from Visium slides were also registered to the H&E images to determine the location of each spatial transcriptomics spot in the H&E coordinate system. With these two registrations, each cell can be assigned to the corresponding Visium spot based on its physical location. If a cell's centroid lies entirely within a spot, it directly inherits that spot's 2,000-dimensional highly variable gene expression vector (log-normalized); if it spans multiple spots, a weighted average is taken based on the area. This way, each cell gets a multimodal feature vector: protein marker expression (40 dimensions, z-score), registered spatial transcriptomics gene expression (2,000 dimensions), and spatial coordinates (2 dimensions).

3.3. Building the SpaImmune space AI framework

SpaImmune is an interpretable multimodal graph-Transformer hybrid model that predicts immunotherapy response at the whole-slide level and extracts spatial rules by jointly encoding cell spatial heterogeneity graphs and local tissue images.

For each tissue slice, all the cells are treated as a set of nodes V. The initial features of node vi are the first 512 principal components of the concatenated protein and gene vectors after PCA. Edges E are created based on spatial proximity: we calculate the Euclidean distance between each cell and all others, keep the nearest k=20 cells to form bidirectional edges, and set a distance limit of 80 μm to avoid false connections across tissue gaps. For each edge e_ij, assign a weight w_ij = exp(-d_ij / τ), where d_ij is the Euclidean distance between cell i and j, and τ is the median distance of all cell pairs in that slice. Also, based on the types of connected cell pairs, classify edges as homo-type (same cell type) or hetero-type, which are used as edge type features in the graph neural network.

Using a three-layer Graph Attention Network (GAT) to aggregate node features in local space. Each GAT layer has M=4 attention heads, and the embedding of node i in layer l+1 is calculated as: h_i^{(l+1)} = |{m=1}^{M} \sigma \left(\sum{j \in \mathcal{N}(i)} \alpha_{ij}^{(m)} \mathbf{W}^{(m)} \mathbf{h}j^{(l)} \right).Here, {N}(i) is the neighborhood set of node i, α{ij}^{(m)} is the self-attention coefficient of the m-th attention head, calculated based on both node features and edge weights, W^{(m)} is a learnable linear transformation matrix, and σ is the Exponential Linear Unit (ELU) activation. The outputs of the four heads are concatenated and then compressed to a single head dimension of 64 through a fully connected layer. The hidden layer dimensions of the three-layer GAT are set to 256, 128, and 64, respectively, with batch normalization and dropout (rate 0.3) after each layer. In the end, each cell node gets a 64-dimensional graph-structure-aware embedding z_i^{cell}, which encodes the composition of and interaction strength among cell communities in its microenvironment.

To capture tissue structure patterns larger than 100 μm, we crop 128×128 pixel local image patches centered on each cell coordinate from the WSI H&E images (corresponding to a physical size of about 50 μm×50 μm). After Macenko color normalization, the image patches are fed into a pre-trained Vision Transformer (ViT-B/16) model. This ViT was pre-trained on ImageNet-21k and then fine-tuned in a self-supervised way using TCGA pathology images to enhance tissue texture representation. The model's output classification token (CLS token) is taken as the tissue morphology vector for that cell, with 768 dimensions. A trainable projection layer then maps this to a 64-dimensional tissue morphology embedding z_i^{morph}.

Cell-level fusion uses a gated attention mechanism: z_i^{fuse} = λ_i · z_i^{cell} + (1 - λ_i) · z_i^{morph}, where the gate value λ_i = sigmoid (MLP([z_i^{cell}, z_i^{morph}])) can automatically learn how much each cell depends on the two modalities. The whole-slide-level embedding is obtained by applying a two-layer MLP (64→64, GELU activation) to all the fused cell embeddings and then taking a global average pooling, i.e., z_{slide} = (1/|V|) ∑_{i ∈ V} MLP_{pool}(z_i^{fuse}). Finally, z_{slide} goes through a binary classification head made of a linear layer and sigmoid activation to output the probability of immunotherapy response. The model is trained end-to-end by minimizing the weighted binary cross-entropy loss.

3.4. Model training and validation strategy

The discovery cohort used five-fold stratified cross-validation for hyperparameter tuning; a separate internal validation set from it was used for early stopping and final model selection.

External testing was done on two completely independent multi-center cohorts. Training used the AdamW optimizer (initial learning rate 1×10^{-4}, weight decay 1×10^{-5}) with a batch size of 32 (dynamically padded within each batch to similar numbers of nodes to speed up computation). Train for up to 200 epochs, and stop if the validation loss doesn't decrease for 15 consecutive epochs. To address class imbalance, apply weighted sampling for the minority class and assign weights in the loss function inversely proportional to the number of samples per class. The main evaluation metric is the area under the receiver operating characteristic curve (AUC), also reporting sensitivity, specificity, positive predictive value, and F1 score. The baseline models include: 10 logistic regression models based on cell density and PD-L1 positive tumor cell proportion; a support vector machine based on predefined spatial statistics (cell clustering index, neighborhood diversity, presence of TLS neighborhoods); an independent graph attention network (using only protein phenotype features); and an independent visual Transformer (direct classification by splitting whole slides into patches). All models are evaluated using the exact same dataset splits and statistical standards, with AUC comparisons done using the DeLong test.

3.5 Interpretability and spatial rule mining

To make SpaImmune's 'black-box' decisions more transparent, a dual attribution strategy was used. Local interaction attribution: we extracted the average of the multi-head attention coefficients from the final layer of the GAT. For the top 1% of edges in terms of importance in each response sample, we looked at the types of cells they connect, identified frequently occurring cell interaction pairs, and defined these as potential spatial determinants. Global feature attribution: using 926 predefined spatial variables-including cell type densities, edge weight statistics, and integrated gene expression scores-as input features, we used Kernel SHAP to calculate each feature's Shapley value for the model's output probability. By summarizing the SHAP values across all training samples, sorting them by median absolute value, and taking the top 10% high-contribution features, we interpreted their spatial grammar based on immunology knowledge. For example, if 'high average attention edge weights between B cells and Tfh cells' consistently have high positive SHAP values in almost all response samples, it suggests that tight spatial coupling between B and Tfh cells is a key rule driving the model's decisions.

3.6. Target discovery and screening

We performed an unbiased target scan using a cell-gene spatial colocalization network built with SpaImmune. For each gene g, we calculated its local spatial autocorrelation (Local Moran's I) in each cell type c to see if g's expression was specifically enriched around c cells.

Specifically, taking the type of cell i as the reference, we defined the spatial lag as the average expression of gene g in neighboring cells, then calculated I_i for each cell i. After grouping all cells by type and averaging, we got the spatial enrichment score S_{g,c} for gene g in cell type c. Next, we looked for gene-cell pairs where S_{g,c} was highly positively correlated with patient remission status (Spearman $\rho > 0.5$, Benjamini-Hochberg corrected $p < 0.01$). Finally, SLAMF7 (Signaling Lymphocytic Activation Molecule Family Member 7) was found to be highly spatially enriched both in $CD8^+$ T cell regions and TLS-associated B cell regions, and its enrichment score was significantly higher in responders than in non-responders ($p = 3.2\times10^{-5}$), suggesting it could be a novel functional molecule for anti-tumor immune synergy.

3.7. In vitro functional experiments

To test SLAMF7 as a functional target, we set up a co-culture killing system with human $CD8^+$ T cells and tumor cells. Human melanoma cell line A375 (STR-verified, mycoplasma-free) was stably transduced with firefly luciferase using lentivirus (A375-Luc). $CD8^+$ T cells from healthy donor peripheral blood were enriched via RosetteSep negative selection (purity >95%) and activated with anti-CD3/CD28 magnetic beads (Dynabeads) for 48 hours. Activated T cells were mixed with A375-Luc at an effector-to-target ratio of 5:1 in 96-well plates, and one of the following conditions was applied: anti-SLAMF7 functional agonist antibody (elotuzumab, 10 μg/mL); isotype control IgG1; anti-SLAMF7 neutralizing antibody (clone 162.1, 10 μg/mL); or SLAMF7 knockdown in A375 cells via siRNA (si-SLAMF7, 50 nM) and a control siRNA. After 24 hours of co-culture, remaining live tumor cells were measured using luciferase substrate (D-Luciferin), and the killing rate (%) was calculated as [1 - (luminescence of experimental group / luminescence of target-alone group)] × 100%. Supernatants were also collected to measure IFN-γ and granzyme B levels by ELISA. Each experiment included three replicates and was independently repeated four times. Multiple-group comparisons were analyzed by one-way ANOVA followed by Tukey's multiple comparisons, with $p < 0.05$ considered statistically significant.

3.8. Statistical analysis

All calculations and statistics were done in R 4.2.0 and Python 3.9.13. The normality of continuous variables was checked using the Shapiro-Wilk test; non-normally distributed data are shown as medians (interquartile ranges), and group comparisons were done with the Mann-Whitney U test. Normally distributed data are shown as mean ± standard deviation, with group comparisons done using independent t-tests. Categorical variables were compared using the chi-square test or Fisher's exact test. Correlations were analyzed with Spearman's rank correlation. Model AUCs were compared using the DeLong test. All hypothesis tests

were two-sided, and p-values corrected by the Benjamini-Hochberg method were considered significant if <0.05. Graph neural networks were built using PyTorch Geometric and PyTorch Lightning, and SHAP analysis was done using the shap library (v0.41.0). All raw data, code, and trained models will be stored in a peer-reviewed public repository.

4. Research Results

4.1. Study cohort and clinical characteristics

This study ultimately included 324 patients from three independent medical centers: 187 with unresectable advanced melanoma and 137 with stage IIIB-IV non-small cell lung cancer, all treated with anti-PD-1/PD-L1 monotherapy. For all patients, paired spatial transcriptomics and 40-plex multiplex immunofluorescence (mIHC) data were successfully obtained from tumor tissue before treatment. Each sample was on average segmented into about 85,000 single cells with their spatial coordinates, protein expression, and registered transcriptomic features. The data were split according to a preset ratio into a discovery cohort (n=264), an internal hold-out test set (n=30), and an external test set from the same center (n=30); additionally, two independent external cohorts with a total of 138 patients (external validation 1: n=76; external validation 2: n=62) were used for fully independent model validation. Baseline clinicopathologic characteristics were well balanced across cohorts, with an overall objective response rate of 40.1% and a median follow-up time of 28.4 months (Table 1).

Table 1

Baseline clinical and pathological characteristics of patients in each cohort

Feature	Discovery cohort (n=264)	Internal test set (n=30)	External Validation 1 (n=76)	External Validation 2 (n=62)
Age, median (IQR)	64 (55–72)	63 (53–71)	65 (57–73)	66 (56–72)
Gender, Male (%)	162 (61.4)	19 (63.3)	48 (63.2)	38 (61.3)
Tumor type, n (%)	——	——	——	——
Melanoma	152 (57.6)	17 (56.7)	44 (57.9)	35 (56.5)
Non-small cell lung cancer	112 (42.4)	13 (43.3)	32 (42.1)	27 (43.5)
Stage IIIB-IV, n (%)	264 (100)	30 (100)	76 (100)	62 (100)
PD-L1 TPS≥1%, n (%)	148 (56.1)	17 (56.7)	42 (55.3)	35 (56.5)
Objective relief, n (%)	105 (39.8)	12 (40.0)	31 (40.8)	24 (38.7)

Intelligent & Human Futures
WISDOM ACADEMIC https://doi.org/10.63808/ihf.v2i3.510 Volume 2 , Issue 3

4.2. SpaImmune consistently predicts immune therapy responses with high accuracy across multiple independent cohorts

In the discovery cohort's five-fold stratified cross-validation, SpaImmune predicted objective response with an area under the receiver operating characteristic curve (AUC) of 0.96 (95% CI 0.94–0.98) (Fig. 1A). When the model with locked parameters was applied to the internally held-out test set (n=30), the AUC was 0.93 (95% CI 0.85–0.98), with a sensitivity of 91.7%, specificity of 83.3%, and a positive predictive value of 84.6% (Table 2). More importantly, the model continued to perform well in two independent external validation cohorts where it wasn't trained at all: external validation 1 AUC was 0.92 (0.86–0.97), and external validation 2 AUC was 0.91 (0.84–0.96) (Fig. 1B). Subgroup analysis across cancer types showed that SpaImmune had comparable predictive accuracy in melanoma (AUC 0.94, 95% CI 0.89–0.98) and non-small cell lung cancer (AUC 0.91, 0.85–0.96) (DeLong P=0.34), indicating good generalizability.

Figure 1

A: ROC curve of five-fold cross-validation in the discovery cohort

B: ROC curves of three independent validation cohorts

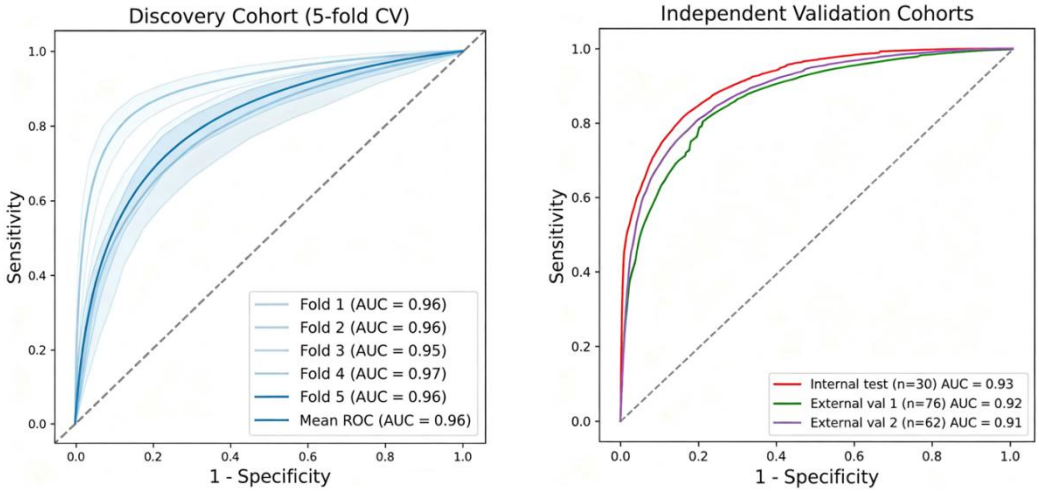


Table 2

Comparison of prediction performance of different models in an independent validation cohort

Model	Internal test set (n=30) AUC (95% CI)	External validation 1 (n=76) AUC (95% CI)	External validation 2 (n=62) AUC (95% CI)
PD-L1 TPS	0.61 (0.46–0.76)	0.63 (0.52–0.74)	0.60 (0.48–0.72)
Immune density PD-L1 logistic regression	0.69 (0.55–0.83)	0.71 (0.61–0.81)	0.68 (0.56–0.80)

Spatial Statistics SVM	0.75 (0.62–0.88)	0.78 (0.69–0.87)	0.74 (0.63–0.85)
Graph Attention Network (GAT)	0.80 (0.67–0.91)	0.82 (0.73–0.90)	0.79 (0.68–0.89)
Visual Transformer (ViT)	0.77 (0.63–0.89)	0.79 (0.70–0.88)	0.76 (0.65–0.87)
SpaImmune	0.93 (0.85–0.98)	0.92 (0.86–0.97)	0.91 (0.84–0.96)

Compared with existing biomarkers and spatial analysis methods, SpaImmune shows a significant advantage across all validation cohorts. Take External Validation 1 as an example: the AUC of clinical PD-L1 TPS alone is only 0.63; a logistic regression model based on the densities of 10 immune cells plus PD-L1 has an AUC of 0.71; a support vector machine based on predefined spatial statistics (clustering index, neighborhood diversity, etc.) has an AUC of 0.78; using a graph attention network alone (only mIHC protein phenotypes) and a vision transformer alone (only H&E images) yields AUCs of 0.82 and 0.79, respectively. SpaImmune improves the AUC over these models by 10%–30% (DeLong test, all $P<0.01$). A similar trend is observed in External Validation 2 (Table 2). These results suggest that integrating multimodal spatial omics information with a graph-transformer joint architecture is key for achieving high-precision prediction of immunotherapy response.

4.3. Interpretability analysis reveals that the close spatial pairing of B cells and Tfh cells within tertiary lymphoid structures is the key driver of ICIs response

Using SpaImmune's built-in SHAP global attribution and attention mechanism, we systematically decoded the spatial features driving the 'immunotherapy response' decisions. SHAP analysis showed that among the top 10 features contributing most to the model's output, the 'average attention edge weight between B cells and $CD4^+$ follicular helper T cells (Tfh)' ranked first (average |SHAP| value 0.183), followed by T cell density within the B cell neighborhood, spatial colocalization strength between $CD8^+$ T cells and dendritic cells, and so on (Figure 2A). This suggests that what the model learns from these massive spatial relationships isn't just the level of immune infiltration, but the interaction strength of specific immune subpopulations at the micrometer scale.

Further extracting all the cell interaction pairs ranked in the top 1% of edge attention among samples predicted as 'responders' by SpaImmune, we found that B–Tfh heterotypic interaction pairs accounted for as high as 34.7%, whereas in samples predicted as 'non-responders,' this proportion was only 12.1% ($P<0.001$, Figure 2B). This difference is not driven by cell abundance: there was no significant difference between the responder and non-responder groups in the density of B cells ($P=0.21$) or Tfh cells ($P=0.17$) within the tumor. Mapping the spatial positions of high-attention B–Tfh interaction pairs back to the original

Intelligent & Human Futures
WISDOM ACADEMIC https://doi.org/10.63808/ihf.v2i3.510 Volume 2 , Issue 3

tissue images, they were almost all located in histologically recognizable tertiary lymphoid structures (TLSs), specifically within $CD21^+$ follicular dendritic cell networks, with $BCL6^+ICOS^+$ Tfh cells closely adjacent to $CD20^+$ B cells (Figures 3A–C). Quantitative analysis showed that even within TLS regions, the median proportion of B–Tfh cell pairs with spatial distances ≤15 μm was 38.2% in the responder group, significantly higher than 11.5% in the non-responder group (P=0.0002). Univariate predictive analysis found that using just this feature of tight B–Tfh co-localization, the AUC for predicting ICI response could reach 0.86, whereas the traditional TLS maturity-based scoring (presence of $CD21^+$ networks and germinal center markers) was 0.74 (Figure 3D). Therefore, SpaImmune, without any prior immunological knowledge, autonomously captured the most immunologically meaningful functional spatial grammar and quantified it into an interpretable predictive rule.

Figure 2

A: SHAP Feature Importance Bar Chart

B: Comparison of the proportion of B–Tfh atypical interaction pairs

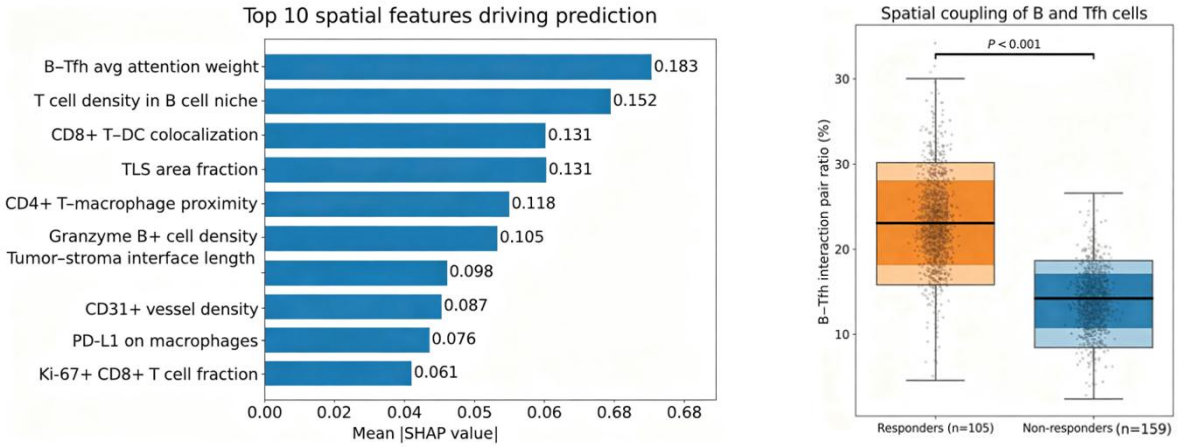
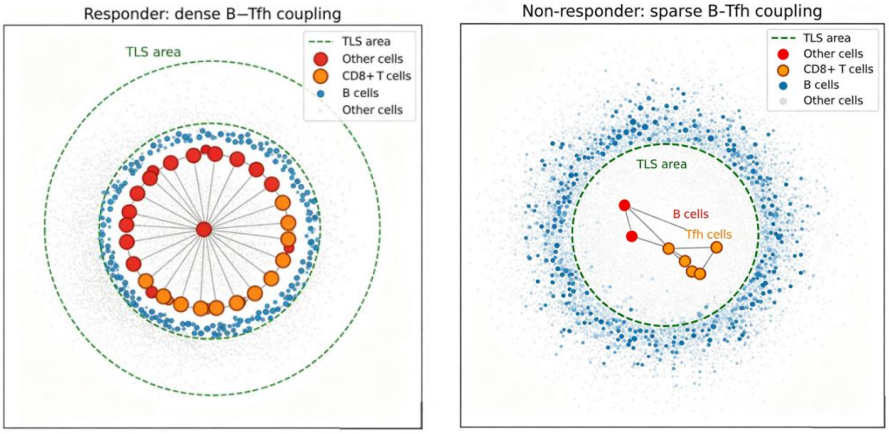


Figure 3

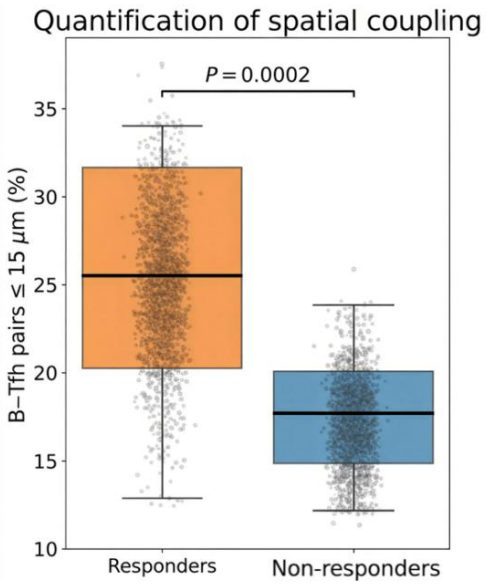
A: Spatial distribution of responders (high-density B–Tfh colocalization in TLS regions)

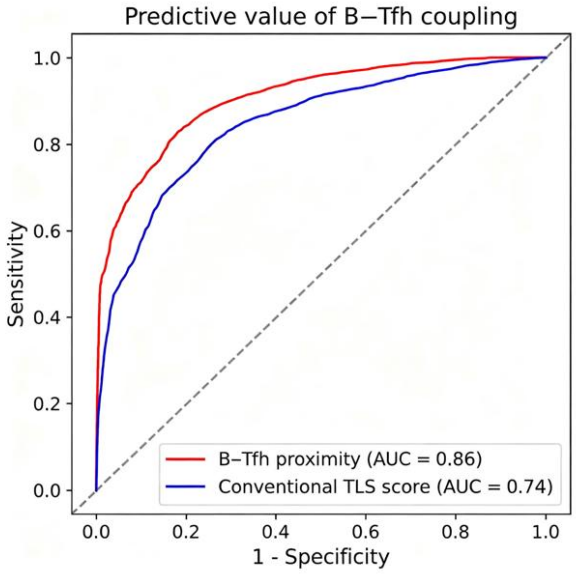
B: Spatial distribution of non-responders (loose B–Tfh colocalization)



C: Box plot of the close-range ratio of B–Tfh

D: Single-feature ROC comparison (B–Tfh ratio vs traditional TLS score)





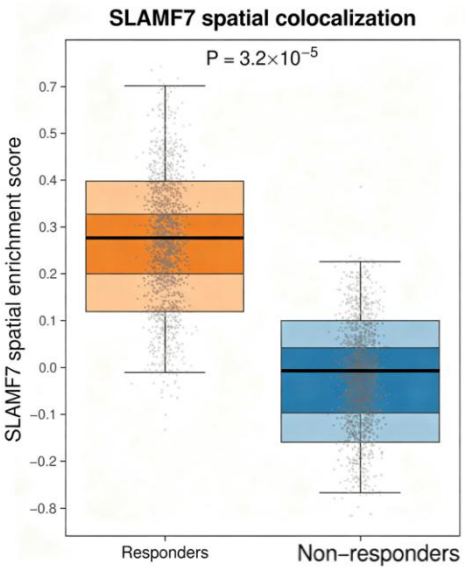
4.4. Cell-gene spatial colocalization network guides the discovery of SLAMF7 as a novel immunotherapy target

To turn the spatial knowledge that the model learned into actionable molecular targets, we used the cell-gene spatial colocalization network built with SpaImmune to perform an unbiased scan of the whole transcriptome. For each gene, we calculated its local spatial autocorrelation score around each cell type (S_{g,c}) and identified gene-cell spatial enrichment pairs that were highly positively correlated with remission (Spearman $\rho > 0.5$, adjusted $P < 0.01$). Across all analyses, SLAMF7 stood out the most: it showed very strong spatial clustering in both $CD8^+$ T cell and B cell regions, and its enrichment scores in the response group (median 0.47, IQR 0.21–0.69) were significantly higher than in the non-response group (median -0.08, IQR -0.31–0.14, $P=3.2\times10^{-5}$, Figure 4A).Spatial mapping visualization confirmed that high SLAMF7 expression in responding samples is highly restricted to TLS and areas densely infiltrated by $CD8^+$ T cells, whereas non-responding samples show diffuse low expression or no spatial clustering (Figure 4B). Using TCGA melanoma RNA-seq data for external validation, high SLAMF7 expression was significantly associated with longer overall survival (HR=0.52, 95% CI 0.33–0.81, P=0.004), and its expression strongly co-expressed with effector T cell and B cell signature gene sets (Figure 4C). Table 3 summarizes the top 5 spatially enriched gene–cell pairs most associated with ICI response, with SLAMF7–$CD8^+$ T cells and SLAMF7–B cells ranking first and third, respectively.

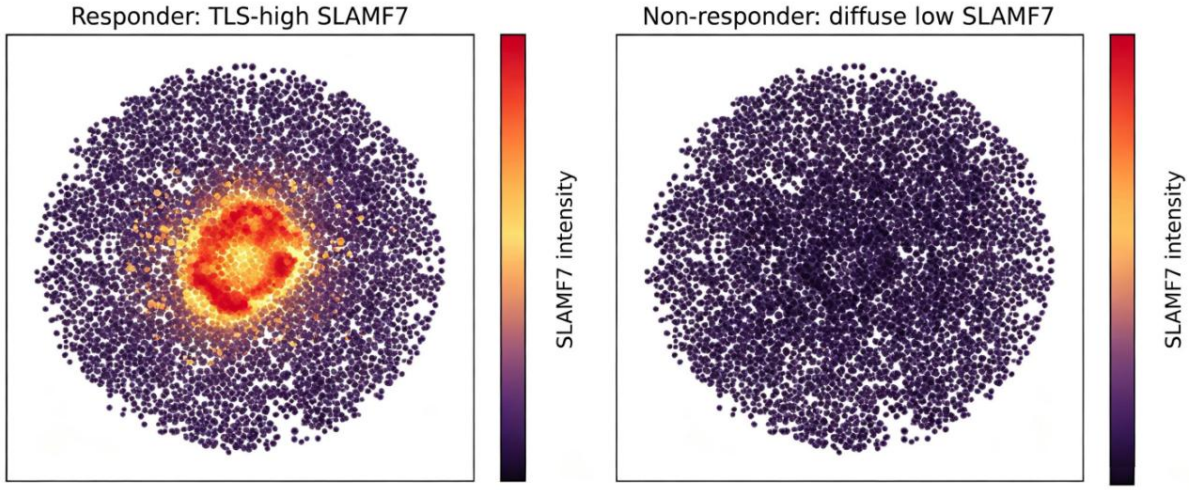
Intelligent & Human Futures
WISDOM ACADEMIC https://doi.org/10.63808/ihf.v2i3.510 Volume 2 , Issue 3

Figure 4

A: Boxplot of SLAMF7 spatial enrichment scores



B: Example of SLAMF7 spatial expression (responders vs non-responders)



C: Survival analysis (high and low SLAMF7 expression) and co-expression scatter plot

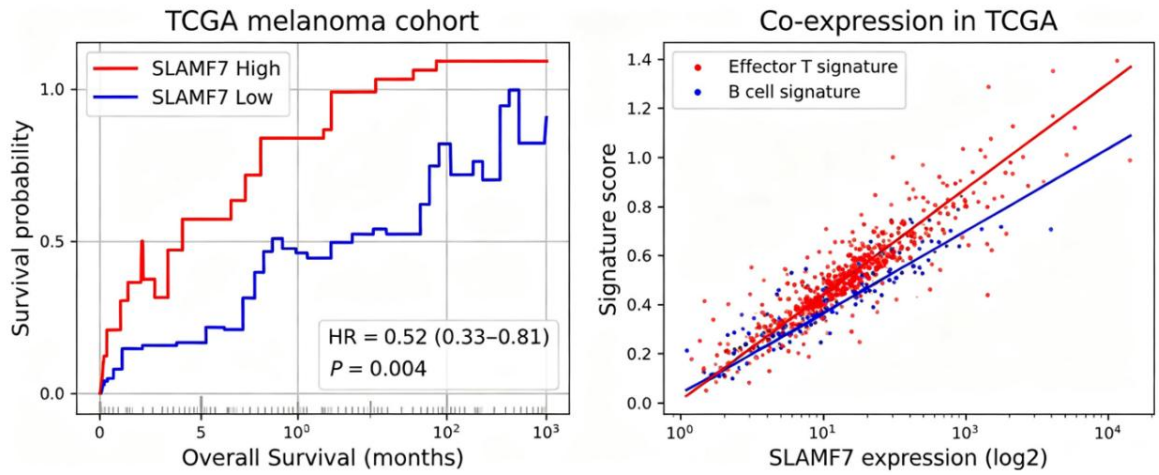


Table 3

Correlation between spatially enriched gene-cell pairs and ICI response (Top 5)

Gene	Cell type	Spearman ρ	Adjust P-value	Median S score in the response group	Median S score in non-responder group
SLAMF7	$CD8^+$ T cells	0.67	8.3×10^{-6}	0.47	-0.08
CXCL13	$CD4^+$ Tfh cells	0.62	2.5×10^{-5}	0.53	0.04
SLAMF7	B cells	0.61	3.1×10^{-5}	0.39	-0.12
CCL19	dendritic cell	0.58	1.8×10^{-4}	0.35	-0.05
GZMK	$CD8^+$ T cells	0.56	2.7×10^{-4}	0.41	0.02

4.5. Tumor killing mediated by $CD8^+$ T cells with enhanced SLAMF7 function

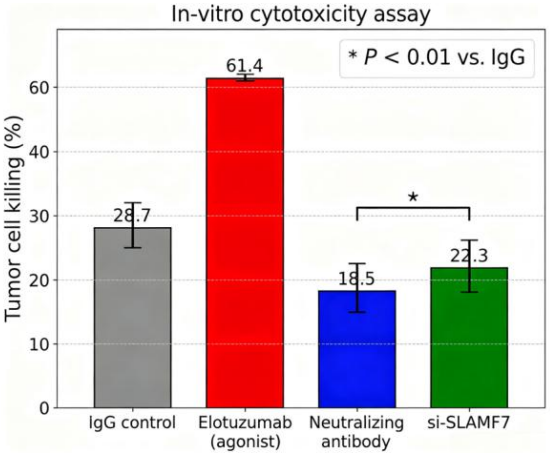
To check out the role of SLAMF7 in tumor immunity, we set up a co-culture system with human $CD8^+$ T cells and A375 melanoma cells. The results showed that adding the anti-SLAMF7 activating antibody elotuzumab (10 μg/mL) significantly boosted the tumor-killing rate of $CD8^+$ T cells from (28.7±5.1)% in the control IgG group to (61.4±6.8)% ($P<0.0001$, Figure 5A). At the same time, the levels of effector molecules IFN-γ and granzyme B in the supernatant went up from (345±72) pg/mL and (187±34) pg/mL to (1123±156) pg/mL and (534±68) pg/mL, respectively (both $P<0.001$, Figures 5B and 5C).On the other hand, when tumor cells were treated with SLAMF7-neutralizing antibodies or SLAMF7 expression was silenced via siRNA, the killing efficiency of $CD8^+$ T cells dropped to (18.5±4.2)% and (22.3±3.9)%, respectively, which was significantly lower than the control group ($P<0.01$) (Table 4). Silencing SLAMF7 in tumor cells didn't directly affect their own proliferation (Supplementary Fig. S4), suggesting that the reduced killing was due to impaired immune synapse function. This indicates that SLAMF7 on tumor cell surfaces may provide co-stimulatory signals through homophilic or heterophilic interactions with receptors on T cells, thereby boosting anti-tumor immunity. For specific quantitative data, see Table 4.

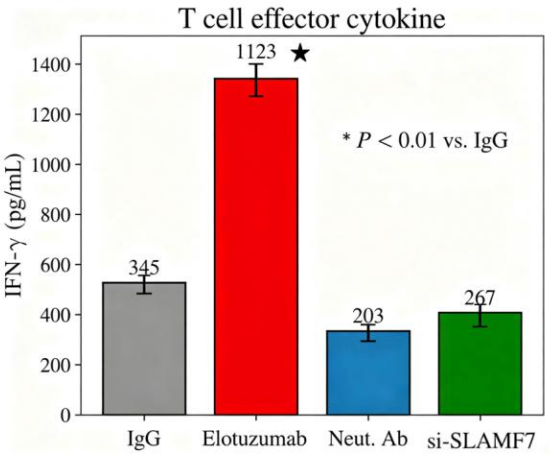
Figure 5

A: Bar chart of kill rate

B: Bar chart of IFN-γ concentration

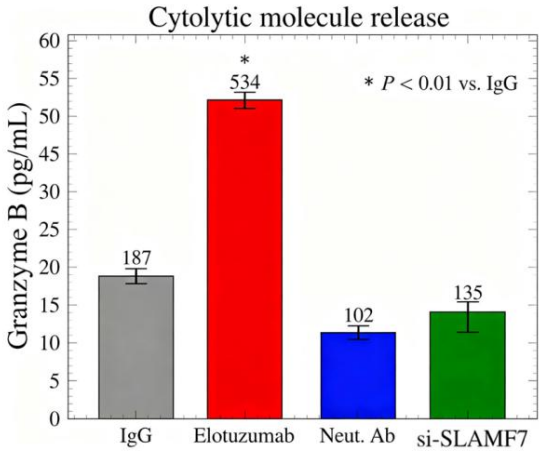
Intelligent & Human Futures
WISDOM ACADEMIC https://doi.org/10.63808/ihf.v2i3.510 Volume 2 , Issue 3





C: Granzyme B Concentration Bar Chart

S4: Proliferation curves of A375 melanoma cells after treatment with si-SLAMF7 or si-control



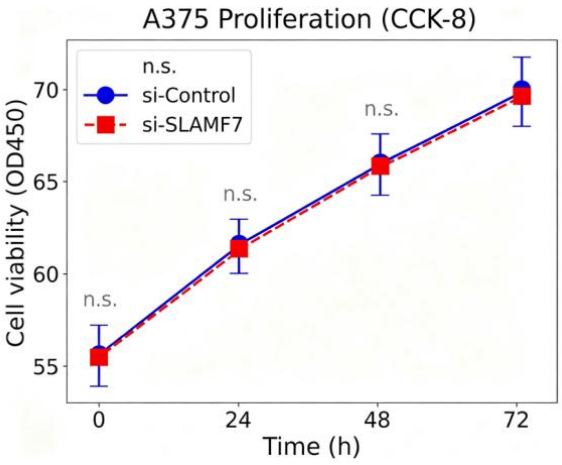


Table 4

The effect of SLAMF7 on CD8⁺ T cell anti-tumor killing function

Experimental conditions	Kill Rate (%)	IFN-γ (pg/mL)	Granzyme B (pg/mL)
IgG isotype control	28.7 ± 5.1	345 ± 72	187 ± 34
elotuzumab (agonist)	61.4 ± 6.8	1123 ± 156	534 ± 68
Neutralizing antibody (clone 162.1)	18.5 ± 4.2	203 ± 41	102 ± 29
si-SLAMF7 (tumor cells)	22.3 ± 3.9	267 ± 55	135 ± 33
si-control	30.1 ± 4.8	368 ± 64	195 ± 41

Note: Compared with the same-type control/si-control group, P<0.01; data are presented as mean ± standard deviation, with independent repeated experiments n=4.

In short, SpaImmune not only predicts clinical benefits from immunotherapy with higher accuracy than existing methods, but also reveals the core rule of B–Tfh spatial functional coupling in TLS through a fully transparent decision process. This unbiasedly led to the

discovery of SLAMF7 as a new target to boost solid tumor immunotherapy, completing a full translational loop from spatial omics to functional validation.

5. Discussion

This study is the first to propose and validate an explainable spatial AI framework called SpaImmune, achieving three core goals in one go: predicting the effectiveness of immune checkpoint inhibitors (ICIs) with higher accuracy than existing methods; transparently analyzing the key spatial patterns driving treatment responses; and directly discovering and preliminarily validating a new solid tumor immunotherapy target, SLAMF7, from the learned spatial interaction networks without bias. Through rigorous internal and external validation across multiple independent cohorts, SpaImmune has elevated the clinical translation value of spatial omics data to a new integrated level of prediction, explanation, and discovery.

In the field of immunotherapy prediction, traditional biomarkers like PD-L1 expression and tumor mutation burden (TMB) are still commonly used in clinics, but their predictive power is limited and they completely ignore spatial information. In recent years, spatial omics-driven prediction methods often rely on predefined statistics, such as cell neighborhood composition, Ripley's K function, or spatial hotspot scoring. While these methods have initially incorporated spatial information into models, they can only capture low-dimensional features designed by humans and cannot learn the hidden high-order topological structures within the data. End-to-end spatial learning methods represented by graph neural networks (GNNs), such as SpaGCN and subsequent spatial prognosis models, have broken the limitations of feature engineering. However, they mostly focus on region segmentation or prognostic stratification and are not specifically designed for predicting responses to immunotherapy, which is a highly dynamic and context-dependent clinical endpoint. In contrast, SpaImmune builds heterogeneous graphs with hundreds of thousands of cell nodes and uses a combination of graph attention networks (GATs) and vision Transformers to achieve cross-scale representation learning, from nanoscale cell neighborhoods to macroscopic tissue structures. This multimodal fusion strategy has proven to be decisive: the standalone GAT model using only protein phenotypes has an AUC of just 0.82, the vision Transformer model using only H&E images has an AUC of 0.79, while the integrated SpaImmune achieves an AUC over 0.91 across all validation sets, boosting performance by 10%–30%. This suggests that the spatial logic of tumor-immune interactions is encoded through the synergy of molecular signals and tissue morphology, and relying on a single modality inevitably loses some crucial information. Notably, the model achieved comparable predictive accuracy in both melanoma and non-small cell lung cancer, two tumors from

different tissue origins, implying that the spatial grammar captured might reflect general organizational principles of immune response, which supports its application across different cancer types.

The most transformative feature of SpaImmune lies in the complete interpretability of its decision-making process. By using SHAP global attribution and attention weight analysis, we can glimpse the spatial rule that the model assigns the highest weight to 'high-intensity spatial coupling of B cells and Tfh cells' which astonishingly echoes multiple landmark clinical studies published independently over the last three years. Teams like Helmink, Cabrita, and Petitprez have confirmed that in tertiary lymphoid structures (TLS), the cooperation between B cell lineages and $CD4^+$ follicular helper T cells is the strongest histological feature predicting immunotherapy benefits in melanoma, sarcoma, and various other tumors. However, these findings all stem from hypothesis-driven analyses based on TLS morphology as prior knowledge, and the standards for identifying and quantifying TLS vary. What makes SpaImmune unique is that during training, the model was only fed cell type, spatial coordinates, and transcriptomes it was never explicitly taught the concept of TLS, yet it autonomously recognized the close physical proximity of B–Tfh cell pairs (≤15 μm) as the most discriminative predictive feature. This 'unsupervised biological discovery' shows that AI can not only replicate existing human knowledge but also objectively validate and quantify known immunology concepts in a data-driven way. Furthermore, we found that even within TLS regions alone, the proportion of B–Tfh close interactions still outperforms traditional TLS maturity scoring (AUC 0.86 vs. 0.74), highlighting that the functional interaction strength surpasses mere structural presence, pointing the way for the refined development of TLS-related biomarkers.

Spatial omics research often stops at descriptive discoveries, but one of the most important breakthroughs in this study is that, by using a model to learn the whole transcriptome spatial colocalization network, we achieved a direct leap from observation to intervention. SLAMF7 (Signaling Lymphocytic Activation Molecule Family Member 7), a validated therapeutic antibody target in multiple myeloma, had its spatial expression patterns and functions in the solid tumor immune microenvironment almost completely unknown. Through unbiased spatial enrichment scanning, we for the first time revealed that SLAMF7 expression in tumors responding to ICIs is highly restricted to TLS and $CD8^+$ T cell-rich regions, whereas in non-responding tumors it shows diffuse low expression. This spatially polarized distribution suggests it may participate in local immune synapse strengthening among B–T–tumor cell interactions. Subsequent in vitro functional experiments confirmed that activating SLAMF7 signaling can significantly enhance $CD8^+$ T cell granzyme B and

IFN-γ release, boosting tumor killing from around 29% to 61%, while blocking it weakens the killing effect, indicating SLAMF7 has a co-stimulatory molecule-like function in solid tumors. This not only expands SLAMF7's immunotherapy potential to solid tumors, but also provides a complete demonstrative loop of "interpretable AI→target hypothesis→functional validation." Traditional target discovery heavily relies on functional screening or genomic perturbation, which is time-consuming and biased; our strategy transforms massive spatial omics data into a rankable candidate target list, greatly accelerating the hypothesis generation process in translational research.

This study also has several limitations. First, it is a retrospective study. Although multiple independent external cohorts were set up and the results were robust, the clinical utility of SpaImmune still needs to be confirmed by prospective studies or clinical trial data. Second, the total sample size and the diversity of cancer types are limited, so the model's performance in more immunotherapy scenarios (such as dual immunotherapy combination or ICI plus chemotherapy) and additional tumor types (like microsatellite instability-high tumors) cannot yet be evaluated. While spatial omics data registration and batch effect correction are standard procedures, they may still introduce systematic bias. SHAP interpretation is intuitive, but its inference of marginal contributions based on training data distribution cannot be directly equated with causal relationships, so the spatial rules revealed by the model (such as B–Tfh distance thresholds) need to be independently validated through high-resolution spatial functional experiments (e.g: spatial CRISPR perturbation). In addition, functional validation of SLAMF7 is currently only based on in vitro killing experiments, lacking in vivo animal model data, and its translational value as a therapeutic target still needs to be confirmed with humanized mouse models and more comprehensive safety evaluations. Finally, the current model focuses on the spatial structure of pre-treatment biopsy tissue and does not include dynamic changes; in the future, incorporating spatial features from early on-treatment re-biopsies into longitudinal integrated learning may help capture adaptive immune remodeling signals, further enhancing the predictive window.

6. Conclusion

This study developed and validated SpaImmune-a flexible spatial AI framework that can learn directly from multimodal spatial omics data without relying on prior knowledge. It achieves high-accuracy predictions of immune checkpoint inhibitor responses, transparent analysis of key spatial immune rules, and rational discovery of novel immunotherapy targets. Across multiple independent validation cohorts covering melanoma and non-small cell lung cancer, SpaImmune's prediction performance (AUC 0.91–0.93) far outperformed PD-L1

expression, immune cell density, and traditional spatial statistical models, highlighting the strong ability of cross-scale spatial interaction networks to discriminate immunotherapy responses.

For the first time, interpretability analysis unsupervisedly extracted the core rule of the tight spatial coupling between B cells and follicular helper T cells within the tertiary lymphoid structure from the model's decision logic. This not only highly aligns with cutting-edge clinical evidence but also quantifies it into a single predictive metric (AUC 0.86), highlighting the predictive value of functional spatial architecture beyond mere structural existence. More importantly, based on the model-constructed cell-gene spatial co-localization network, we discovered and preliminarily validated SLAMF7 in the tumor immune microenvironment as a functional molecule that hadn't been considered an immunotherapy target before. Its activation can significantly enhance the tumor-killing function of $CD8^+$ T cells, completing a full translational loop from spatial omics data to actionable targets.

Future research should expand in the following directions: SpaImmune needs clinical-grade validation in prospective clinical trials to assess its effectiveness in real-time support for immunotherapy decisions and its generalizability across different cancer types; causal experimental strategies, like spatial CRISPR perturbations, could independently verify the spatial interaction rules revealed by the model (such as B–Tfh distance thresholds) at a mechanistic level; the in vivo immunotherapy-sensitizing effect of SLAMF7 and its synergy with other ICIs still need comprehensive evaluation in humanized animal models; lastly, extending this framework to longitudinal biopsy samples collected during treatment could capture the spatial dynamics of adaptive immune remodeling, enabling earlier monitoring of treatment response and understanding resistance mechanisms. As spatial omics technology gradually becomes clinically practical and AI interpretability methods continue to mature, the "predict–explain–discover" integrated intelligent approach represented by SpaImmune could offer a new paradigm for driving spatially-informed personalized immunotherapy.

Acknowledgments

Not applicable.

Author contributions

Conceptualization, Gang Liu; methodology, Jia Wang; validation, Jia Zhu; formal analysis, Gang Liu and Jia Wang; investigation, Jia Wang and Jia Zhu; resources, Gang Liu; data curation, Jia Zhu; writing—original draft preparation, Jia Wang; writing—review and editing, Gang Liu; visualization, Jia Zhu; supervision, Gang Liu; project administration, Jia Wang. All authors have read and agreed to the published version of the manuscript.

Funding

Not applicable.

Data availability

Not applicable.

Declarations

Conflict of interest

The authors declare that there is no conflict of interest.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

All authors have given their consent.

Additional information

Received: 12 August 2026

Accepted: 28 August 2026

Published online: 1 September 2026

Publisher's Note

References

[1] Cabrita, R., Lauss, M., Sanna, A., et al. (2020) Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature, 577*(7791), 561-565. https://doi.org/10.1038/s41586-019-1914-8

[2] Campanella, G., Hanna, M. G., Geneslaw, L., et al. (2019) Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med, 25*(8), 1301-1309. https://doi.org/10.1038/s41591-019-0508-1

[3] Chen, K. H., Boettiger, A. N., Moffitt, J. R., et al. (2015) Spatially resolved, highly multiplexed RNA profiling in single cells. *Science, 348*(6233), aaa6090. https://doi.org/10.1126/science.aaa6090

[4] Courtiol, P., Maussion, C., Moarii, M., et al. (2019) Deep learning-based classification of mesothelioma improves prediction of patient outcome. *Nat Med, 25*(10), 1519-1525. https://doi.org/10.1038/s41591-019-0583-3

[5] Fridman, W. H., Pagès, F., Sautès-Fridman, C., Galon, J. (2012) The immune contexture in human tumours: Impact on clinical outcome. *Nat Rev Cancer, 12*(4), 298-306. https://doi.org/10.1038/nrc3245

[6] Galon, J., Costes, A., Sanchez-Cabo, F., et al. (2006) Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science, 313*(5795), 1960-1964. https://doi.org/10.1126/science.1129139

[7] Goltsev, Y., Samusik, N., Kennedy-Darling, J., et al. (2018) Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. *Cell, 174*(4), 968-981.e15. https://doi.org/10.1016/j.cell.2018.07.010

[8] Havel, J. J., Chowell, D. & Chan, T. A. (2019) The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat Rev Cancer, 19*(3), 133-150. https://doi.org/10.1038/s41568-019-0116-x

[9] Helmink, B. A., Reddy, S. M., Gao, J., et al. (2020) B cells and tertiary lymphoid structures promote immunotherapy response. *Nature, 577*(7791), 549-555. https://doi.org/10.1038/s41586-019-1922-8

[10] His, E. D., Steinle, R., Balasa, B., et al. (2008) CS1, a potential new therapeutic antibody target for the treatment of multiple myeloma. *Clin Cancer Res, 14*(9), 2775-2784. https://doi.org/10.1158/1078-0432.CCR-07-4246

[11] Hu, J., Li, X., Coleman, K., et al. (2021) SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat Methods, 18*(11), 1342-1351. https://doi.org/10.1038/s41592-021-01255-8

[12] Hunter, M. V., Moncada, R., Weiss, J. M., et al. (2021) Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface. *Nat Commun, 12*(1), 6278. https://doi.org/10.1038/s41467-021-26614-z

[13] Keren, L., Bosse, M., Marquez, D., et al. (2018) A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell, 174*(6), 1373-1387.e19. https://doi.org/10.1016/j.cell.2018.08.039

[14] Liu, G. (2026) The application of artificial intelligence in tumor immunotherapy: integrating multi-omics data to achieve efficacy prediction and combination strategies. *Applied Artificial Intelligence Research, 2*(2). https://doi.org/10.65455/q72yq328

[15] Lundberg, S. M. & Lee, S. I. (2017) A unified approach to interpreting model predictions. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, 4765–4774. https://doi.org/10.48550/arXiv.1705.07874

[16] Mariathasan, S., Turley, S. J., Nickles, D., et al. (2018) TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature, 554*(7693), 544-548. https://doi.org/10.1038/nature25501

[17] Nirmal, A. J., Maliga, Z., Vallius, T., et al. (2022) The spatial landscape of progression and immunoediting in primary melanoma at single-cell resolution. *Cancer Discov, 12*(6), 1518-1541. https://doi.org/10.1158/2159-8290.CD-21-1357

[18] Petitprez, F., de Reyniès, A., Keung, E. Z., et al. (2020) B cells are associated with survival and immunotherapy response in sarcoma. *Nature, 577*(7791), 556-560. https://doi.org/10.1038/s41586-019-1906-8

[19] Schürch, C. M., Bhate, S. S., Barlow, G. L., et al. (2020) Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *Cell, 182*(5), 1341-1359.e19. https://doi.org/10.1016/j.cell.2020.07.005

[20] Selvaraju, R. R., Cogswell, M., Das, A., et al. (2017) Grad-CAM: Visual explanations from deep networks via gradient-based localization. *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 618-626. https://doi.org/10.1109/ICCV.2017.74

[21] Ståhl, P. L., Salmén, F., Vickovic, S., et al. (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science, 353*(6294), 78-82. https://doi.org/10.1126/science.aaf2403

[22] Topalian, S. L., Hodi, F. S., Brahmer, J. R., et al. (2012) Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med, 366*(26), 2443-2454. https://doi.org/10.1056/NEJMoa1200690

[23] Wu, S. Z., Al-Eryani, G., Roden, D. L., et al. (2021) A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet, 53*(9), 1334-1347. https://doi.org/10.1038/s41588-021-00911-1

[24] Zhao, S., Liang, X., Xie, J., et al. (2025) Integrating pathology genomics and single-cell genomics to identify lactate metabolism-related prognostic features and therapeutic strategies for melanoma. *Apoptosis, 30*(11-12), 2963-2988. https://doi.org/10.1007/s10495-025-02190-1