Vision-Based Deep Learning of Biomedical Tabular Data via Self-Supervised Cartographic Representation

Sakib Mostafa[1], Yuming Jiang[2], James Zou[3,4,5], Tarik F. Massoud[6,7], Ash A. Alizadeh[9], Maximilian Diehn[1,7,8], Lei Xing[1,10,11*], and Md Tauhidul Islam[1,*]

[1]Department of Radiation Oncology, Stanford University, Stanford, California, USA
[2]Department of Radiation Oncology, Wake Forest University School of Medicine, Winston Salem, North Carolina, USA
[3]Department of Biomedical Data Science, Stanford University, Stanford, California, USA
[4]Department of Computer Science, Stanford University, Stanford, California, USA
[5]Department of Electrical Engineering, Stanford University, Stanford, California, USA
[6]Department of Radiology, Stanford University, Stanford, California, USA
[7]Department of Medicine, Division of Oncology, Stanford University, Stanford, CA, USA
[8]Stanford Cancer Institute, Stanford University, Stanford, California, USA
[9]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, California, USA
[10]Institute of Computational and Mathematical Engineering, Stanford University, Stanford, CA, USA
[11]Department of Electrical Engineering, Stanford University, Stanford, CA, USA
[*]Correspondence: tauhid@stanford.edu and lei@stanford.edu

ABSTRACT

Tabular data constitute a dominant representation in biomedical science. Unlike images, texts, or time series, tabular representation of a dataset generally lacks structural organization: features are treated as unordered dimensions, and their interrelationships must be inferred implicitly by learning algorithms. This fundamental structural limitation constrains the ability of convolutional neural networks, vision transformers, and other structure-aware vision architectures to exploit local correlations and higher-order interactions that encode underlying biological and clinical signals. Here we introduce Dynamic Feature Mapping (Dynomap), an end-to-end deep learning cartography framework that learns a task-optimized spatial topology of features directly from data. Dynomap discovers inter-feature dependencies and jointly optimizes their spatial arrangement with a predictive objective through a differentiable rendering mechanism, without reliance on heuristics, predefined feature groupings, or external priors. By transforming high-dimensional tabular vectors into learned feature maps, Dynomap enables vision-based deep learning architectures to operate effectively on non-spatial biomedical inputs and achieves substantially better predictive performance compared to existing state-of-the-art methods across clinical and biological datasets. When applied to a lung cancer liquid biopsy dataset from Stanford Hospital, Dynomap organizes genes from a curated cancer-associated panel into coherent spatial structures and improves multiclass cancer subtype prediction accuracy by up to 18% relative to classical and state-of-the-art deep learning tabular analysis methods. In a Parkinson's disease voice dataset, Dynomap spatially clusters disease-associated acoustic descriptors, including tunable Q-factor wavelet energy and entropy measures linked to vocal instability, and yields accuracy gains of up to 8% compared to state-of-the-art techniques. Similar performance improvements were observed when Dynomap was applied to 13 additional public tabular benchmarks spanning molecular and non-biomedical domains. By converting unordered feature spaces into learned spatial representations, Dynomap establishes a general and principled strategy for bridging tabular and vision-based deep learning and enables the discovery of structured, task-relevant patterns from unordered high-dimensional biomedical data.

Introduction

Tabular data serve as a foundational format for representing information across biology, medicine, and data-intensive scientific fields, spanning diverse modalities such as gene expression matrices, single-cell profiles, proteomic and metabolomic measurements, liquid biopsy features, electronic health record variables, imaging-derived radiomic and phenomic descriptors, wearable sensor summaries, and experimentally engineered features[1–10]. Across these domains, each sample is represented as a high-dimensional feature vector capturing complex biological, technical, and environmental characteristics of the system under study[1,2]. Although such representations are convenient, a fundamental limitation is their lack of intrinsic structural

organization[11,12]. Unlike images, texts, signals, or time series data, tabular inputs contain no explicit encoding of relationships among features, treating variables as unordered dimensions and leaving dependencies arising from coordinated molecular programs, shared regulatory pathways, and physiologic dependencies entirely implicit[11–13]. The learning algorithm must therefore recover all feature relationships from scratch, with no structural priors such as locality, compositionality, or spatial coherence to guide the search[14,15]. This substantially reduces model capabilities and leads to suboptimal performance in many practical applications[11,12]. The problem is particularly consequential in biomedical research, where data are often acquired in the presence of substantial technical noise, pronounced biological heterogeneity, and severe class imbalance[6,16,17]. Therefore, models must simultaneously denoise measurements, recover coordinated feature interactions, and generalize from limited samples, making both pattern discovery and interpretation substantially challenging[2,12].

Existing methods for analyzing tabular biomedical data span a wide range of modeling strategies, yet each carries fundamental limitations that constrain predictive performance and interpretability. At the simplest end, linear models such as logistic regression, linear discriminant analysis, and ridge regression combine features through a weighted sum and perform reasonably well when feature relationships are approximately linear, but they fail systematically in real biomedical datasets where interactions among variables are nonlinear, conditional, and context-dependent[18,19]. Tree-based ensemble methods such as Random Forests, AdaBoost, XGBoost, LightGBM, and gradient boosted trees address nonlinearity by recursively partitioning the feature space, and in practice they remain strong baselines for many tabular benchmarks[20–23]. However, they operate on raw features without constructing hierarchical representations, which limits their ability to capture the distributed, higher-order interactions characteristic of tabular biomedical data such as genomic and transcriptomic data. Distance and kernel-based approaches such as KNN and SVM sidestep feature modeling altogether by comparing samples directly in input space[24,25], which leaves the internal organization of features entirely unmodeled and provides no view into how predictive variables relate to one another. Neural network models including MLPs and modern tabular architectures such as TabM and ModernNCA can capture complex nonlinear interactions through learned embeddings, and recent benchmarks show they outperform tree-based methods on several biological tasks[11,12,26–28]. Yet the structure they learn remains entirely implicit within the network weights, making it difficult to interpret which feature relationships drive predictions or whether the learned organization is biologically meaningful. Foundation-style models such as TabPFN push performance further through meta-learned priors but still treat features as an unordered vector, with no mechanism to reveal how variables interact or co-organize[11,12,29]. Fixed vector-to-image methods such as DeepInsight[30], IGTD[31], NCTD[32], Genomap[33], and REFINED[34] take a different approach by mapping tabular features onto a spatial grid before training and applying convolutional models on top. While this allows convolutional models to operate on tabular inputs, the spatial layout is constructed once using predefined similarity measures and remains fixed throughout training. Because the layout does not adapt to the prediction objective, it cannot reorganize to reflect the specific biological interactions that matter for a given task[30–34]. This gap motivates a framework that discovers task-adaptive spatial organization directly from data, making the arrangement of predictive biological signals interpretable[35].

Deep learning architectures that use spatial relationships among features, such as convolutional neural networks (CNNs) and vision transformers (ViTs), have demonstrated strong capabilities in domains including natural images and structured signals[14,36,37]. These models leverage locality, hierarchical composition, and parameter sharing to model structured patterns of input data efficiently[14,38]. Such architectural properties are particularly effective when relationships among input dimensions are organized in space, allowing computation of local correlations and progressive abstraction of correlation signals through the layers of the network for reliable decision-making[39]. In these networks, spatial proximity enables convolutional filters or attention mechanisms to capture complex and higher-order interactions that would be difficult to capture in an unstructured vector representation[14,35]. Establishing a data-driven spatial arrangement of features is therefore essential for enabling spatial architectures to capture coordinated biological or phenotypic interactions in tabular data[30,31,34].

Here, to address the limitations of existing methods and develop a general strategy for transforming unordered biomedical tabular data into a spatially organized format suitable for vision deep learning models, we introduce Dynomap, a framework that learns a spatial organization of tabular features in an end-to-end manner (Fig. 1). Dynomap assigns each feature a learnable coordinate in a two-dimensional continuous space. Individual samples are rendered into images through a fully differentiable process, in which feature values modulate localized kernels placed at their learned positions. The feature layout and the predictive model are optimized jointly through backpropagation, allowing the spatial arrangement of features to evolve in response to the prediction objective. In this formulation, spatial neighborhoods emerge directly from the structure of the predictive signal rather than being imposed in advance. This design enables convolutional or vision transformer models to operate directly on tabular data while preserving feature identity and supporting direct visualization of how predictive biological signals are organized across features. By coupling spatial representation learning with prediction, Dynomap provides an adjustable layout transformation of unordered tabular data. The learned spatial organization is task-adaptive and can reorganize as objectives change, allowing coherent biological or phenotypic modules to emerge from the data. Moreover, this representation supports interpretation through direct visualization and attribution analysis, enabling examination of how predictive signal is distributed across features and spatial neighborhoods.

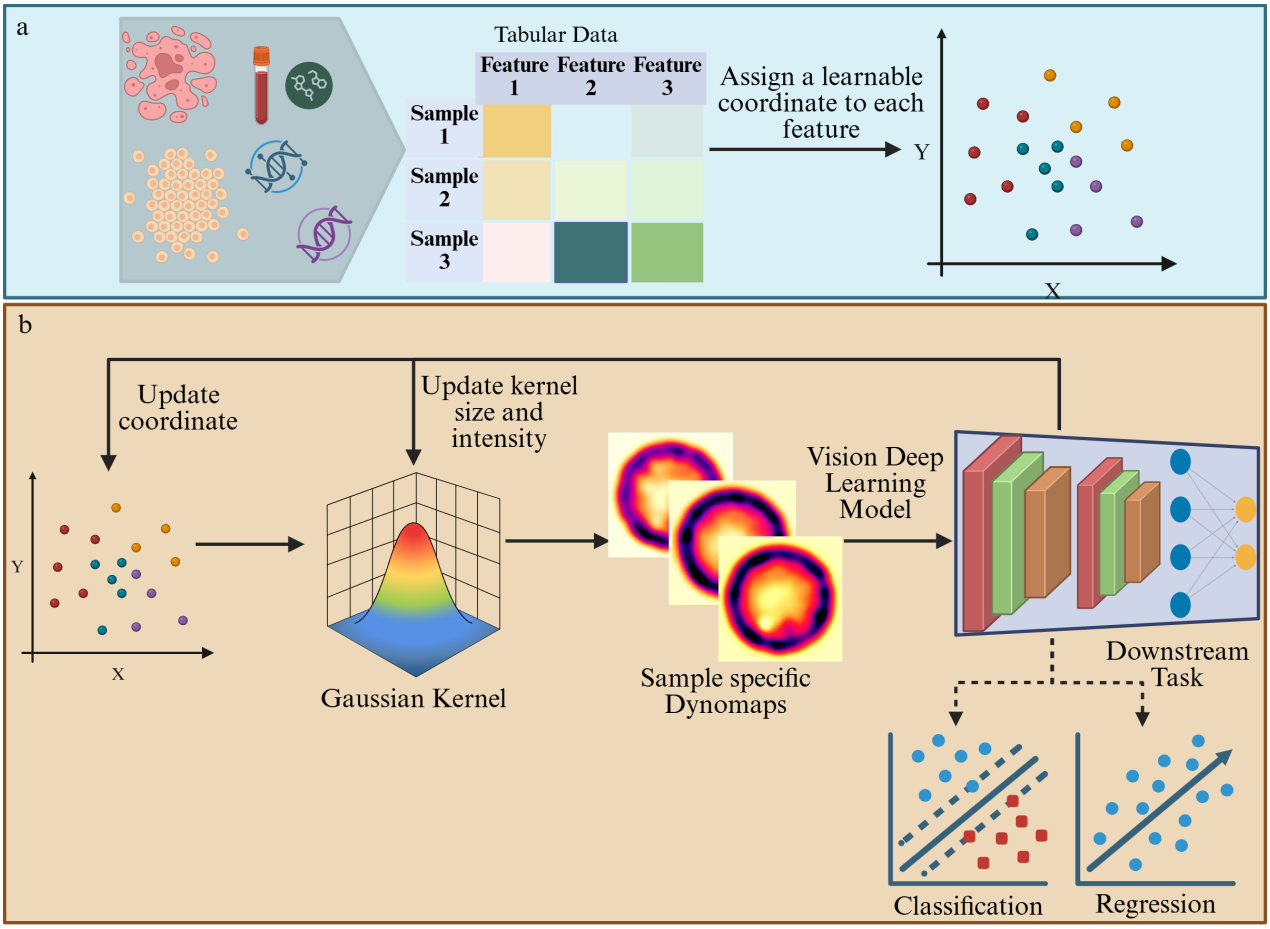

Figure 1. Overview of the Dynomap framework for spatial representation learning of tabular data. a, Biomedical measurements are represented as a feature matrix where each column is a biological variable such as a gene expression value, molecular measurement, or phenotypic descriptor, and each row is a sample. Because these features carry no inherent spatial relationships, Dynomap assigns each one a learnable coordinate in a continuous two-dimensional space, so that proximity in the image can come to reflect biological relatedness. **b**, During training, feature coordinates and rendering parameters are updated jointly with the predictive model. Each feature contributes to the image through a Gaussian kernel whose position and intensity are determined by the learned coordinates and the feature values of that sample, producing a continuous spatial map specific to each sample. A convolutional neural network processes these maps to perform downstream tasks including classification and regression. Because the spatial layout evolves in response to the prediction objective, features that carry related biological signal tend to migrate toward one another and form coherent spatial neighborhoods.

Below, we apply Dynomap across diverse biomedical datasets spanning liquid biopsy transcriptomics, bulk RNA sequencing, single-cell expression profiling, and engineered phenomic measurements. Across these datasets varying in dimensionality, class imbalance, and biological complexity, Dynomap reveals structured and task-adaptive spatial organization of features, enabling identification of coherent, class-dependent biological modules across molecular and phenomic domains. The learned representations exhibit spatial neighborhoods in which predictive variables cluster according to disease state, cellular lineage, or phenotypic condition. These spatial patterns are supported by attribution analysis and null-controlled spatial statistics, indicating that the learned topography reflects biologically meaningful signal. By jointly learning feature coordinates and prediction in an end-to-end framework, Dynomap provides an adaptive spatial representation of tabular data enabling vision models to operate effectively on traditionally non-spatial inputs while preserving interpretability and facilitating discovery of clinically relevant patterns.

Results

Detecting, subtyping, and staging cancer from ultrasensitive cfRNA liquid biopsy

We first applied Dynomap to a circulating cell-free RNA (cfRNA) liquid biopsy cohort profiled on RARE-Seq, an ultrasensitive platform designed to detect low-abundance tumor-derived transcripts in plasma[40]. Liquid biopsy offers the possibility of detecting, subtyping, and staging cancer from a single blood draw, and cfRNA has emerged as a particularly informative analyte because tumor-derived transcripts carry tissue-of-origin signal that genomic assays cannot easily recover[6,7]. In practice, however, the signal is extremely difficult to extract. Tumor-derived transcripts represent a small fraction of plasma cfRNA, measurements exhibit low signal-to-noise ratios that vary across patients, and cohort composition is typically imbalanced across cancer types and disease stages[6,17]. These properties make cfRNA one of the most demanding settings for tabular learning, because predictive models must simultaneously separate weak biological signal from technical noise, generalize across heterogeneous patients, and avoid collapsing onto dominant classes. We therefore treated RARE-Seq as the primary testbed for Dynomap and examined performance across three clinically motivated tasks, namely healthy versus cancer discrimination, cancer subtype classification, and stage prediction. To assess whether Dynomap's performance depends on the dimensionality of the input, we considered two feature regimes on the same cohort. In the whole-transcriptome setting, we retained the 4,000 highly variable genes drawn from the full catalogue of over 60,000 genes. In the targeted setting, we restricted inputs to the 622-gene signature panel defined in the original study[40]. This pairing allows a direct comparison between a broad transcriptomic regime and a compact, biologically curated feature space, and tests whether learned spatial organization adapts to the concentration of available signal.

Across all three tasks—healthy versus cancer discrimination, cancer subtype classification, and cancer stage classification—Dynomap consistently outperformed both classical machine learning methods and state-of-the-art deep tabular models. In the 4,000 HVG setting, Dynomap achieved 93.0% accuracy for binary classification, with a macro-F1 of 92.0% and macro-sensitivity of 91.1% (Fig. 2a). This exceeded classical baselines including Logistic Regression (90.3%), XGBoost (91.2%), and Random Forest (88.9%), as well as SOTA neural tabular models such as TabM (92.9%) and ModernNCA (90.7%)[11,20,22,29]. When compared to TabM, Dynomap achieved much higher macro-sensitivity, indicating more balanced performance across classes. These results demonstrate that Dynomap surpasses both ensemble-based learners, widely regarded as strong baselines for tabular data[19], and recently proposed deep tabular architectures.

The advantage of Dynomap became more notable in the multiclass subtype task. Using 4,000 HVGs, Dynomap achieved 80.7% accuracy with a macro-F1 of 79.3%, substantially outperforming Logistic Regression (73.5%), XGBoost (74.6%), and TabM (78.6%). The margin was even higher when compared to tree-based methods and MLPs. For state classification, Dynomap achieved 79.6% accuracy and a macro-F1 of 78.1%, again substantially exceeding classical and modern deep learning-based tabular approaches. For stage, Dynomap consistently maintained superior macro-sensitivity, indicating robustness under class imbalance. For the 622-gene signature panel in binary classification, Dynomap achieved 90.9% accuracy with a macro-F1 of 90.0%, remaining superior to existing models. Despite nearly a tenfold reduction in dimensionality, binary classification performance declined only modestly relative to the 4,000 HVG regime for Dynomap, indicating that the curated panel preserves the dominant discriminative signal.

For Subtype classification under the signature panel, Dynomap reached 92.0% accuracy, which represents a significant increase relative to the 80.7% accuracy obtained using 4,000 HVGs, suggesting that subtype-discriminative signal is concentrated within the curated gene set. Even under this constrained feature space, Dynomap maintained higher macro-F1 compared with classical and deep learning-based baselines. For stage prediction by using the 622-gene panel, Dynomap achieved 78.8% accuracy, similar to the 4,000 HVG regime, but with reduced macro-F1 and macro-sensitivity. This asymmetry indicates that subtype information is concentrated within curated markers, whereas stage-associated signal appears more distributed across the broader transcriptome. Across both feature regimes, Dynomap remained similar or superior to all baselines, including TabPFN and TabM[11,29], demonstrating that its performance does not rely on brute-force dimensionality but adapts to the concentration of available signal.

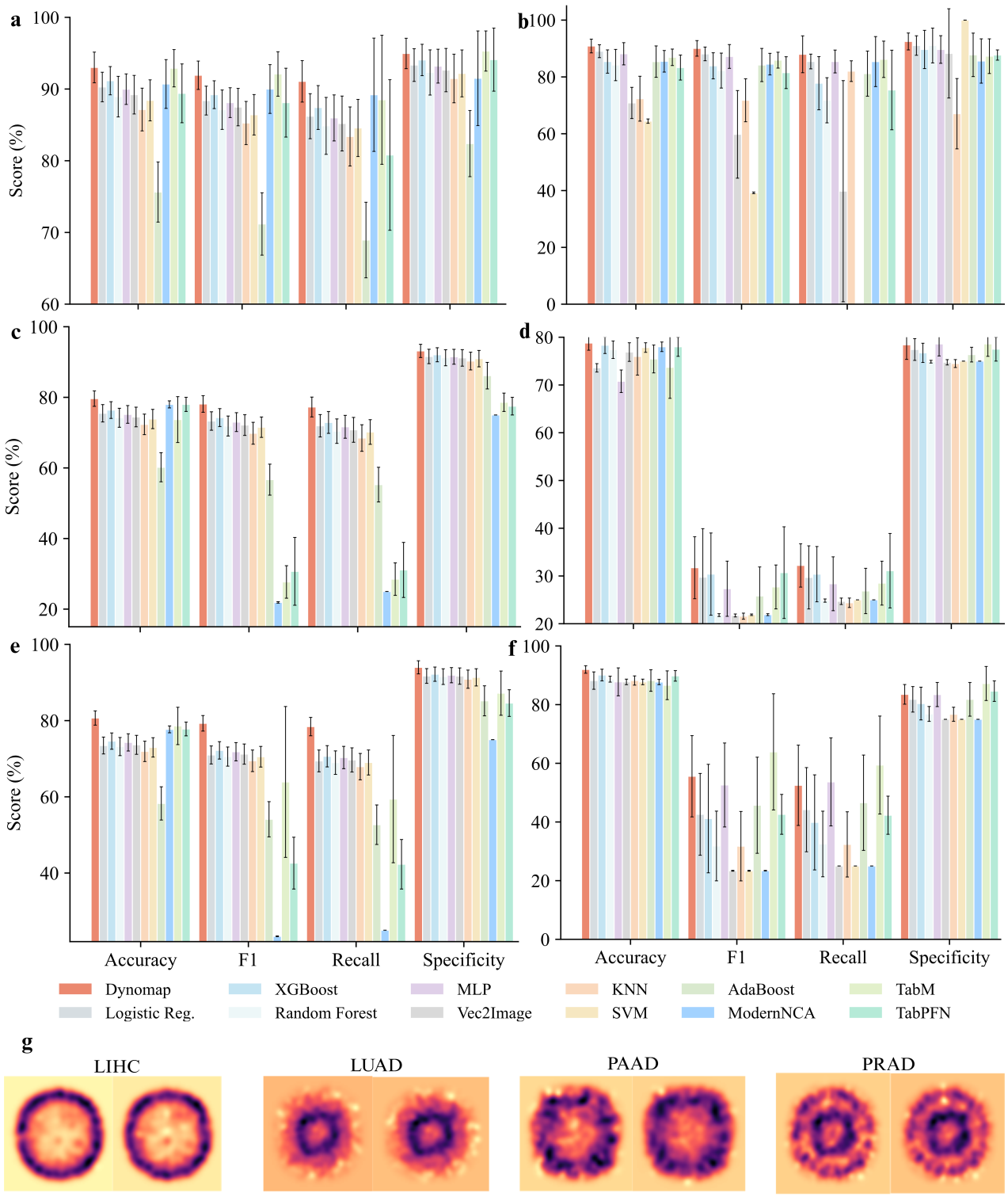

Figure 2. Dynomap performance and learned representations for liquid biopsy cfRNA classification. a–f, Dynomap is benchmarked against eleven baseline methods across three clinically relevant prediction tasks: binary healthy versus cancer classification (**a**, **b**), multiclass cancer subtype prediction (**c**, **d**), and cancer stage classification (**e**, **f**). Each task is evaluated under two feature regimes, with highly variable genes on the left and a curated gene signature panel on the right. Performance is reported as accuracy, macro-F1 score, recall, and specificity. Error bars indicate standard deviation across five cross-validation folds. Dynomap achieves the highest or jointly highest performance across most metrics in both feature settings. **g**, Representative spatial image representations produced by Dynomap for cancer subtype classification. Each group of four images corresponds to samples from a single cancer class. Spatial patterns are reproducible within each class, reflecting task-driven organization that emerges directly from the tabular gene expression input rather than from any predefined feature arrangement.

Beyond predictive accuracy, Dynomap exhibited systematic changes in spatial organization and attribution patterns between the 4,000 HVG and 622-gene regimes. Visualization of generated images revealed reproducible task-specific spatial structures (Fig. 2g). Spatial autocorrelation analysis using Moran's I and neighborhood purity (kNN consistency) showed that the observed spatial organization deviated from null layouts (Figs. S3, S4), indicating non-random clustering of the predictive signal.

To examine how predictive signals are distributed across genes, we computed Integrated Gradients (IG) and summarized normalized mean absolute attribution values for each task and feature regime (Fig. 3)[41]. Attribution patterns were analyzed at the class level to identify genes contributing disproportionately to specific predictions. In subtype classification under the 4,000 HVG setting, distinct tumor classes were associated with separable high-attribution gene sets (Fig. S13). For prostate cancer (PRAD), genes such as *NFATC3*, *RABEP2*, and *BIRC3* exhibited elevated attribution relative to other classes. Lung adenocarcinoma (LUAD) predictions were associated with higher contributions from *CRIPTO* and *CA9*, whereas pancreatic cancer (PAAD) showed stronger attribution for *STAT3* and *NF1*[42–47].

When restricting input to the 622-gene signature panel, the class-dependent attribution structure remained evident, but the dominant genes shifted, and the attribution mass was concentrated within a smaller subset of features. Although subtype separability improved under the signature regime, the specific high-contributing genes differed from those identified in the 4,000 HVG setting. In the binary healthy versus cancer task, butterfly plots revealed clear class-specific attribution asymmetry (Fig. S12). Cancer predictions showed elevated attribution for genes including *PFKM*, *RASGRP2*, and *SMAD4*, whereas healthy predictions were associated with higher attribution for *P2RY10*, *TP63*, and *NFE2L2*[44,48–50]. Stage prediction exhibited partially overlapping but stage-dependent attribution patterns, with genes such as *PAX6* and *MUC2* more prominent in early-stage samples and *BIRC3* and *CTLA4* contributing more strongly in later stages (Fig. S14)[42]. Under the 622-gene regime, stage-associated patterns persisted and used *CA12*, *PPY*, and *TMEFF2*.

Across all tasks, the magnitude and distribution of IG attributions reorganized systematically when transitioning from 4,000 HVGs to the 622-gene signature panel. The full-gene regime distributed predictive contributions across a broader gene set, whereas the curated panel concentrated attribution within fewer features. No single subset of genes dominated across all tasks or regimes. Instead, both the spatial layout and attribution hierarchy adapted to the prediction objective and the available gene space, indicating that Dynomap dynamically reallocates predictive weight according to task and dimensional context.

Taken together, these results show that Dynomap surpasses strong classical learners and modern deep tabular models on ultrasensitive cfRNA data, while exhibiting task-dependent and feature-regime-dependent representational adaptation. Performance gains were accompanied by systematic changes in spatial organization and attribution concentration, indicating that Dynomap learns structured representations that reflect both the predictive objective and the available molecular feature space.

Resolving tumor type from tumor-educated platelet transcriptomes

We next examined whether the performance gains observed on cfRNA dataset by Dynomap extend to a blood-based transcriptomic assay with a higher signal-to-noise profile. Tumor-educated platelets (TEPs) absorb tumor-derived RNA during circulation, and their transcriptional profiles have been proposed as a minimally invasive readout for cancer detection and tissue-of-origin inference[51]. Unlike plasma cfRNA, platelet RNA-seq yields substantially more abundant and stable transcripts, which shifts the analytical challenge from signal recovery toward discrimination among cancer types that share overlapping transcriptional programs. We used the GSE68086 cohort, which contains platelet RNA-seq profiles from patients with breast cancer, colorectal cancer, glioblastoma, lung cancer, and pancreatic cancer alongside healthy controls[52]. This setting tests whether Dynomap's learned spatial organization reflects stable class-specific transcriptional programs when the limiting factor is biological similarity among cancers rather than technical noise in the measurement.

Across both binary and multiclass tasks, Dynomap achieved the strongest performance among all evaluated models (Fig. 4). In the binary healthy versus cancer setting, Dynomap reached 94.4% accuracy with a macro-F1 of 93.3% and macro-sensitivity of 92.1%. This performance exceeded classical linear and tree-based baselines, including Logistic Regression (94.0%) and XGBoost (92.4%), and surpassed MLPs and Random Forests. Notably, Dynomap also outperformed modern deep tabular architectures, TabM (93.7%) and TabPFN (93.7%)[11,29]. For binary classification, Dynomap consistently maintained higher macro-F1 and macro-sensitivity. The performance advantage became more prominent in the multiclass setting spanning five cancer types. Dynomap achieved 59.28% accuracy with a macro-F1 of 57.14%, exceeding Logistic Regression (56.19%), TabM (57.2%), and ModernNCA (57.9% accuracy but with lower macro-F1), and substantially outperforming tree-based ensembles and MLPs. Importantly, Dynomap exhibited the highest macro-sensitivity (55.97%) while maintaining strong macro-specificity (88.79%), indicating that improvements were not confined to dominant classes. In contrast, several deep learning-based baselines displayed elevated specificity but reduced recall across minority classes. This pattern suggests that Dynomap captures structured differences among cancer types rather than exploiting imbalanced separability.

Beyond predictive performance, Dynomap learned spatial layouts that exhibited clear and reproducible class-dependent organization (Fig. 5a). Visualization of the learned embeddings revealed that genes emphasized for each cancer type occupied

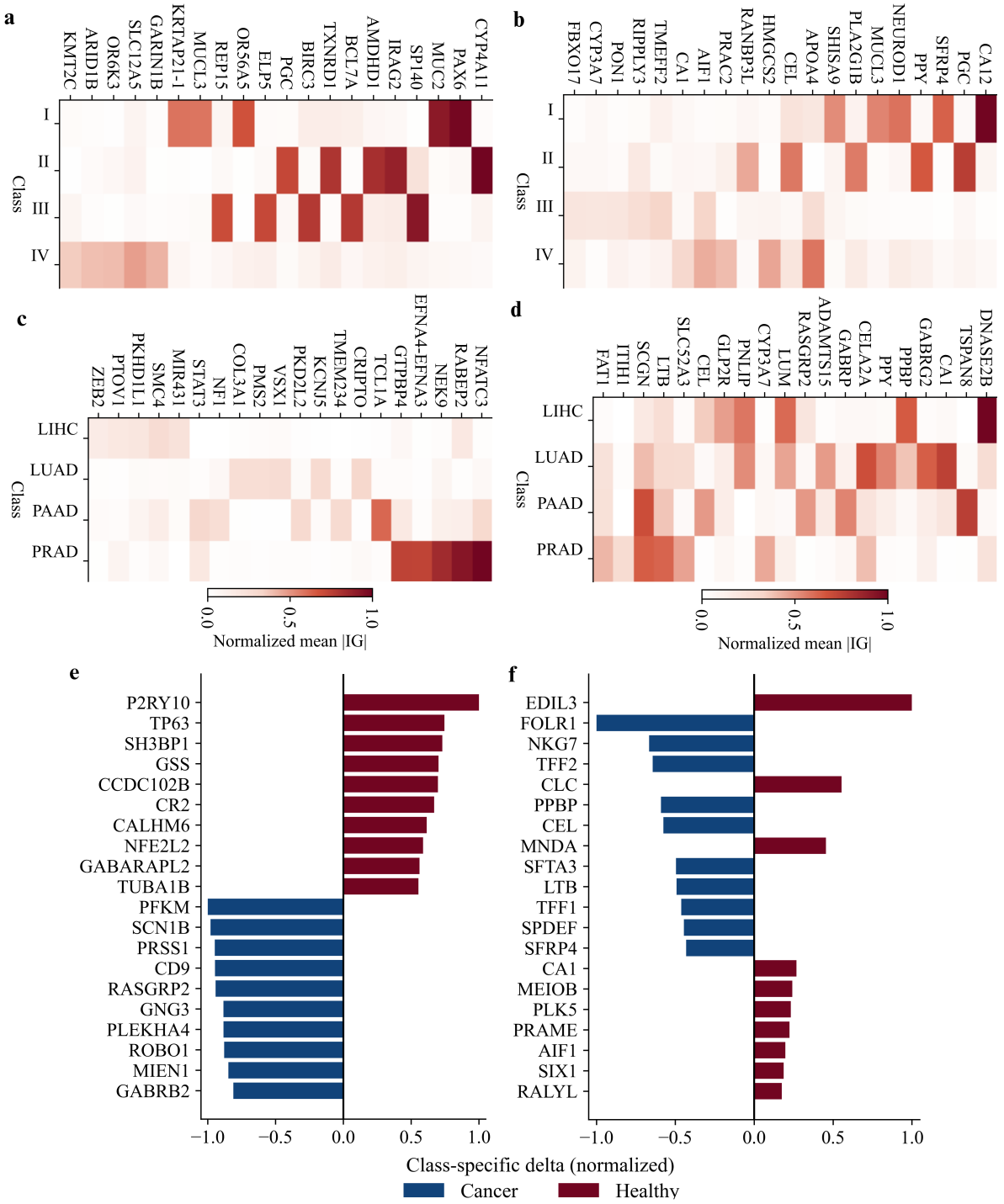

Figure 3. Task-dependent gene attribution patterns learned by Dynomap in liquid biopsy cfRNA. a, Normalized mean absolute Integrated Gradients (|IG|) for cancer stage classification using highly variable genes. Rows correspond to stage classes (I–IV), and columns denote the highest-attribution genes. Distinct stages are associated with different gene attribution profiles. **b**, Stage classification using the 622-gene signature panel. Attribution patterns remain class-dependent. **c**, Multiclass subtype classification using highly variable genes. Each tumor type (LIHC, LUAD, PAAD, PRAD) is characterized by a distinct subset of high-attribution genes. **d**, Subtype classification under the signature panel setting, showing task-adaptive shifts in emphasized genes relative to the full-gene regime. **e**, Differential attribution for binary healthy versus cancer prediction using highly variable genes. Bars indicate normalized class-specific attribution differences (delta |IG|), with positive values favoring healthy prediction and negative values favoring cancer prediction. **f**, Differential attribution under the signature panel setting. Cancer and healthy-associated genes exhibit minimal overlap and form distinct attribution profiles. Together, these panels demonstrate that Dynomap learns task and class-dependent gene importance patterns that vary across prediction objectives and feature regimes.

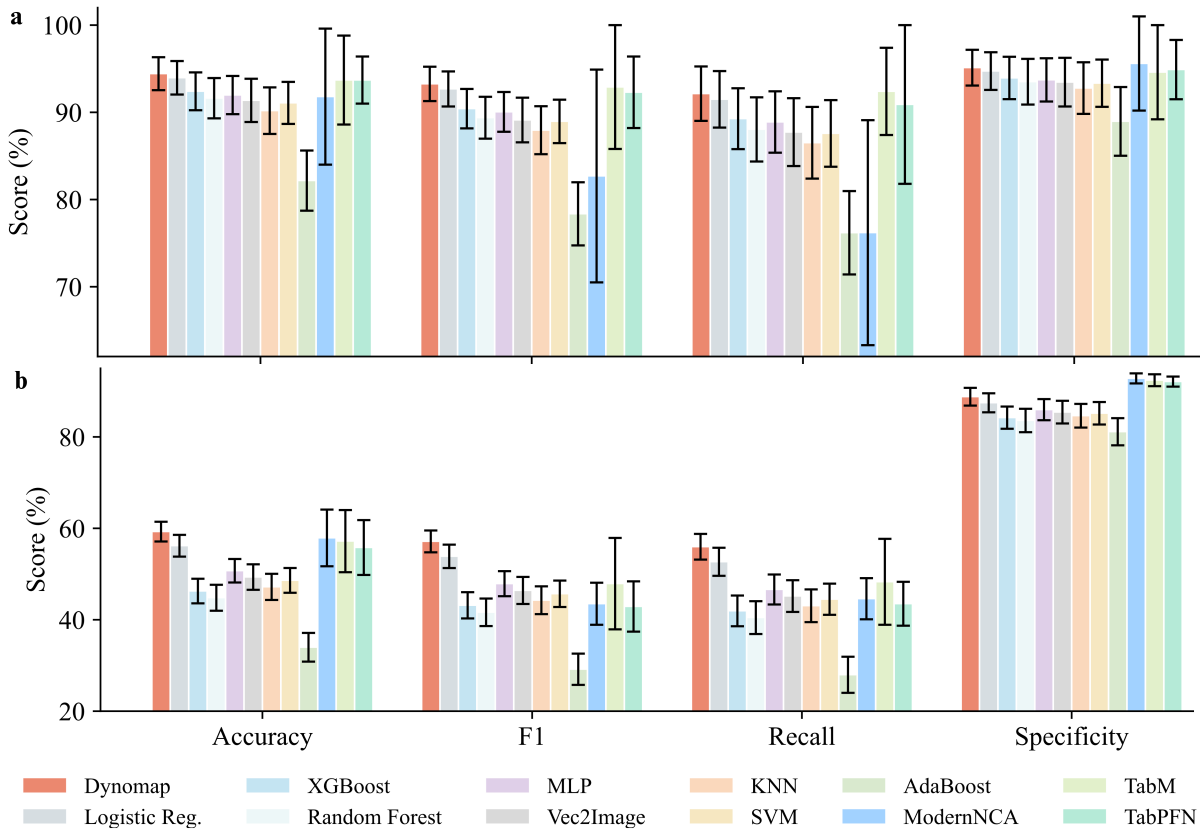

Figure 4. Benchmarking Dynomap on bulk transcriptomic classification (GSE68086). a, Binary healthy versus cancer classification. Performance is reported across accuracy, macro-F1 score, recall, and specificity for Dynomap and eleven baseline methods. Dynomap achieves the highest performance across all reported metrics, with consistent gains in both sensitivity- and specificity-based measures. **b**, Multiclass tissue-of-origin classification. Despite increased class complexity, Dynomap maintains superior accuracy, macro-F1 score, and recall relative to conventional tabular baselines. Specificity remains high across methods, while Dynomap preserves balanced class discrimination. Error bars represent standard deviation across five cross-validation folds.

distinct spatial regions, forming localized clusters rather than diffuse or globally mixed configurations. These clusters were consistent across samples within the same class and differed systematically across classes, indicating that the learned layout preserves class-specific transcriptional structure. We then quantified this organization using global spatial autocorrelation (Moran's I) and local neighborhood consistency (kNN purity[53,54]) in Fig. 5d. For all cancer classes, observed Moran's I values substantially exceeded their corresponding null expectations, indicating that high-attribution genes were spatially clustered rather than randomly dispersed. Similarly, kNN purity scores were elevated across breast, colorectal, glioblastoma, lung, and pancreatic cancers, demonstrating that genes with similar attribution profiles formed coherent local neighborhoods in the learned layout. These findings establish that Dynomap's spatial structure is statistically non-random and consistent across classes.

Integrated Gradients analysis further clarified how predictive signal is organized within this learned spatial manifold[41]. In the binary task, cancer-associated predictions were characterized by elevated attribution for genes including *KDM5B*, *MED12L*, *SLC25A30*, and *MMP1*, spanning epigenetic regulation, Mediator-complex-associated transcription, mitochondrial-carrier biology, and extracellular-matrix remodeling, respectively[55–58], whereas healthy-associated predictions emphasized genes such as *SECISBP2*, *REPS2*, *PTPN18*, and *ENO2*, which have established roles in selenoprotein regulation, receptor and signaling modulation, HER2 phosphorylation and turnover, and neuron-specific enolase biology, respectively[59–62] (Fig. 5c). Importantly, attribution patterns showed minimal overlap between cancer and healthy classes and were spatially localized to distinct regions of the learned map, indicating that class-discriminative signal is not only separable but also spatially segregated.

In the multiclass setting, each cancer type was associated with a distinct subset of high-attribution genes occupying separate spatial regions. Breast cancer predictions emphasized genes such as *F8*, *AP2S1*, and *DOCK11*, whereas pancreatic cancer predictions were associated with elevated attribution for *MMP1* and related extracellular remodeling genes[44,63]. Colorectal and glioblastoma cancers exhibited their own class-specific attribution signatures, including genes such as *RHOC*, *CAMTA1*, and *RPL23A*, which have been reported in tumor-type–specific transcriptional analyses[64,65]. Although certain genes recur across cancer types, their attribution magnitude and spatial localization differed by class, producing non-overlapping attribution hotspots across the map. This pattern indicates class-dependent reweighting of predictive signal.

Notably, the spatial regions highlighted by Integrated Gradients aligned closely with areas of elevated Moran's I and kNN purity. Genes with high attribution for a given cancer type were concentrated within contiguous neighborhoods, whereas low-attribution genes were more diffusely distributed. This concordance between attribution magnitude and spatial clustering suggests that Dynomap organizes transcriptional signals into coherent spatial modules that correspond to class-specific predictive structures. Moreover, genes emphasized in the multiclass task did not necessarily occupy identical spatial regions or exhibit identical attribution strengths in the binary task, demonstrating task-dependent reorganization of the layout. The learned spatial map is therefore not static, but adapts as class structure changes.

Dynomap achieves strong classification performance on bulk transcriptomic data while learning spatial representations in which class-specific predictive signal is organized into statistically quantifiable clusters. The combination of superior macro-balanced performance, elevated spatial autocorrelation, increased neighborhood consistency, and distinct class-dependent attribution patterns indicates that Dynomap captures structured transcriptional variation aligned with the underlying classification objective rather than relying on a fixed gene ordering or purely global separability.

To evaluate Dynomap across a broader set of liquid-biopsy settings, we independently retrained the framework for lung-cancer and colorectal-cancer versus healthy-control classification in plasma cell-free RNA (cfRNA), extracellular-vesicle (EV) RNA, and tumor-educated platelet RNA datasets[40,51,66,67]. The analysis comprised ten cohort- and carrier-specific classification tasks evaluated using donor-disjoint five-fold cross-validation (Supplementary Fig. S32). Pooled donor-level out-of-fold AUROC ranged from 0.835 to 0.982, while mean fold AUROC ranged from 0.817 to 0.982. These results extend the liquid-biopsy application of Dynomap beyond the original RARE-Seq cohort and demonstrate that Dynomap can be independently retrained to achieve high cancer-versus-control discrimination across heterogeneous liquid-biopsy datasets and circulating RNA carriers.

Identifying disease-associated vocal features in Parkinson's disease

We next examined whether the benefits of learned spatial organization in Dynomap extend beyond molecular data to engineered phenotypic measurements. Vocal impairment is among the earliest and most penetrant motor features of Parkinson's disease, and digital voice analysis has been investigated as a low-cost, non-invasive screening tool that can be performed remotely and repeatedly over time[68,69]. Unlike transcriptomic features, acoustic descriptors are constructed quantities derived from signal processing of sustained phonations, and their relationships are governed by the mathematics of feature extraction rather than by any underlying biological pathway. This removes any pathway-informed prior that could assist spatial organization and tests whether Dynomap can recover useful structure from predictive signal alone. We used the dataset reported by Sakar et al., which contains 754 engineered acoustic features spanning time-domain perturbation measures, entropy-based metrics, and tunable Q-factor wavelet transform coefficients, computed from sustained voice recordings of Parkinson's patients and matched

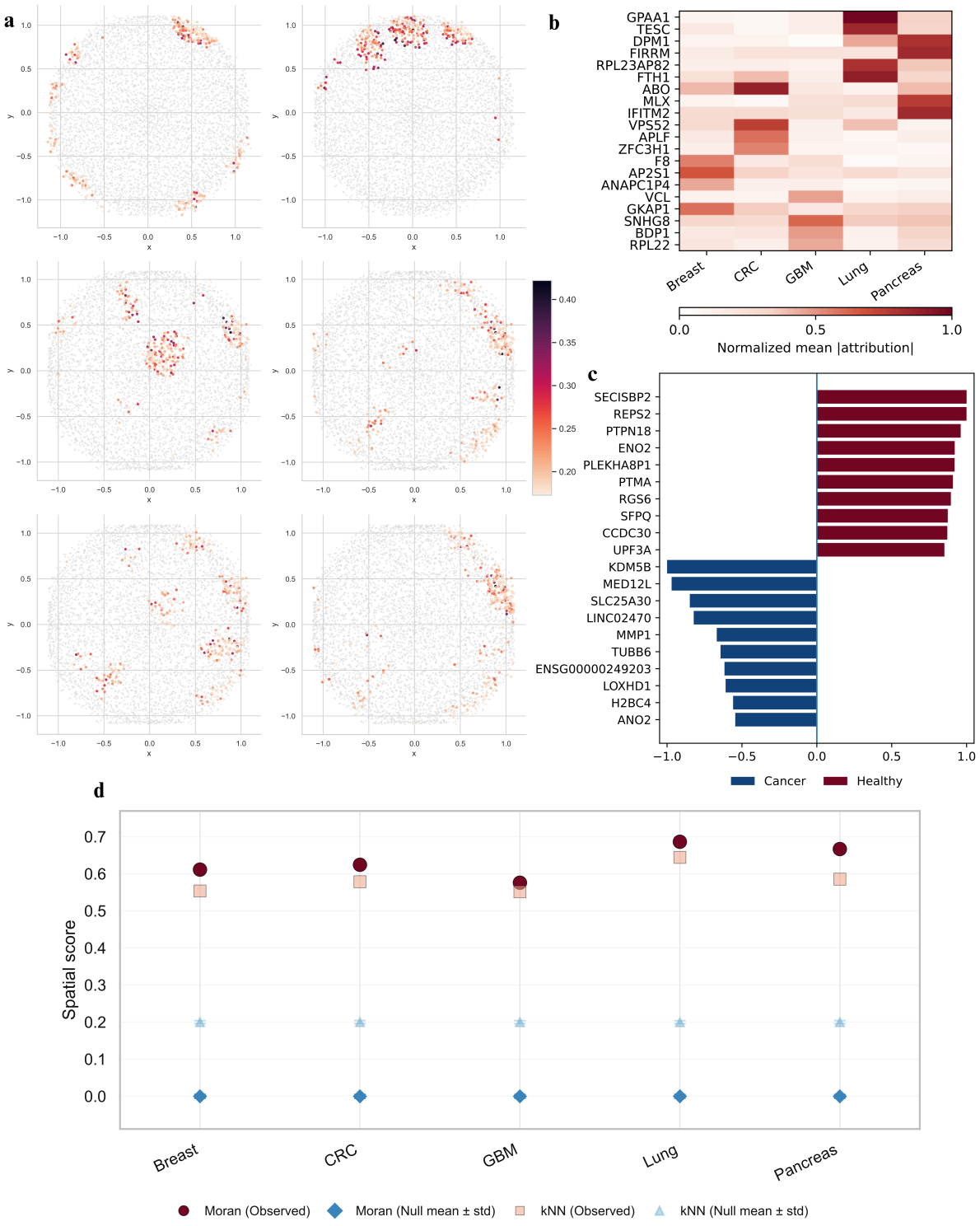

Figure 5. Class-specific spatial organization and attribution structure in bulk transcriptomic data (GSE68086). a, Learned gene layouts with class-specific attribution overlaid. Gray points represent all genes positioned in the optimized two-dimensional layout. Colored points indicate genes with high Integrated Gradients attribution for the specified class. Top row: binary classification (Cancer, left; Healthy, right). Middle row: multiclass setting (Pancreas, left; Breast, right). Bottom row: multiclass setting (Glioblastoma, left; Colorectal cancer, right). Distinct classes activate spatially localized gene regions rather than overlapping uniformly across the map. **b**, Heatmap of normalized mean absolute attribution across cancer types for the highest-ranked genes. Rows correspond to genes and columns to tumor classes. Distinct attribution profiles are observed for different cancer types, with limited overlap among the most emphasized genes. **c**, Differential attribution for binary classification. Bars represent normalized class-specific attribution differences, with positive values favoring healthy prediction and negative values favoring cancer prediction. **d**, Quantification of spatial structure. Moran's I (spatial autocorrelation) and kNN purity scores are shown for each cancer type. Observed values exceed null expectations derived from shuffled layouts, indicating that class-associated genes form statistically significant spatial neighborhoods in the learned representation.

controls[70]. The task is binary classification of Parkinson's versus control, and it serves as a stringent generalization test because none of the architectural assumptions Dynomap makes about features are known to hold in this setting.

Dynomap achieved an accuracy of 93.7% with a macro-F1 score of 91.5% for discrimination between Parkinson's disease and control samples (Fig. 6). Macro-sensitivity reached 96.7%, while macro-specificity was 85.0%. Across all reported metrics, Dynomap substantially outperformed classical machine learning baselines, including Logistic Regression, Random Forest, XGBoost, and PCA+SVM, as well as modern neural tabular models such as TabM, ModernNCA, and TabPFN. Notably, several comparison methods achieved moderate sensitivity but substantially lower specificity, indicating a bias toward predicting the majority class. Dynomap maintained strong sensitivity while improving specificity relative to most competing approaches, resulting in more balanced class discrimination. These differences are reflected in macro-F1 improvements of up to 7.6% compared with existing methods.

The spatial layouts learned by Dynomap for the Parkinson dataset exhibited structured organization. Visualization of the feature embeddings showed that features contributing strongly to Parkinson's disease and control predictions were not uniformly distributed across the learned image space (Fig. 6a). Instead, high-attribution features formed localized regions. Quantitative spatial analysis confirmed this pattern. Observed Moran's I values deviated from null expectations for both Parkinson- and control-associated features, indicating non-random clustering. Similarly, local neighborhood consistency measured by kNN purity exceeded null distributions (Fig. 6a), demonstrating that features with similar attribution profiles occupied nearby spatial neighborhoods. These effects were reproducible across cross-validation folds, and fold-level layout similarity analysis showed stable spatial configurations across training splits (Supplementary Fig. S26).

Integrated Gradients analysis further revealed distinct class-specific attribution patterns (Fig. 6b,c)[41]. Parkinson's disease predictions emphasized a subset of TQWT-derived spectral energy and entropy features, including multiple *tqwt_energy* and *tqwt_entropy_shannon* coefficients, as well as perturbation-related metrics. In contrast, control predictions were associated with a different subset of features, including measures related to vocal stability and dispersion. These class-dependent features occupied partially segregated spatial regions in the learned layout, resulting in minimal overlap between Parkinson- and control-associated attribution hotspots. Many of the emphasized features correspond to acoustic descriptors previously used to characterize vocal impairment in Parkinson's disease, including measures of frequency variation and nonlinear signal dynamics[68–70]. Dynomap was not provided with any grouping or domain-specific constraints on these features. Instead, the model learned a spatial arrangement that concentrated disease-associated signal into coherent neighborhoods aligned with prediction.

Importantly, the representational behavior observed here mirrors patterns seen in transcriptomic benchmarks, despite the fundamentally different feature semantics. In both molecular and phenomic settings, Dynomap learned layouts in which predictive features clustered spatially and reorganized according to the classification objective. This suggests that the spatial learning mechanism is not dependent on biological pathway structure but can operate on engineered acoustic measurements as well. While the Parkinson dataset lacks explicit biological networks, Dynomap nonetheless produced structured, class-dependent spatial organization and competitive predictive performance.

The analysis of the Parkinson dataset shows that even when applied to engineered acoustic descriptors with no explicit biological topology, the model organizes predictive features into coherent spatial regions and reallocates attribution in a class-dependent manner. The learned layouts remain stable across folds and reflect systematic clustering of disease-associated signal rather than random arrangement. These observations indicate that Dynomap's representational behavior is driven by the predictive structure present in the data itself, rather than by assumptions about molecular organization. In this phenomic setting, Dynomap maintains balanced discrimination while preserving interpretable spatial structure, demonstrating that the framework extends naturally to high-dimensional engineered feature spaces.

Stratifying breast cancer molecular subtypes in TCGA-BRCA

We next evaluated Dynomap on a large, clinically annotated breast cancer cohort where molecular subtyping has direct therapeutic implications. PAM50 classification separates breast tumors into Luminal A, Luminal B, Basal-like, HER2-enriched, and Normal-like subtypes, and these categories inform decisions about endocrine therapy, anti-HER2 agents, and chemotherapy selection in routine clinical practice[71,72]. The analytical challenge is that Luminal A and Luminal B share substantial transcriptional overlap, HER2-enriched and Basal-like tumors are clinically aggressive but distinguishable only through specific gene programs, and cohort composition remains imbalanced even at scale[72]. We used the TCGA-BRCA cohort profiled on the HiSeqV2 RNA-seq platform and retained the five PAM50 classes along with a secondary binary task grouping Basal-like and HER2-enriched tumors as aggressive relative to the remaining subtypes[63]. This setting tests whether Dynomap maintains discriminative structure across thousands of genes and hundreds of samples without collapsing transcriptionally similar subtypes, and whether the learned spatial organization reflects the clinically defined subtype hierarchy.

As seen in Fig. 7a, for the binary task distinguishing aggressive from non-aggressive tumors, Dynomap achieved 98.3% accuracy with a macro-F1 of 97.4%, recall of 97.8%, and specificity of 97.1%. This represents the highest overall accuracy

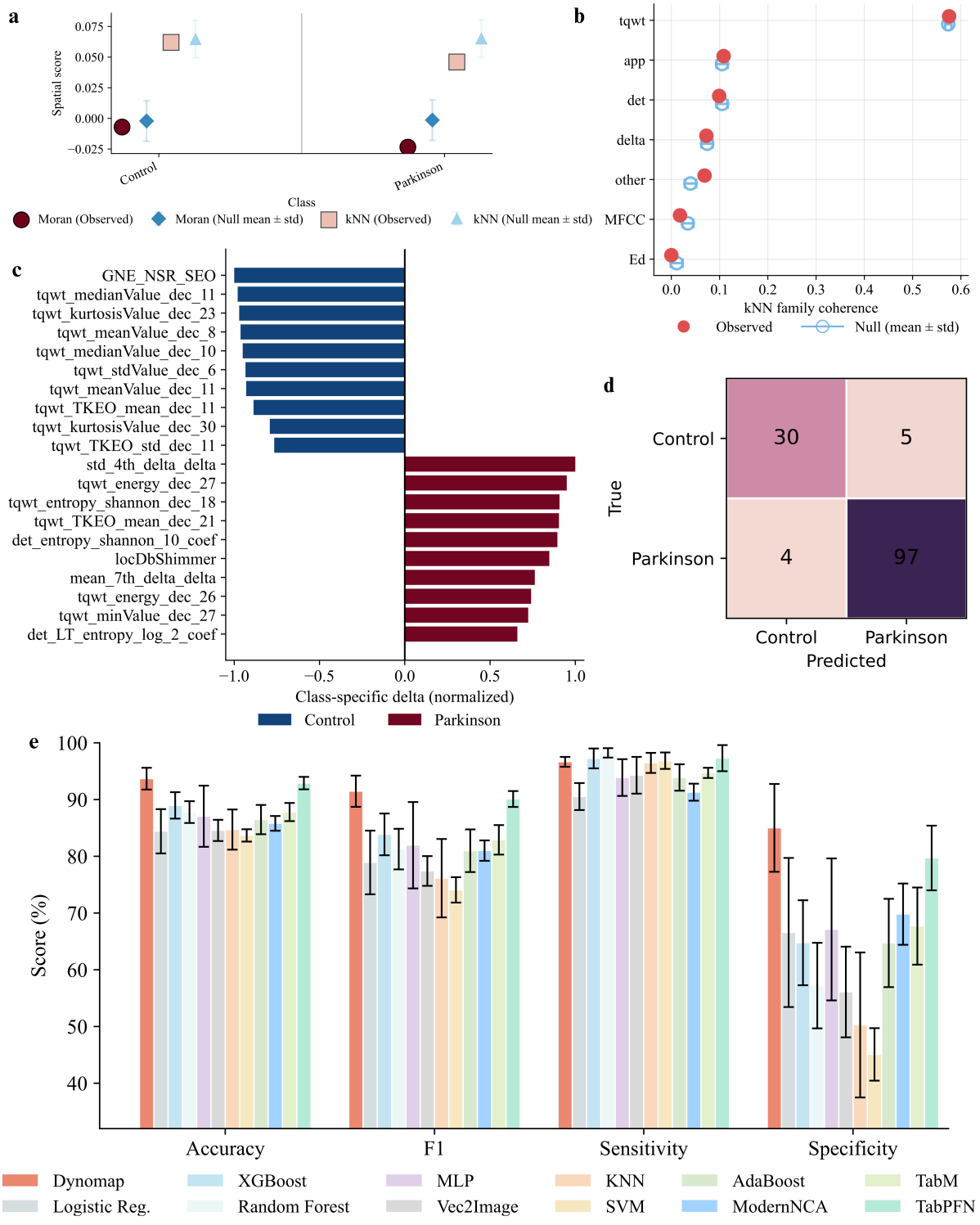

Figure 6. Spatial organization and phenomic feature attribution in Parkinson's disease classification. a, Quantification of spatial structure in the learned feature layout. Moran's I (spatial autocorrelation) and kNN purity are shown for Control and Parkinson classes. Observed values are compared to null expectations derived from shuffled layouts, indicating non-random spatial clustering of class-associated features. **b**, kNN family coherence across engineered feature families (e.g., tqwt, MFCC, delta, and related descriptors). Observed coherence exceeds null distributions, indicating that features derived from similar extraction families occupy nearby spatial regions in the learned layout. **c**, Differential feature attribution for binary classification. Bars represent normalized class-specific attribution differences (delta |IGl), with positive values favoring Parkinson prediction and negative values favoring Control prediction. Distinct sets of acoustic features contribute to each class. **d**, Confusion matrix showing counts of true and predicted labels for Control and Parkinson samples. **e**, Comparative performance across models. Accuracy, macro-F1 score, recall, and specificity are reported for Dynomap and eleven baseline methods. Dynomap achieves the strongest overall performance across class-balanced metrics. Error bars represent standard deviation across five cross-validation folds.

among all evaluated models and a balanced tradeoff between sensitivity and specificity. Tree-based ensembles such as XGBoost achieved comparable recall and F1, while TabM exhibited marginally higher specificity. However, these methods did not simultaneously maximize accuracy and maintain comparable balance across both classes. Precision–recall and ROC curves further show that Dynomap maintains high precision across a broad recall range rather than achieving performance through a narrow high-confidence regime (Supplementary Fig. S22). These results indicate that Dynomap remains competitive with strong ensemble methods while preserving stable class discrimination at scale.

The multiclass PAM50 subtype task presents a more demanding benchmark in Fig. 7b. Dynomap achieved 91.9% accuracy and a macro-F1 of 91.0%, exceeding the strongest baseline (TabM, 89.6% accuracy) and outperforming XGBoost and MLP across macro-balanced metrics. Notably, improvements were most pronounced in macro-F1 and recall, indicating that performance gains were not driven by dominant classes but reflected improved discrimination across all subtypes. While TabM achieved higher macro-specificity, it did so with lower recall, suggesting a more conservative classification boundary. Precision–recall curves show that Dynomap maintains consistently higher precision across recall thresholds (Fig. 7c, Supplementary Fig. S23). These results demonstrate that Dynomap preserves subtype separability even when transcriptional profiles exhibit substantial inter-class similarity.

Integrated Gradients analysis further revealed subtype-specific redistribution of gene attribution (Fig. 7d). Distinct PAM50 classes were characterized by different high-attribution gene sets with limited overlap among top-ranked features. For example, LumB predictions showed elevated attribution for *PRODH*, *TCEAL2*, and *FAM189A2*, whereas HER2-associated predictions emphasized *B4GALNT2* and *FABP6*. LumA exhibited stronger attribution for *TMEM59L*, *SYBU*, and *GRP*, while Basal predictions highlighted genes including *FBP1* and *TMCC2*. Normal-like samples emphasized a separate subset, including *RYR3*, *PLK1*, and *CCDC64*. These patterns indicate that Dynomap does not rely on a single global gene ranking across classes. Instead, attribution magnitude and spatial localization adapt to the multiclass structure of the task.

A similar redistribution was observed in the binary grouping. Aggressive cancer predictions were associated with genes such as *C6orf211* (*ARMT1*), *ARSG*, *ACADSB*, *CMYA5*, and *MS4A2*, whereas non-aggressive predictions emphasized *KCNQ5*, *DLGAP1*, *SCAND3*, *FAM129A* (*NIBAN1*), and *HHIPL2*. Several of these genes are linked to distinct biological programs, including immune receptor signaling (*MS4A2*)[73], mitochondrial metabolism (*ACADSB*)[74], and PCNA-associated regulation (*ARMT1*)[75], as well as neuronal or synaptic gene programs (*KCNQ5*, *DLGAP1*)[76,77] and stress-response or proliferative regulation (*FAM129A*, *SCAND3*, *HHIPL2*)[78–80]. These class-specific attribution profiles were spatially localized and showed minimal overlap in high-contribution regions. Importantly, genes emphasized in the binary task did not necessarily occupy identical spatial positions or maintain identical attribution magnitudes in the multiclass setting, indicating task-dependent reorganization of predictive weight.

Overall, the TCGA-BRCA analysis demonstrates that Dynomap scales to large, high-dimensional transcriptomic cohorts while preserving clinically defined subtype structure. The model achieves strong balanced performance in both binary and multiclass settings and learns stable, non-random spatial layouts in which subtype-associated genes form coherent modules. Rather than enforcing a static feature hierarchy, Dynomap reallocates attribution across genes according to task structure and class definition, maintaining interpretability and discriminative performance under substantial heterogeneity.

Classifying cell identity across a large single-cell transcriptomic atlas

Finally, we tested Dynomap on single-cell transcriptomics, where cell-type classification is a foundational step in the analysis of atlas-scale datasets. Accurate cell identity assignment underpins downstream analyses ranging from trajectory inference and gene regulatory network reconstruction to the discovery of rare or disease-associated populations[81]. The challenge at atlas scale is not binary discrimination but fine-grained separation among many transcriptionally similar populations, particularly closely related immune subsets and stromal lineages that share core gene programs. We used the Tabula Muris atlas, which contains 54,865 cells annotated into 55 cell types across 20 mouse organs[81]. This benchmark probes whether Dynomap can maintain balanced performance across high class cardinality and whether the learned spatial organization recovers lineage-aligned structure when class boundaries are subtle and partially hierarchical.

Across this multiclass setting, Dynomap achieved 97.6% accuracy with a macro-F1 score of 92.0%, a macro-sensitivity of 90.0%, and a macro-specificity of 100.0% (Fig. 8a). While overall accuracy approached the ceiling for several high-capacity models, macro-balanced metrics differentiated performance under extensive class heterogeneity. Dynomap maintained higher macro-F1 than logistic regression (87.9%), random forest (87.9%), and KNN (85.7%), and substantially outperformed PCA+SVM and AdaBoost. Tree-based ensemble methods achieved comparable overall accuracy, with XGBoost reaching 97.5% accuracy and a macro-F1 of 93.2%, and TabM achieving 97.8% accuracy with a macro-F1 of 93.7%. Despite these narrow margins among the strongest ensemble-based approaches, Dynomap preserved competitive macro-balanced performance while maintaining near-perfect specificity, indicating stable discrimination across a large number of closely related cell identities. Precision–recall analysis further demonstrated that Dynomap sustained high precision across a broad recall range rather than concentrating performance on a limited subset of highly separable cell populations (Fig. 8c, Supplementary Fig. S15).

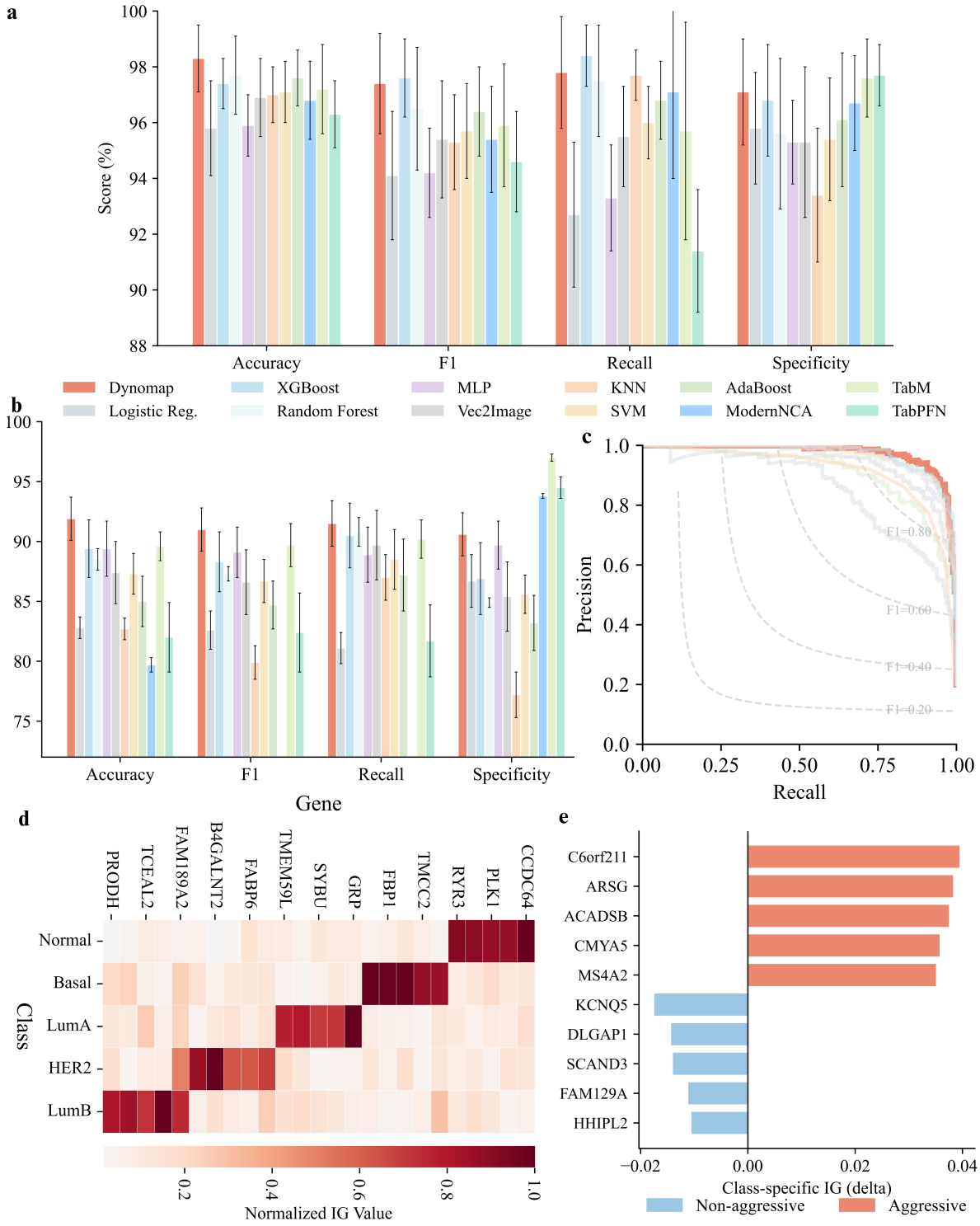

Figure 7. Binary and molecular subtype classification in breast cancer transcriptomes (TCGA-BRCA). a, Comparative performance of Dynomap and eleven baseline methods for binary classification of non-aggressive versus aggressive tumors. Metrics include accuracy, macro-F1 score, recall, and specificity. Dynomap achieves the highest overall performance across class-balanced measures. Error bars indicate standard deviation across five cross-validation folds. **b**, Multiclass molecular subtype classification (Basal, HER2, LumA, LumB, Normal). Dynomap maintains superior accuracy and balanced performance across macro-F1, recall, and specificity relative to conventional and advanced tabular baselines. **c**, Precision-Recall curve for subtype classification. **d**, Normalized Integrated Gradients attribution for representative genes across molecular subtypes. Distinct gene groups are preferentially associated with individual subtypes. **e**, Differential attribution for binary stratification into non-aggressive and aggressive tumors. Bars represent class-specific attribution differences (delta |IG|), with positive values indicating greater contribution to aggressive prediction and negative values to non-aggressive prediction.

Given the high number of distinct classes, we next examined whether the learned representation exhibits structured spatial organization rather than diffuse attribution. Spatial autocorrelation analysis revealed that observed Moran's I values consistently exceeded null expectations derived from shuffled layouts across representative cell types (Fig. 8b; Supplementary Fig. S16)[53,54]. Similarly, kNN purity scores were high relative to null distributions, indicating that genes contributing strongly to a given cell type formed coherent local neighborhoods in the learned layout.

Integrated Gradients analysis provided a global view of class-specific attribution across cell types (Fig. 8d; Supplementary Fig. S20)[41]. Instead of showing a uniform importance pattern, high-attribution genes appeared in lineage-restricted bands with minimal cross-class overlap. T-cell–associated populations had elevated attribution for *Cd3g*, a core component of the T-cell receptor complex required for antigen recognition and signal transduction[82]. B-cell–related populations emphasized *Cd79b* and *Cd74*, central mediators of B-cell receptor signaling and MHC class II–dependent antigen presentation[83,84]. Stromal and mesenchymal populations preferentially highlighted extracellular matrix genes such as *Col1a1* and *Col3a1*, which encode fibrillar collagens characteristic of connective tissue and fibroblast identity[85]. Epithelial subtypes exhibited stronger attribution for keratin-associated genes including *Krt5* and *Krt15*, established markers of epithelial differentiation programs[86]. These lineage-aligned attribution modules are consistent with coordinated gene programs underlying cell identity in single-cell atlases[81,87].

Importantly, high-attribution genes were not globally dominant across multiple classes. Pairwise similarity analysis of attribution profiles showed strong diagonal dominance with limited off-diagonal overlap (Supplementary Fig. S17), indicating class-specific redistribution of predictive weight. Hierarchical clustering of cell-type attribution profiles further revealed biologically coherent groupings, with closely related immune populations forming neighboring clusters and stromal and epithelial lineages segregating into distinct branches (Supplementary Figs. S18, S19). Attribution profiles were lineage-specific, with limited overlap across cell types.

In summary, these results show that Dynomap scales to high-dimensional single-cell data with extensive class cardinality while preserving structured spatial organization and lineage-specific attribution patterns. In a setting characterized by subtle transcriptional boundaries and hierarchical relationships among 55 cell types, the model maintains strong macro-balanced performance and learns non-random spatial layouts in which predictive signal concentrates into biologically coherent modules. Rather than relying on a static ranking of genes, Dynomap dynamically redistributes attribution according to cell identity, indicating that its representational behavior remains structured and task-adaptive under large-scale single-cell heterogeneity.

Discussion

Tabular biomedical data do not have a natural spatial structure. Each feature is treated as a separate column, and relationships among features must be learned from the data alone[11,12]. This differs from images and structured signals, where nearby pixels or measurements often carry related information. Vision deep learning models are effective in those settings because they can exploit locality, weight sharing, and hierarchical feature aggregation[14,38,39]. When the same models are applied directly to tabular data, the row-column layout does not tell the model which features are biologically or clinically related[11]. Dynomap addresses this representational gap by learning spatial organization jointly with prediction (Fig. 1 and Fig. S1). Feature coordinates are optimized through a differentiable rendering process that is coupled directly to the prediction objective, rather than imposed through predefined similarity measures or external grouping rules. Across all datasets examined here, there are two consistent observations. First, Dynomap achieves strong predictive performance across regimes that vary in feature dimensionality, class imbalance, biological heterogeneity, and data modality (Figs. 2–8). Second, the learned spatial representations exhibit structured organization in which predictive features self-organize into coherent, task-dependent neighborhoods that are statistically distinguishable from null spatial arrangements (Figs. 3, 5, 6, 7, 8; Supplementary Figs. S3–S5, S16, S25). Taken together, these findings support the view that end-to-end spatial representation learning can extend convolutional inductive biases to traditionally non-spatial tabular domains while preserving interpretability.

To test how broadly the framework extends beyond the five focused biomedical datasets analyzed above, we evaluated Dynomap on 13 additional public benchmarks (Supplementary Figs. S27, S28, S29, S30, and S31). The collection comprises nine pairwise microarray cancer classification tasks across diverse tissue origins, a heavily imbalanced ovarian cancer microarray dataset, two mass spectrometry cancer cohorts from the Arcene benchmark, and a six-class smartphone-based human activity recognition dataset drawn entirely from outside the biomedical domain[10,88–90]. Sample sizes range from 100 to 10,299, and feature counts span 561 to 10,935. Across this collection, Dynomap achieves a mean improvement of 1.5% over the strongest classical baseline (Logistic Regression, Random Forest, XGBoost, AdaBoost, KNN, MLP, PCA+SVM, and Vec2Image) and 0.8% over the strongest deep tabular baseline among TabM, TabPFN, and ModernNCA. The gain is largest on small, high-dimensional microarray cohorts where classical learners saturate, reaching up to 9.0% absolute accuracy improvement on AP Endometrium vs Omentum relative to MLP. The advantage persists on the imbalanced OVA Ovary task, where accuracy and macro-F1 diverge by more than ten percentage points across all evaluated methods, indicating that Dynomap captures discriminative signal beyond majority-class agreement. On the activity recognition benchmark, which uses engineered inertial

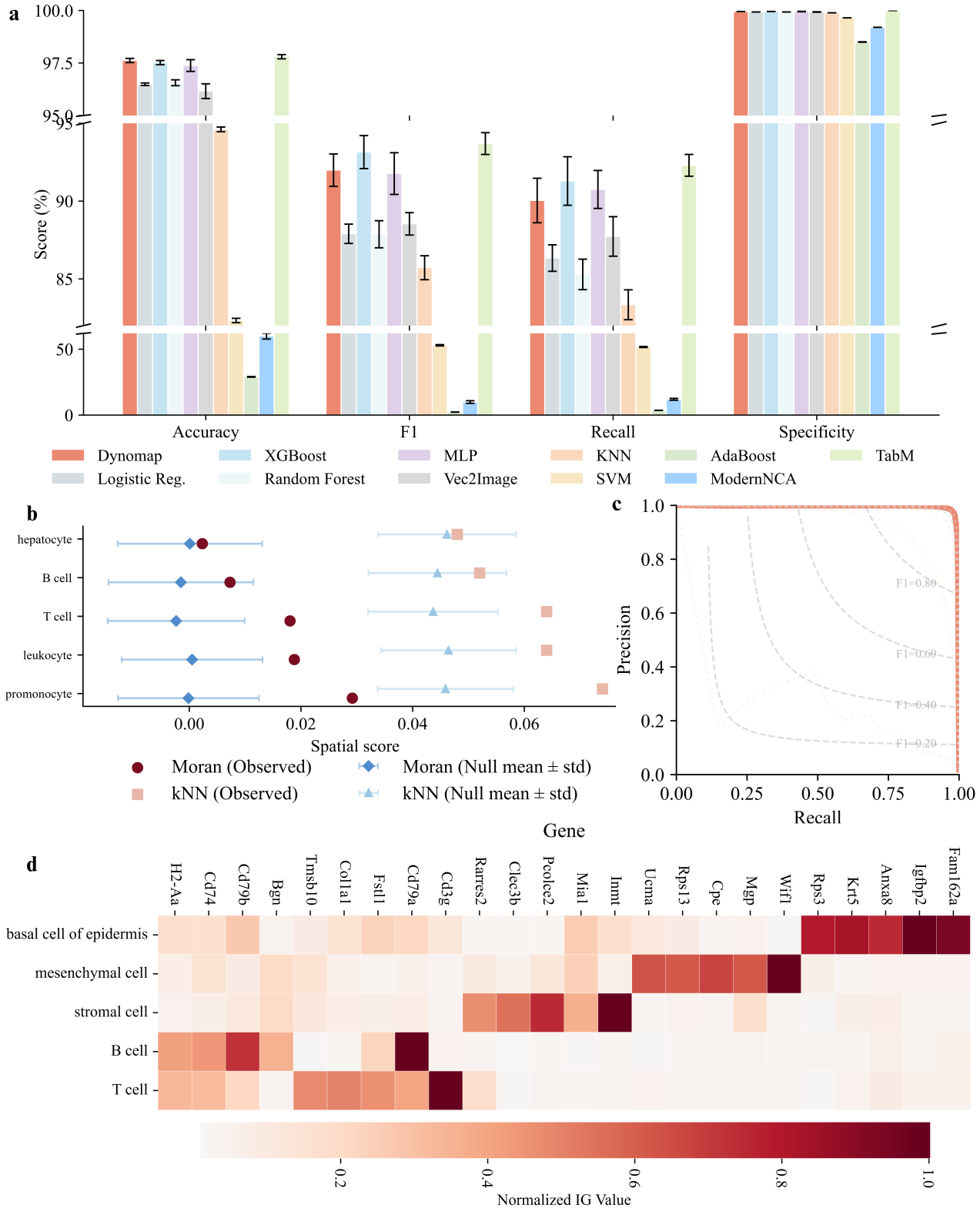

Figure 8. Cell-type classification and spatial organization in single-cell transcriptomes (Tabula Muris). a, Comparative performance of Dynomap and eleven baseline methods for multiclass cell-type classification. Metrics include accuracy, macro-F1 score, recall, and macro-specificity. Dynomap achieves the highest overall performance across class-balanced measures. Error bars indicate standard deviation across cross-validation folds. **b**, Spatial organization analysis for representative cell types. Moran's I (spatial autocorrelation) and kNN purity are shown alongside null expectations derived from shuffled layouts. Observed values exceed null distributions, indicating non-random clustering of cell-type–associated genes in the learned layout. **c**, Precision–recall curve demonstrating stable performance across recall thresholds. **d**, Normalized Integrated Gradients attribution for representative genes across selected cell types. Distinct gene groups are preferentially associated with individual cellular identities, indicating class-dependent attribution structure within the learned spatial representation.

sensor features and spans six classes, Dynomap retains its lead at 98.96% accuracy, exceeding both the strongest classical method (XGBoost, 98.07%) and the strongest deep tabular method (TabPFN, 98.54%). These results extend the trend observed on the focused biomedical datasets and show that the learned spatial cartography yields a consistent advantage over both classical and state-of-the-art deep tabular learners across heterogeneous data regimes, including settings well outside molecular biology.

A central observation across experiments is that Dynomap's performance advantage holds across widely varying feature spaces, from compact curated panels to full transcriptome-scale inputs. In RARE-Seq, strong performance is maintained under both full-gene and restricted signature-panel settings, and across binary detection, stage prediction, and subtype classification (Fig. 2). When the number of highly variable genes is systematically varied over a broad range, performance remains stable relative to competitive baselines (Supplementary Fig. S11). In high-dimensional biomedical settings, predictive signal is often distributed and weak, and model performance can depend strongly on feature filtering and regularization strategies[12,19]. The observed stability indicates that Dynomap does not rely on a single curated subset of features but adapts its spatial organization as the feature space expands or contracts. Similar robustness is observed in TCGA-BRCA, where subtype prediction under class imbalance remains competitive when evaluated using precision–recall curves, a metric better suited to imbalanced classification than accuracy or ROC analysis[91] (Fig. 7; Supplementary Figs. S22, S23). In Tabula Muris, Dynomap maintains competitive performance despite a large number of closely related cell types with overlapping transcriptional programs[81] (Fig. 8). Across these regimes, the representation reorganizes with the task rather than depending on a fixed feature ordering, supporting generalization across heterogeneous biomedical contexts.

The multi-cohort liquid-biopsy analysis further illustrates the breadth of the Dynomap framework. Following independent cohort-specific training, Dynomap achieved high cancer-versus-control discrimination across plasma cfRNA, EV RNA, and platelet RNA datasets (Supplementary Fig. S32). This finding broadens the application of Dynomap from a single ultrasensitive cfRNA cohort to heterogeneous liquid-biopsy assays with different RNA compositions and measurement characteristics[51,66,67]. Together, these results support Dynomap as a common, cohort-adaptive representation-learning framework that can be applied across liquid-biopsy datasets while allowing feature selection, spatial organization, and predictive parameters to be learned separately within each cohort.

Interpretability analyses further demonstrate that the learned spatial organization reflects spatial organization that is statistically distinguishable from noise. Integrated Gradients attribution maps reveal localized regions that differ across objectives within the same dataset (Figs. 3 and 7), consistent with the principle that meaningful interpretation requires alignment between model structure and task-specific feature interactions[92]. In RARE-Seq, shifting from binary detection to stage or subtype prediction reorganizes high-attribution genes into distinct spatial neighborhoods despite identical underlying features (Fig. 3; Supplementary Figs. S3–S5). Spatial statistics provide quantitative support for this structure. Moran's I[53] and kNN-based neighborhood coherence consistently exceed shuffled-layout baselines across bulk transcriptomic, single-cell, and phenomic datasets (Fig. 5d; Fig. 6c; Supplementary Figs. S16, S25). These results indicate that clustering of predictive features is unlikely to arise from arbitrary coordinate assignment. In Tabula Muris, the Jaccard similarity matrix shows strong within-class neighborhood stability and structured overlap among related lineages (Supplementary Fig. S17), consistent with shared developmental and transcriptional programs across immune and stromal populations[81]. High-attribution genes include canonical lineage markers that occupy coherent spatial regions (Fig. 8d; Supplementary Fig. S20), supporting the biological plausibility of the learned layouts.

The emergence of coherent spatial neighborhoods follows directly from the optimization framework in the proposed Dynomap approach. Convolutional filters aggregate information locally and are most effective when spatial proximity corresponds to meaningful relationships[14,35,38]. When tabular features are arranged without structure, local aggregation combines unrelated variables, limiting efficiency. By learning feature coordinates jointly with prediction, Dynomap aligns spatial proximity with predictive co-activation patterns. This alignment allows convolutional layers to capture distributed interactions through localized receptive fields while preserving feature identity. The repeated observation of task-dependent spatial reorganization across independent datasets suggests that the learned layouts reflect objective-driven structure rather than arbitrary embeddings.

Outlook

Dynomap demonstrates that learning spatial organization directly from the prediction objective provides a general strategy for transforming heterogeneous tabular biomedical data into structured and interpretable representations. Across tasks that vary in dimensionality, imbalance, and biological complexity, the framework produces stable spatial layouts in which predictive features form coherent, task-adaptive neighborhoods associated with disease state, cellular identity, or phenotypic condition. In multiple benchmarks, Dynomap improves predictive accuracy relative to established approaches and maintains performance when the feature regime is restricted from a full feature set to a signature panel, indicating that the learned topography adapts to the available biological signal. By providing an explicit, task-specific feature organization through a differentiable mapping,

Dynomap enables attribution-guided inspection and hypothesis generation in settings where feature interactions are otherwise implicit. More broadly, spatial representation learning enables relationships among molecular, cellular, and clinical variables to be encoded into structured spatial maps, revealing latent organizational patterns underlying tumor heterogeneity, disease progression, and treatment response. This framework is applicable across multiple biological scales, including genomics, transcriptomics, radiomics, digital pathology, and multimodal clinical datasets, where complex interactions may otherwise remain obscured in conventional tabular analysis. These capabilities support improved biomarker discovery, patient stratification, and outcome prediction, and provide a foundation for integrative analyses that unify molecular and clinical data in precision medicine settings.

Methods

Overview of the Dynomap framework

Dynomap converts high-dimensional tabular inputs into structured spatial representations that can be processed by convolutional models. The framework assigns each feature a learnable coordinate in a two-dimensional space and renders each sample into an image using a differentiable kernel-based mapping. Unlike fixed vector-to-image transformations, both spatial coordinates and rendering parameters are optimized jointly with the prediction objective. Let $\mathbf{x} \in \mathbb{R}^G$ denote an input feature vector with G features. The model learns a mapping from $\mathbf{x}$ to a spatial image $I \in \mathbb{R}^{P\times P}$, while simultaneously optimizing a classifier defined on the rendered representation.

Feature gating

Before spatial organization, feature magnitudes are adaptively reweighted to allow task-dependent emphasis or suppression. Given an input vector $\mathbf{x}$, a dense layer with sigmoid activation produces gating weights $\mathbf{g} \in (0,1)^G$,

$$\mathbf{g} = \sigma(\mathbf{W}_g\mathbf{x} + \mathbf{b}_g),$$

where σ denotes the sigmoid function. The gated feature vector is

$$\tilde{\mathbf{x}} = \mathbf{x} \odot \mathbf{g}.$$

This operation provides a differentiable mechanism for feature modulation prior to spatial rendering, allowing the representation to adapt to the prediction objective.

Learnable spatial layout

To introduce structure into unordered features, Dynomap assigns each feature a trainable coordinate in a two-dimensional space. For each feature i, a coordinate $\mathbf{c}_i \in \mathbb{R}^2$ is learned. Coordinates are initialized uniformly in $[-1,1]$ and optimized jointly with model parameters. To prevent degenerate configurations, spatial dispersion is enforced through three regularization terms.

A centering term ensures zero-mean layout,

$$\mathscr{L}_{\text{center}} = \left\| \frac{1}{G}\sum_{i=1}^{G} \mathbf{c}_i \right\|_2^2.$$

A spread term constrains global variance,

$$\mathscr{L}_{\text{spread}} = \left(\text{std}(\{\mathbf{c}_i\}) - 1\right)^2.$$

A repulsion term discourages coordinate collapse.

$$\mathscr{L}_{\text{repel}} = \mathbb{E}_{i\neq j}\left[\frac{1}{\|\mathbf{c}_i - \mathbf{c}_j\|_2^2 + \varepsilon}\right],$$

with $\varepsilon = 10^{-4}$. The total layout regularization is

$$\mathscr{L}_{\text{layout}} = 0.01\mathscr{L}_{\text{center}} + 0.01\mathscr{L}_{\text{spread}} + 0.05\mathscr{L}_{\text{repel}}.$$

These terms maintain spatial dispersion while allowing coordinates to reorganize according to predictive structure.

Differentiable rendering

Once spatial coordinates are defined, each sample is rendered into a two-dimensional image of resolution $P \times P$, where P denotes the number of pixels per spatial dimension and is treated as a model hyperparameter. This step converts feature values into spatial intensity patterns while preserving differentiability. For each grid location (u, v) in the $P \times P$ image,

$$I(u,v) = \sum_{i=1}^{G} \tilde{x}_i \exp\left(-\frac{\|(u,v) - \mathbf{c}_i\|_2^2}{2\sigma_i^2}\right).$$

Kernel width is feature-dependent,

$$\sigma_i = 0.5 + 4.5 \tanh(|\tilde{x}_i|).$$

The rendering grid is dynamically defined based on the coordinate bounding box with fixed padding. Each rendered image is standardized per sample. Because the rendering operation is differentiable with respect to both $\tilde{\mathbf{x}}$ and $\mathbf{c}_i$, gradients propagate through the spatial mapping during optimization.

End-to-end joint optimization

All components of Dynomap, including feature gating, spatial coordinates, rendering, and classification layers, are optimized jointly under a unified objective. Let $f(\mathbf{x}; \theta)$ denote the full model with parameters $\theta = \{\theta_{\text{gating}}, \theta_{\text{layout}}, \theta_{\text{CNN}}, \theta_{\text{head}}\}$. The total loss is defined as

$$\mathscr{L}_{\text{total}} = \mathscr{L}_{\text{CE}}(f(\mathbf{x}; \theta), y) + \mathscr{L}_{\text{layout}},$$

where $\mathscr{L}_{\text{CE}}$ denotes categorical cross-entropy and $\mathscr{L}_{\text{layout}}$ represents the spatial regularization terms described above. Model parameters are updated via gradient-based optimization,

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta \mathscr{L}_{\text{total}},$$

where η denotes the learning rate.

Because the rendering operation is differentiable with respect to both feature values and spatial coordinates, gradients propagate from the classification loss through the convolutional layers and rendering module to the coordinate parameters. As a result, spatial organization evolves directly in response to predictive error, allowing feature neighborhoods to adapt to the task objective.

Layout stabilization

During early training, spatial coordinates may undergo substantial movement as the model searches for predictive organization. To prevent continuous drift and to stabilize topology, coordinate velocity is monitored across epochs. Mean displacement is defined as

$$v^{(t)} = \frac{1}{G} \sum_{i=1}^{G} \left\| \mathbf{c}_i^{(t)} - \mathbf{c}_i^{(t-1)} \right\|_2.$$

If $v^{(t)}$ falls below a predefined threshold for a fixed number of consecutive epochs, coordinate updates are halted and the layout is frozen. For the primary TensorFlow experiments, model selection was performed only after stabilization to ensure evaluation under a fixed spatial configuration. The auxiliary multi-cohort liquid-biopsy analysis used the separate PyTorch training and checkpoint-selection protocol described below.

Convolutional modeling and training

Rendered images of resolution $P \times P$ with $P = 64$ are processed using a CNN designed to capture hierarchical spatial structure through shared local filters[14,38]. The backbone consists of sequential two-dimensional convolutional layers with ReLU activation[93], applied with same padding to preserve spatial dimensions. Max pooling layers reduce spatial resolution, and global pooling aggregates spatial responses into a compact feature vector.

To preserve direct access to original tabular information, the convolutional feature vector is concatenated with the gated feature vector $\tilde{\mathbf{x}}$. The combined representation is passed through a fully connected classifier head consisting of a dropout layer with rate 0.5[94], a hidden dense layer with 64 units and ReLU activation, and a final linear layer producing class logits.

For the primary benchmark experiments, the model was trained by minimizing categorical cross-entropy loss with label smoothing[95]. Optimization was performed using the Adam optimizer[96] with learning rate 1×10^{-3}. Mini-batch size was set to

16, and models were trained for up to 1000 epochs. During training, spatial coordinates were monitored for stabilization as described above (patience = 15 epochs, velocity threshold = 0.002).

The primary benchmark models were implemented in TensorFlow with reproducible random seed initialization across Python, NumPy, and TensorFlow operations. GPU acceleration was used during training, and identical hyperparameters were applied across the primary benchmark datasets to ensure comparability. The auxiliary multi-cohort liquid-biopsy analysis used a nested evaluation protocol described in its dataset-specific Methods subsection.

Integrated Gradients attribution

To quantify feature contributions to model predictions, we employed Integrated Gradients (IG)[41]. Let $f(\mathbf{x})$ denote the logit corresponding to a target class and $\mathbf{x} \in \mathbb{R}^G$ the input feature vector. Given a baseline input $\mathbf{x}'$ (set to the zero vector), the integrated gradient for feature i is defined as

$$\mathrm{IG}_i(\mathbf{x}) = (x_i - x_i') \int_0^1 \frac{\partial f\left(\mathbf{x}' + \alpha(\mathbf{x} - \mathbf{x}')\right)}{\partial x_i} \, d\alpha.$$

This formulation accumulates gradients along the straight-line path from the baseline to the input. In practice, the integral is approximated using a Riemann sum with m discrete steps. For multiclass settings, IG was computed with respect to the logit of the predicted class. Attribution magnitudes were aggregated across samples within each class to obtain class-specific importance profiles. These attribution vectors were subsequently used for spatial and clustering analyses.

Spatial autocorrelation using Moran's I

To quantify whether high-attribution features cluster spatially beyond random expectation, we computed Moran's I statistic[53]. Let a_i denote the attribution score for feature i and let $\mathbf{c}_i$ denote its learned spatial coordinate. A spatial weight matrix $W = \{w_{ij}\}$ is defined based on neighborhood relationships in the learned layout, where $w_{ij} = 1$ if feature j is among the k nearest spatial neighbors of feature i, and $w_{ij} = 0$ otherwise. Let $\bar{a}$ denote the mean attribution value and N the number of features.

Moran's I is defined as

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_i \sum_j w_{ij}(a_i - \bar{a})(a_j - \bar{a})}{\sum_i (a_i - \bar{a})^2}.$$

Positive values indicate spatial clustering of similar attribution magnitudes, whereas values near zero indicate spatial randomness. To assess statistical significance, we generated null distributions by randomly permuting attribution values across fixed coordinates and recomputing Moran's I across multiple iterations.

k-nearest neighbor spatial coherence

In addition to global spatial autocorrelation, we evaluated local neighborhood consistency using a k-nearest neighbor (kNN) coherence measure. For each feature i, let $\mathcal{N}_k(i)$ denote the set of its k nearest neighbors in the learned coordinate space. Define a binary indicator z_i reflecting whether feature i belongs to the top-q% highest attribution values for a given class.

The local coherence score for feature i is defined as

$$C_i = \frac{1}{k} \sum_{j \in \mathcal{N}_k(i)} \mathbf{1}(z_j = z_i),$$

where $\mathbf{1}(\cdot)$ denotes the indicator function. The overall spatial coherence is computed as the mean of C_i across features classified as high attribution. This statistic quantifies whether high-importance features preferentially cluster within their local spatial neighborhoods.

Significance was assessed by comparing observed coherence scores to null distributions obtained by randomly permuting attribution labels across fixed coordinates. Together with Moran's I, this provides complementary global and local measures of spatial organization in the learned layout.

Datasets

Unless otherwise specified, highly variable genes (HVGs) were selected based on variance across samples or cells within each dataset. All tabular inputs were standardized prior to modeling.

Circulating cell-free RNA profiling (RARE-Seq).

We evaluated Dynomap on a circulating cell-free RNA (cfRNA) cohort derived from a multi-cancer liquid biopsy study[40]. The dataset contains plasma-derived cfRNA profiles spanning healthy individuals and multiple cancer types with annotated clinical stage and tissue of origin. Gene-level expression matrices were constructed from TPM values and restricted to Ensembl gene identifiers. Features with zero variance across samples were removed. Three classification tasks were defined. For binary detection, samples were grouped into Healthy (Control, LDCT Control, Meta-Reference Control) and Cancer (LUAD, LUAD (TKI), LIHC, PAAD, PRAD). For subtype prediction, cancer samples were mapped to LUAD, LIHC, PAAD, and PRAD categories, with LUAD (TKI) grouped under LUAD. For stage classification, only cancer samples with available stage annotations were retained, and samples with missing stage information were excluded. For the full-gene setting, we selected the top 4,000 highly variable genes based on variance across samples. In the signature setting, we used all genes within the curated 622-gene cfRNA panel without additional HVG filtering. Expression matrices were standardized using z-score normalization prior to modeling. All downstream analyses, including attribution and spatial statistics, were performed separately for the HVG and signature feature regimes.

Platelet RNA sequencing in solid tumors (GSE68086).

To assess performance on bulk transcriptomic data derived from tumor-educated platelets, we used the GSE68086 dataset[51]. This cohort comprises platelet RNA-seq profiles from patients with multiple cancer types and healthy controls. Samples include breast cancer (BRCA), colorectal cancer (CRC), glioblastoma (GBM), lung cancer, pancreatic cancer, and healthy individuals. Raw gene expression matrices were obtained from the Gene Expression Omnibus (GEO). After alignment with series matrix metadata and removal of zero-variance genes, the top 5,000 highly variable genes based on variance across samples were retained for modeling. Expression values were standardized prior to modeling. Two tasks were defined: (i) binary cancer versus healthy classification and (ii) five-class tumor type prediction (Breast, CRC, GBM, Lung, Pancreas).

Multi-cohort liquid-biopsy analysis.

We evaluated Dynomap across ten retrospective cancer-versus-healthy-control tasks constructed from RARE-Seq plasma cfRNA[40], GSE174302 plasma cfRNA[66,97], GSE216561 plasma and EV RNA[67,98], and GSE68086 platelet RNA[51,52]. Lung-cancer analyses included RARE-Seq Stanford (179 libraries from 165 donors; 99 cancer and 66 control donors), RARE-Seq MGH (117 libraries from 102 donors; 52 cancer and 50 control donors), GSE174302 cfRNA (35 cancer and 19 control donors), GSE216561 plasma RNA (19 cancer and 19 control donors), GSE216561 EV RNA (19 cancer and 14 control donors), and GSE68086 platelet RNA (59 cancer and 54 control donors). Colorectal-cancer analyses included GSE174302 cfRNA (54 cancer and 19 control donors), GSE216561 plasma RNA (23 cancer and 19 control donors), GSE216561 EV RNA (19 cancer and 14 control donors), and GSE68086 platelet RNA (44 cancer and 54 control donors).

Each dataset, RNA carrier, and cancer contrast was modeled independently. Expression matrices were transformed using $\log[1+10^6 x_{ij}/(\sum_g x_{ig})]$. Samples were partitioned using five-fold stratified group cross-validation, with all repeated libraries from a donor assigned to the same fold. Within each Dynomap outer-training fold, donors were separated into group-disjoint inner-training and validation partitions. Genes detected in at least 5% of inner-training samples and having non-zero variance were ranked by variance, and the top 1,000 genes were retained. Standardization parameters were estimated exclusively from the inner-training partition.

For these additional experiments, we used a separate PyTorch implementation of the Dynomap architecture. A new model was initialized for every dataset and fold and trained for 75 epochs using batch size 16, Adam optimization with learning rate 10^{-3}, and 64×64 rendered feature maps. The checkpoint achieving the highest inner-validation accuracy was retained. No trained weights, feature sets, decision thresholds, or spatial coordinates were transferred between datasets. Pooled donor-level out-of-fold AUROC was used as the primary performance summary. Mean and standard deviation of the five fold-level AUROCs were used for visualization.

For descriptive comparison, we trained class-balanced logistic-regression models using 1,000 HVGs selected exclusively from each outer-training fold. A QC-only logistic-regression model was evaluated using three library-level covariates: $\log(1+\text{library sum})$, $\log(1+\text{number of detected genes})$, and $\log(1+\text{maximum observed expression})$.

TCGA breast cancer (TCGA-BRCA).

We analyzed RNA-seq expression data from The Cancer Genome Atlas (TCGA) breast cancer cohort[63]. Gene expression and clinical annotations were obtained from the UCSC Xena repository[99]. Expression values correspond to HiSeqV2 RNA-seq profiles and were aligned to clinical PAM50 subtype annotations[71]. Two classification tasks were defined: (i) binary aggressive versus non-aggressive tumor classification, where Basal and HER2-enriched tumors were labeled as aggressive, and (ii) multiclass PAM50 subtype prediction (LumA, LumB, Basal, HER2, Normal-like). From the aligned expression matrix, we selected the top 4,000 highly variable genes based on variance across samples. Expression values were standardized prior to modeling. No additional feature engineering was performed.

Parkinson's disease voice measurements.

To evaluate generalization beyond transcriptomics, we used the Parkinson's disease voice dataset originally described by Sakar *et al.*[70]. This dataset contains engineered acoustic features extracted from sustained phonations of individuals with Parkinson's disease and healthy controls. Features include jitter, shimmer, harmonicity, nonlinear dynamical metrics, and wavelet-derived descriptors. We used the dataset in its published tabular form without feature selection or dimensionality reduction. All features were standardized prior to modeling. The task was binary classification of Parkinson's disease versus control.

Single-cell transcriptomics (Tabula Muris).

To evaluate Dynomap in a large multicellular context, we used the Tabula Muris single-cell RNA-seq atlas[81]. The dataset comprises 54,865 cells across 55 annotated cell types collected from 20 murine organs. We utilized a preprocessed expression matrix and selected the top 1,089 highly variable genes based on variance across cells to represent core transcriptional programs. Expression values were standardized prior to modeling. The task was multiclass cell-type classification across all annotated populations.

Additional benchmark datasets.

To assess generalization beyond the five focused biomedical cohorts, we evaluated Dynomap on 13 additional publicly available tabular datasets covering microarray gene expression, mass spectrometry, and engineered sensor measurements. Nine of these are pairwise microarray cancer classification datasets obtained from the OpenML repository[88,89]: AP Breast vs Lung (ID 1150), AP Breast vs Omentum (ID 1127), AP Colon vs Lung (ID 1126), AP Endometrium vs Breast (ID 1123), AP Endometrium vs Omentum (ID 1151), AP Endometrium vs Prostate (ID 1141), AP Endometrium vs Uterus (ID 1164), AP Omentum vs Uterus (ID 1124), and AP Breast vs Kidney (ID 1158). Each task distinguishes primary tumors of one tissue origin from another. Cohort sizes range from 130 to 604 samples, and all share 10,935 gene-level features. The OVA Ovary dataset (OpenML ID 1166) is drawn from the same microarray collection and contains 1,545 samples with 10,935 features, in which ovarian tumors are distinguished from all other tissue types. This dataset is heavily imbalanced toward the negative class and is included as a stress test for class-balanced performance. Two further datasets are versions of the Arcene mass spectrometry cancer benchmark released in the NIPS 2003 feature selection challenge (OpenML IDs 1458 and 41157), containing 200 and 100 samples respectively across approximately 10,000 mass spectrum features[90]. The thirteenth dataset is the UCI Human Activity Recognition Using Smartphones benchmark, comprising 10,299 samples with 561 engineered inertial sensor features across six activity classes[10]. This dataset was selected to test whether the framework generalizes to non-biomedical multiclass settings with structurally different feature semantics. For each dataset, feature matrices were standardized prior to modeling. For tasks with more than 4,000 features, we retained the 4,000 highest-variance features. The UCI HAR dataset was used with all 561 features. All datasets were evaluated using five-fold stratified cross-validation, and identical Dynomap hyperparameters were applied across all datasets to ensure comparability.

Data availability

All datasets analyzed in this study are publicly available from established repositories. The circulating cell-free RNA (RARE-Seq) cohort was obtained from the multi-cancer liquid biopsy study reported by Nesselbush *et al.*[40]. Tumor-educated platelet RNA-sequencing data (GSE68086) were obtained from the Gene Expression Omnibus (GEO) under accession number GSE68086[51]. Breast cancer RNA-sequencing data (TCGA-BRCA) and associated clinical annotations were obtained from The Cancer Genome Atlas (TCGA)[63] via the UCSC Xena repository[99]. The Parkinson's disease voice dataset was obtained from the publicly released dataset described by Sakar *et al.*[70]. The Tabula Muris single-cell RNA-sequencing atlas was obtained from the Tabula Muris consortium[81]. Additional plasma cfRNA data were obtained from GEO under accession GSE174302[66,97], and paired plasma cfRNA and extracellular-vesicle RNA data were obtained under accession GSE216561[67,98].

All processed expression matrices, feature-selection outputs, and intermediate data files required to reproduce the analyses reported in this study are provided within the associated Code Ocean capsule prepared for peer review. Detailed preprocessing procedures and model training protocols are described in the Methods. No new human or animal data were generated for this study.

Code availability

All code used to perform the analyses in this study has been deposited in a Code Ocean capsule and shared with the editors and reviewers as part of the peer-review process. The capsule contains the complete implementation of the Dynomap framework, including data preprocessing, image construction, model training, and evaluation pipelines.

Upon acceptance of the manuscript, the Code Ocean capsule will be made publicly available with a persistent digital object identifier (DOI) to enable full reproducibility of the results.

References

1. Hasin, Y., Seldin, M. & Lusis, A. J. Multi-omics approaches to disease. *Genome Biol.* **18**, 1–15 (2017).
2. Libbrecht, M. W. & Noble, W. S. Machine learning applications in genetics and genomics. *Nat. Rev. Genet.* **16**, 321–332 (2015).
3. Luecken, M. D. & Theis, F. J. Current best practices in single-cell rna-seq analysis: a tutorial. *Mol. Syst. Biol.* **15**, e8746, DOI: 10.15252/msb.20188746 (2019).
4. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347–355, DOI: 10.1038/nature19949 (2016).
5. Patti, G. J., Yanes, O. & Siuzdak, G. Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell Biol.* **13**, 263–269, DOI: 10.1038/nrm3314 (2012).
6. Heitzer, E., Haque, I. S., Roberts, C. E. S. & Speicher, M. R. Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nat. Rev. Genet.* **20**, 71–88 (2019).
7. Wan, J. C. M. *et al.* Liquid biopsies come of age: towards implementation of circulating tumour dna. *Nat. Rev. Cancer* **17**, 223–238 (2017).
8. Johnson, A. E. W. *et al.* Mimic-iii, a freely accessible critical care database. *Sci. Data* **3**, 160035, DOI: 10.1038/sdata.2016.35 (2016).
9. Lambin, P. *et al.* Radiomics: extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **48**, 441–446, DOI: 10.1016/j.ejca.2011.11.036 (2012).
10. Anguita, D., Ghio, A., Oneto, L., Parra, X. & Reyes-Ortiz, J. L. A public domain dataset for human activity recognition using smartphones. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 437–442 (2013).
11. Gorishniy, Y., Rubachev, I., Khrulkov, V. & Babenko, A. Revisiting deep learning models for tabular data. In *Advances in Neural Information Processing Systems*, vol. 34, 18932–18943 (2021).
12. Grinsztajn, L., Oyallon, E. & Varoquaux, G. Why do tree-based models still outperform deep learning on typical tabular data? In *Advances in Neural Information Processing Systems*, vol. 35, 507–520 (2022).
13. Shavitt, I. & Segal, E. Regularization learning networks: Deep learning for tabular datasets. In *Advances in Neural Information Processing Systems*, vol. 31 (2018).
14. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
15. Goyal, A. & Bengio, Y. Inductive biases for deep learning of higher-level cognition. *Proc. Royal Soc. A* **478**, 20210068 (2022).
16. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).
17. Marusyk, A. & Polyak, K. Tumor heterogeneity: causes and consequences. *Biochimica et Biophys. Acta - Rev. on Cancer* **1805**, 105–117 (2010).
18. Hosmer, D. W., Lemeshow, S. & Sturdivant, R. X. *Applied Logistic Regression* (Wiley, 2013), 3 edn.
19. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning* (Springer, 2009), 2 edn.
20. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
21. Freund, Y. & Schapire, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **55**, 119–139 (1997).
22. Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794 (2016).
23. Ke, G. *et al.* LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, vol. 30, 3146–3154 (2017).
24. Cover, T. M. & Hart, P. E. Nearest neighbor pattern classification. *IEEE Transactions on Inf. Theory* **13**, 21–27 (1967).
25. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).
26. Kadra, A., Lindauer, M., Hutter, F. & Grabocka, J. Well-tuned simple nets excel on tabular datasets. In *Advances in Neural Information Processing Systems*, vol. 34, 23928–23941 (2021).

27. Gorishniy, Y., Kotelnikov, A. & Babenko, A. TabM: Advancing tabular deep learning with parameter-efficient ensembling. In *The Thirteenth International Conference on Learning Representations* (2025).

28. Ye, H.-J., Yin, H.-H., Zhan, D.-C. & Chao, W.-L. Revisiting nearest neighbor for tabular data: A deep tabular baseline two decades later. In *The Thirteenth International Conference on Learning Representations* (2025).

29. Hollmann, N., Müller, S., Eggensperger, K. & Hutter, F. TabPFN: A transformer that solves small tabular classification problems in a second. In *The Eleventh International Conference on Learning Representations* (2023).

30. Sharma, A., Vans, E., Shigemizu, D., Boroevich, K. A. & Tsunoda, T. DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architecture. *Sci. Reports* **9**, 11399, DOI: 10.1038/s41598-019-47765-6 (2019).

31. Zhu, Y. *et al.* Converting tabular data into images for deep learning with convolutional neural networks. *Sci. reports* **11**, 11325 (2021).

32. Alenizy, H. A. & Berri, J. Transforming tabular data into images via enhanced spatial relationships for cnn processing. *Sci. Reports* **15**, 17004 (2025).

33. Islam, M. T. & Xing, L. Cartography of genomic interactions enables deep analysis of single-cell expression data. *Nat. Commun.* **14**, 679 (2023).

34. Bazgir, O. *et al.* Representation of features as images with neighborhood dependencies for compatibility with convolutional neural networks. *Nat. communications* **11**, 4391 (2020).

35. Bengio, Y., Courville, A. & Vincent, P. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis Mach. Intell.* **35**, 1798–1828 (2013).

36. Esteva, A. *et al.* A guide to deep learning in healthcare. *Nat. Medicine* **27**, 166–176 (2021).

37. Dosovitskiy, A., Beyer, L., Kolesnikov, A. *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations* (2021).

38. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).

39. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).

40. Nesselbush, M. C., Luca, B. A., Jeon, Y.-J. *et al.* An ultrasensitive method for detection of cell-free RNA. *Nature* **641**, 759–768, DOI: 10.1038/s41586-025-08834-1 (2025).

41. Sundararajan, M., Taly, A. & Yan, Q. Axiomatic attribution for deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, vol. 70 of *Proceedings of Machine Learning Research*, 3319–3328 (PMLR, 2017).

42. The Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. *Cell* **163**, 1011–1025, DOI: 10.1016/j.cell.2015.10.025 (2015).

43. The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550, DOI: 10.1038/nature13385 (2014).

44. The Cancer Genome Atlas Research Network. Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell* **32**, 185–203.e13, DOI: 10.1016/j.ccell.2017.07.007 (2017).

45. Chen, H., Bian, A., Yang, L.-f. *et al.* Targeting STAT3 by a small molecule suppresses pancreatic cancer progression. *Oncogene* **40**, 1440–1457, DOI: 10.1038/s41388-020-01626-z (2021).

46. Giatromanolaki, A., Harris, A. L., Banham, A. H. *et al.* Carbonic anhydrase 9 (CA9) expression in non-small-cell lung cancer: correlation with regulatory FOXP3+ t-cell tumour stroma infiltration. *Br. J. Cancer* **122**, 1205–1210, DOI: 10.1038/s41416-020-0756-3 (2020).

47. Ramakrishnan, G., Parajuli, P., Singh, P. *et al.* NF1 loss of function as an alternative initiating event in pancreatic ductal adenocarcinoma. *Cell Reports* **41**, 111623, DOI: 10.1016/j.celrep.2022.111623 (2022).

48. Tsanov, K. M. *et al.* SMAD4 induces opposite effects on metastatic growth from pancreatic tumors depending on the organ of residence. *Nat. Cancer* **6**, 1839–1856, DOI: 10.1038/s43018-025-01047-5 (2025).

49. Bilodeau, C., Shojaie, S., Goltsis, O. *et al.* TP63 basal cells are indispensable during endoderm differentiation into proximal airway cells on acellular lung scaffolds. *npj Regen. Medicine* **6**, 12, DOI: 10.1038/s41536-021-00124-4 (2021).

50. Panda, H. *et al.* NRF2 immunobiology in cancer: implications for immunotherapy and therapeutic targeting. *Oncogene* **44**, 3641–3651, DOI: 10.1038/s41388-025-03560-4 (2025).

51. Best, M. G. *et al.* RNA-Seq of tumor-educated platelets enables blood-based pan-cancer, multiclass, and molecular pathway cancer diagnostics. *Cancer Cell* **28**, 666–676, DOI: 10.1016/j.ccell.2015.09.018 (2015).

52. Gene Expression Omnibus. Gene expression omnibus series GSE68086. NCBI Gene Expression Omnibus (2015). Accession GSE68086.

53. Moran, P. A. P. Notes on continuous stochastic phenomena. *Biometrika* **37**, 17–23 (1950).

54. Anselin, L. Local indicators of spatial association—lisa. *Geogr. Analysis* **27**, 93–115 (1995).

55. Jose, A., Shenoy, G. G., Rodrigues, G. S. *et al.* Histone demethylase KDM5B as a therapeutic target for cancer therapy. *Cancers* **12**, 2121, DOI: 10.3390/cancers12082121 (2020).

56. MED12L mediator complex subunit 12l [homo sapiens (human)]. NCBI Gene (2026). Gene ID: 116931. Accessed 2026-08-07.

57. Wohlrab, H., Signoretti, S., Rameh, L. E., DeConti, D. K. & Hansen, S. H. Mitochondrial transporter expression patterns distinguish tumor from normal tissue and identify cancer subtypes with different survival and metabolism. *Sci. Reports* **12**, 17035, DOI: 10.1038/s41598-022-21411-0 (2022).

58. Xu, X., Lu, X., Chen, L., Peng, K. & Ji, F. Downregulation of MMP1 functions in preventing perineural invasion of pancreatic cancer through blocking the NT-3/TrkC signaling pathway. *J. Clin. Lab. Analysis* **36**, e24719, DOI: 10.1002/jcla.24719 (2022).

59. Taguchi, T., Kurata, M., Onishi, I. *et al.* SECISBP2 is a novel prognostic predictor that regulates selenoproteins in diffuse large b-cell lymphoma. *Lab. Investig.* **101**, 218–227, DOI: 10.1038/s41374-020-00495-0 (2021).

60. Penninkhof, F., Grootegoed, J. A. & Blok, L. J. Identification of REPS2 as a putative modulator of NF-kappaB activity in prostate cancer cells. *Oncogene* **23**, 5607–5615, DOI: 10.1038/sj.onc.1207750 (2004).

61. Wang, H.-M., Xu, Y.-F., Ning, S.-L. *et al.* The catalytic region and PEST domain of PTPN18 distinctly regulate the HER2 phosphorylation and ubiquitination barcodes. *Cell Res.* **24**, 1067–1090, DOI: 10.1038/cr.2014.99 (2014).

62. Xu, C.-M. *et al.* Multifunctional neuron-specific enolase: its role in lung diseases. *Biosci. Reports* **39**, BSR20192732, DOI: 10.1042/BSR20192732 (2019).

63. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70, DOI: 10.1038/nature11412 (2012).

64. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).

65. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061–1068 (2008).

66. Chen, S., Jin, Y., Wang, S. *et al.* Cancer type classification using plasma cell-free rnas derived from human and microbes. *eLife* **11**, e75181, DOI: 10.7554/eLife.75181 (2022).

67. Wang, H., Zhan, Q., Ning, M. *et al.* Depletion-assisted multiplexed cell-free rna sequencing reveals distinct human and microbial signatures in plasma versus extracellular vesicles. *Clin. Transl. Medicine* **14**, e1760, DOI: 10.1002/ctm2.1760 (2024).

68. Little, M. A., McSharry, P. E., Hunter, E. J., Spielman, J. & Ramig, L. O. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *BioMedical Eng. OnLine* **6**, 23 (2007).

69. Tsanas, A., Little, M. A., McSharry, P. E. & Ramig, L. O. Novel speech signal processing algorithms for high-accuracy classification of parkinson's disease. *IEEE Transactions on Biomed. Eng.* **59**, 1264–1271 (2012).

70. Sakar, C. O. *et al.* A comparative analysis of speech signal processing algorithms for parkinson's disease classification and the use of the tunable Q-factor wavelet transform. *Appl. Soft Comput.* **74**, 255–263, DOI: 10.1016/j.asoc.2018.10.022 (2019).

71. Parker, J. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).

72. Pernas, S. & Tolaney, S. M. Molecular subtypes of breast cancer: current knowledge and clinical implications. *Clin. Adv. Hematol. & Oncol.* **16**, 535–545 (2018).

73. MS4A2 membrane spanning 4-domains a2 [homo sapiens (human)]. NCBI Gene (2026). Gene ID: 2206. Accessed 2026-02-27.

74. ACADSB acyl-coa dehydrogenase short/branched chain [homo sapiens (human)]. NCBI Gene (2026). Gene ID: 36. Accessed 2026-02-27.

75. Dai, X., Holt, M. V., McBride, A. *et al.* Human C6orf211 encodes Armt1, a protein carboxyl methyltransferase that targets PCNA and is linked to the DNA damage response. *Cell Reports* **10**, 1288–1296, DOI: 10.1016/j.celrep.2015.01.053 (2015).

76. KCNQ5 potassium voltage-gated channel subfamily q member 5 [homo sapiens (human)]. NCBI Gene (2026). Gene ID: 56479. Accessed 2026-02-27.

77. DLGAP1 dlg associated protein 1 [homo sapiens (human)]. NCBI Gene (2026). Gene ID: 9229. Accessed 2026-02-27.

78. Diana, P. & Carvalheira, G. M. G. NIBAN1, exploring its roles in cell survival under stress context. *Front. Cell Dev. Biol.* **10**, 867003, DOI: 10.3389/fcell.2022.867003 (2022).

79. SCAND3 scan domain containing 3 [homo sapiens (human)]. NCBI Gene (2026). Gene ID: 114821. Accessed 2026-02-27.

80. Mu, N., Liu, F., Song, X. *et al.* HHIPL2 positively governs hedgehog signaling to accelerate non-small cell lung cancer progression via enhancing HNRNPC-mediated HNF1A mrna stabilization. *Cell Death & Dis.* **17**, 103, DOI: 10.1038/s41419-025-08331-3 (2025).

81. Tabula Muris Consortium. Single-cell transcriptomics of 20 mouse organs creates a tabula muris. *Nature* **562**, 367–372, DOI: 10.1038/s41586-018-0590-4 (2018).

82. Love, P. E. & Hayes, S. M. ITAM-mediated signaling by the t-cell antigen receptor. *Cold Spring Harb. Perspectives Biol.* **2**, a002485, DOI: 10.1101/cshperspect.a002485 (2010).

83. Reth, M. Antigen receptor tail clue. *Nature* **338**, 383–384, DOI: 10.1038/338383b0 (1989).

84. Neefjes, J., Jongsma, M. L. M., Paul, P. & Bakke, O. Towards a systems understanding of mhc class i and mhc class ii antigen presentation. *Nat. Rev. Immunol.* **11**, 823–836 (2011).

85. Ricard-Blum, S. The collagen family. *Cold Spring Harb. Perspectives Biol.* **3**, a004978 (2011).

86. Moll, R., Divo, M. & Langbein, L. The human keratins: biology and pathology. *Histochem. Cell Biol.* **129**, 705–733 (2008).

87. Zheng, G. X. Y., Terry, J. M., Belgrader, P. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049, DOI: 10.1038/ncomms14049 (2017).

88. Vanschoren, J., van Rijn, J. N., Bischl, B. & Torgo, L. Openml: networked science in machine learning. *SIGKDD Explor. Newsl.* **15**, 49–60 (2014).

89. de Souto, M. C. P., Costa, I. G., de Araujo, D. S. A., Ludermir, T. B. & Schliep, A. Clustering cancer gene expression data: a comparative study. *BMC Bioinforma.* **9**, 497 (2008).

90. Guyon, I., Gunn, S., Ben-Hur, A. & Dror, G. Result analysis of the nips 2003 feature selection challenge. In *Advances in Neural Information Processing Systems*, vol. 17 (2005).

91. Saito, T. & Rehmsmeier, M. The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE* **10**, e0118432 (2015).

92. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **1**, 206–215 (2019).

93. Nair, V. & Hinton, G. E. Rectified linear units improve restricted boltzmann machines. In *ICML* (2010).

94. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).

95. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the inception architecture for computer vision. In *CVPR* (2016).

96. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. In *ICLR* (2015).

97. Gene Expression Omnibus. Gene expression omnibus series gse174302 (2022). NCBI GEO accession GSE174302.

98. Gene Expression Omnibus. Gene expression omnibus series gse216561 (2024). NCBI GEO accession GSE216561.

99. Goldman, M. J. *et al.* Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **38**, 675–678, DOI: 10.1038/s41587-020-0546-8 (2020).