Neighbor-Embedded Block Imagification for Single-Cell Gene Expression Classification

Ali Anaissi[1,2] and Seid Miad Zandavi[1,3]

[1] University of Sydney, Australia
[2] University of Technology Sydney, Australia
[3] Broad Institute of MIT and Harvard, USA
ali.anaissi@sydney.edu.au, szandavi@broadinstitute.org

Abstract. The innovative concept of "block imagification" transforms high - dimensional molecular measurements into two-dimensional RGB images, maintaining a one-to-one relationship with each sample while enabling associations across different classes. This transformation from non-image data to image format creates a holistic molecular representation of a sample, enhancing phenotype classification through advanced computer vision techniques.
Each transformed RGB image encodes molecular information in a 2D blocked diagonal neighbor-embedded space, where the RGB channels represent molecular abundance and gene intensity. The proposed method was applied to single cell RNA sequencing data (scRNA-seq) to "imagify" the gene expression profiles of individual cells.
The results demonstrate that even a simple convolutional neural network (CNN) trained on these transcriptomics images can accurately classify diverse cell types, outperforming traditional scRNA sequence classifiers such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Random Forest (RF).

Keywords: Block Imagification, Single-cell RNA Sequencing (scRNA-seq), Transcriptomics Image Representation Convolutional Neural Networks (CNN), Phenotype / Cell Type Classification

Introduction

Omics refers to a set of innovative and comprehensive strategies used to analyze the complete molecular profile of humans and other species. The term encompasses a growing number of disciplines in molecular biology ending with "-omics," such as genomics, transcriptomics, proteomics, metabolomics, and metagenomics. For example, transcriptomics involves analyzing the complete set of messenger RNA (mRNA) molecules within a cell or tissue, indicating the abundance or concentration of each RNA molecule.

The main objective of omics is to quantify a wide range of biological molecules in a non-targeted and unbiased manner to characterize the structure, function, evolution, and dynamics of an organism or population. This enables researchers to explore how gene or molecular compositions and their associations influence phenotypes [1].

With the advent of high-throughput profiling technologies, such as next generation sequencing (NGS) [2] and high-throughput mass spectrometry screening platforms [1],

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

biology has become increasingly dependent on large-scale molecular data generated at multiple levels, from the genome to the metabolome. The availability of omics data is revolutionizing life sciences by enabling a holistic understanding of physiological processes and supporting more accurate phenotype predictions. However, despite these advances, distinguishing phenotypes (e.g., diseases or cellular subtypes) remains challenging due to the high-dimensional, heterogeneous, and complex nature of omics data.

RNA sequencing (RNA-seq) is an advanced NGS technology used to detect and quantify gene expression, revealing transcript abundance within a biological sample [3]. Understanding the transcriptome is essential for decoding cellular behavior, discovering genome functionality, and identifying molecular signatures of cells and tissues [3]. Unlike bulk RNA-seq, which captures average expression across thousands of cells, single-cell RNA sequencing (scRNA-seq) profiles gene expression at the individual cell level, providing high-resolution insights into cellular heterogeneity.

The development of scRNA-seq has transformed gene expression studies, from cataloguing cell types [4] to uncovering transcriptional regulation [5] [6] [7]. It enables the classification of distinct cell states and phenotypes [8], contributing to precision medicine through better molecular understanding of patient states. Although complete RNA information for each cell is often unavailable, machine learning techniques can be used to identify gene expression patterns that reveal previously unknown cell types. However, the data complexity, sometimes requiring more than 20,000 features per cell, adds significant analytical challenges due to high dimensionality and potential feature interactions.

In typical scRNA-seq datasets, the number of features (genes per cell, often 10,000 – 60,000) far exceeds the number of samples (cells, typically 100–8,000). This imbalance is known as the "curse of dimensionality" [9], which often forces researchers to apply dimensionality reduction techniques. Unfortunately, such methods can lead to information loss, especially for cell-type identification. For instance, models like the Cox proportional hazards model [10], elastic net [11], and random forests [12] have been applied to survival prediction, but often rely on reduced feature sets. Neural networks (NNs) have also been explored for predictive modeling [13], but they typically underperform on right-censored survival data compared to traditional methods [14].

The emergence of deep learning, particularly convolutional neural networks (CNNs), has achieved remarkable success in computer vision, offering a powerful means of interpreting high-dimensional data. Computer vision involves stages such as image acquisition, encoding, and interpretation, making it well-suited for extracting patterns from structured data like images. CNNs excel when applied to image data, even with limited training samples, and can outperform other models in terms of accuracy and efficiency.

Despite the impressive success of CNNs in processing image data, applying CNNs directly to high-dimensional non-image data, such as gene expression profiles, remains challenging. CNNs are specifically designed to work on two-dimensional matrices (e.g., images), leveraging spatial coherence and local dependencies among neighboring pixels to extract meaningful features. In contrast, gene expression features (individual genes) are typically unordered and lack an inherent spatial structure, making the direct application of CNNs ineffective. To overcome this, converting non-image data into image-formatted

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

data enables CNNs to exploit their spatial feature extraction capabilities effectively. Such a transformation not only enhances the interpretability of complex biological data like scRNA-seq but also allows CNNs to achieve strong classification performance with relatively less training data and computational resources compared to other machine learning approaches.

The key lies in reorganizing gene expression data into image structures that mimic the spatial dependencies CNNs are designed to exploit. To achieve this, we introduce the innovative concept of *block imagification*, which transforms high-dimensional molecular measurements into two-dimensional RGB images, maintaining a one-to-one correspondence with each sample while preserving biological relevance.

Each transformed image encodes molecular information in a 2D blocked diagonal neighbor-embedded space, where the RGB channels represent molecular abundance and gene intensity. This novel representation enables the effective application of CNNs to complex, high-dimensional biological datasets such as scRNA-seq, enhancing phenotype classification through powerful image-based learning.

Related work

Transforming non-image machine learning (ML) problems into the image recognition problems provides an opportunity to solve them using powerful deep learning algorithms. Converting a vector into the 2D image data was proposed in [15] as the first preliminary approach to using the relative positions of the genes in a chromosome to classify tumor type; i.e., by using transcription locus as the rearrangement criterion. The *OmicsMapNet* approach analyzes RNA sequence expression data from diffuse glioma samples of TCGA by organizing the functions and ontology relationships into spatially adjacent and patterned locations [16]. *OmicsMapNet* employes a trained CNN model to predict the malignancy grade of tumor samples. Single-cell image segmentation and analysis techniques are frequently employed to convert cell images into detailed quantitative descriptors. This method facilitates scalable and reliable interpretation and allows for the application of representation learning techniques, which are usually limited to numerical matrix data in omics studies [17, 18]. Interestingly, although image data are often converted into quantitative formats for analysis, "imagification" have begun to leverage advanced/powerful image modeling methods, for the analysis of omics data [15] [19].

The *CPC-R algorithm* [20] has been proposed to convert non-image data into images by visualizing non-image data through pair values mapping. The general idea is to group numerical attributes in pairs (x_i, x_{i+1}), where are mapped to a 2D plane corresponding to the horizontal (x_i) and vertical (x_{i+1}) coordinates. Then, cells are labeled with various intensities of a greyscale from black (x_1, x_2) to extreme light grey (x_{n-1}, x_n). Additionally, the visualization of genomics data via the Hilbert curve was proposed by [21]. The Hilbert curve is a continuous fractal space-filling curve that is an effective technique for the linear mapping of multidimensional space. The linear space adequately preserves the locality of features in multidimensional space [22], so the clustering properties of the Hilbert space-filling curve suggest it has potential to deal with data mapping between dimensions.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

DeepInsight [23] introduced a domain-agnostic method for converting various types of non-image data, including text, audio, and synthetic datasets, into image form for CNN processing. Unlike previous methods, DeepInsight does not rely on domain-specific knowledge to rearrange features. More recently, Fotomics [19] applied the Fourier Transform to map scRNA-seq data into 2D coordinates in a neighbor space. However, it does not account for cluster-specific variability, which may affect classification performance. A biologically informed approach was proposed by [24], where genes are rearranged using molecular function hierarchies. This method uses the KEGG ontology [25] and BRITE hierarchies to map genes into a 2D layout, constructing structured gene-expression images for cancer survival prediction. Despite the variety of existing approaches, most methods generate grayscale images and often overlook the intensity or domain-specific distribution inherent in omics data.

In this paper, we propose a novel image transformation method named Block Imagification to address cell-type identification in scRNA-seq data through a classification framework. Our method transforms gene-expression vectors into a 2D block diagonal neighbor-embedded space represented as RGB images, assuming cluster independence among gene groups. The RGB channels capture both gene expression levels and intensities for each cell, encoded via the Discrete Fourier Transform (DFT). This structured representation enables CNNs to better extract local features and enhance classification performance.

Problem Formulation

Depending on the underlying technology, different omics profiles can encompass measurements of several hundreds to several thousands of molecules. We define the concept of *block imagification* as the process of transforming a numerical vector representing these measurements into an image that maintains a one-to-one correspondence with the original sample, independently clustered. This resulting image is then used to classify the sample phenotype via automated image recognition techniques.

In general, an image consists of spatially coherent pixels, where meaningful information is locally shared among neighboring pixels. If pixels are arranged arbitrarily, their placement can adversely affect feature extraction and classification accuracy. Therefore, when converting non-image data into image form, maintaining spatial coherence of pixels within local regions is crucial to improve classification performance. Additionally, color images typically have three channels per pixel that encode intensity and chrominance information. The imagification process should capture the relative relationships among different molecules in a spatial coordinate system, including their intensity and chrominance attributes.

Accordingly, we propose a method that transforms a numerical vector (representing gene abundances in a sample) into a Cartesian coordinate plane using t-Distributed Stochastic Neighbor Embedding (t-SNE) [26] as a nonlinear dimensionality reduction technique applied across samples. Furthermore, we apply a discrete Fourier transform (DFT) on a finite number of points to extract frequency-domain features representing individual behavior. This approach enables capturing all possible associations across different classes.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

This transformation locates features (i.e., genes) in the Cartesian plane, represented as a colorful 2D image. The pixel colors are determined by gene expression values as the first channel, while the real and imaginary components of the DFT form the second and third channels, respectively. Thus, the resulting color image integrates multiple layers of information: gene expression intensity and frequency-domain characteristics within each sample. Each channel thereby encodes complementary information, facilitating comprehensive feature representation for downstream classification.

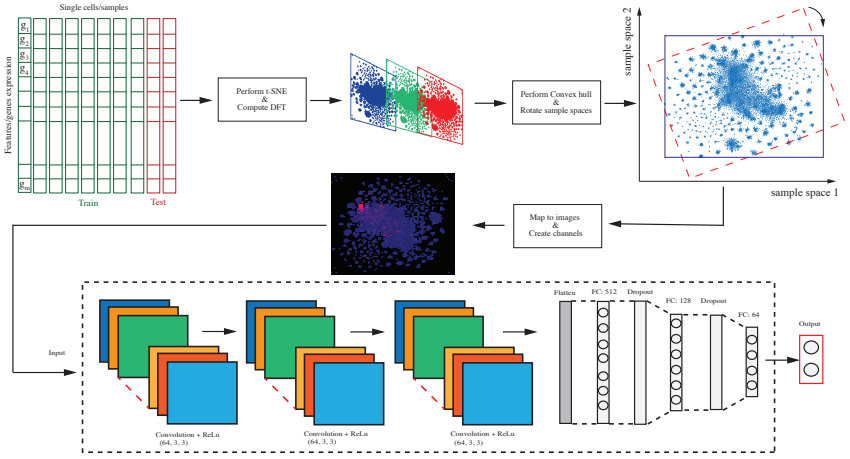

Fig. 1: Omics Imagification structure and CNN architecture for classification of scRNA-seq datasets

To begin with the transformation (see Fig. 1), the non-image scRNA-seq dataset is split into training and test subsets. Then, the training data is partitioned into distinct classes. Each class-wise partition is transformed into a reference Cartesian plane (sample space) using the t-SNE transformation. This transformation identifies the spatial locations of features (i.e., each gene) within the plane.

Once the locations of all features are determined in the feature matrix, the next step is to find the minimum convex set enclosing the feature matrix. The convex hull [27] is used to identify this minimum convex boundary. Since the image needs to be aligned horizontally or vertically for input to the CNN, a rotation step is applied, as formulated in Eq. 1.

$$\begin{bmatrix} X_{rp} \\ Y_{rp} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} X_p \\ Y_p \end{bmatrix} \tag{1}$$

where $[X_p, Y_p]^T$ is the sample spaces measured by t-SNE and $[X_{rp}, Y_{rp}]^T$ is the rotated X_pY_P-plane.

The Cartesian plane needs to be converted to spatially pixel forms to locate the genes/features at the right points. This is organized by pre-set parameters, such as the

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

pixel sizes in the X_{rp} and Y_{rp} directions as P_x and P_y, respectively. The image pixel size (P_x,P_y) is required to coordinate (X_{rp},Y_{rp}) into the pixel frame (x_p,y_p). This conversion is formulated in Eqs. 2 and 3.

$$x_p = round\left(1 + \frac{(X_{rp} - \min(X_{rp})) \times P_x}{\max(X_{rp}) - \min(X_{rp})}\right) \tag{2}$$

$$y_p = round\left(1 + \frac{(Y_{rp} - \min(Y_{rp})) \times P_y}{\max(Y_{rp}) - \min(Y_{rp})}\right) \tag{3}$$

The conversion to pixel frames is achieved using the average of the features (i.e., gene expression) in the same location, as the image size has a pixel limit. A single pixel in the pixel frame may consist of more than one feature for a sample x_k (for $k = 1, 2, ..., n$). When mapping these features into the pixel frame, the pixel intensity is the average value of the features with the same position. Consequently, if the image or grid size is very small, many of the features overlap and the image representation may not be accurate compared with the given number of features. Selecting an appropriate processing resolution depends on the amount of hardware resources available and the required features.

Since the class-wise Cartesian planes are set, the intensity and behavior of each sample must be extracted using DFT [28]. Let $\{f(t); t = 0, \pm1, \pm2, ...\}$ be considered as a deterministic, discrete sequence of data. We assume that $y(t) = f(t)$ has finite energy ($\sum_{-\infty}^{+\infty} |y(t)|^2 < \infty$). According to the finite energy assumption, the sequence data ($y(t)$) can be converted to DFT, as represented in Eq.4.

$$Y(\omega) = \sum_{-\infty}^{+\infty} y(t)e^{-i\omega t} \tag{4}$$

where ω is measured in cycles per sampling interval, which means the frequency of each feature per unit interval in the scRNA-seq. $i \equiv \sqrt{-1}$ is the basis for complex numbers. $Y(\omega)$, the DFT at frequency ω, is a measure of the amplitude and phase of a complex sinusoid, from which Euler's identity introduces $Y(\omega) \equiv \cos(\omega t) - i\sin(\omega t)$. In the complex plane of the DFT, the sinusoidal module shows the system's behavior (i.e., here, the system is defined as omic profile like scRNA-seq). Therefore, the system behavior is fundamentally dependent on the phase and amplitude, which are measured from the real and imaginary parts of the complex sinusoid.

In scRNA-seq data, the DFT of an scRNA-seq expression represents the behavior of each gene across its sample. Due to the nature of sequence data, only a positive form of a complex number is considered. The negative frequency is the rotation vector in the opposite direction to the positive frequency, so it can be removed from the analysis. In this regard, positive frequencies in the real and imaginary parts of DFT are mapped into the second and third channels of imagification to intensify the luminance in the transformed image. Finally for each sample, all the generated class-wise RGB images are encapsulated in a block diagonal form. This bock diagonal image is then resized to a 200×200 RGB image, which is used as the input to CNN model. Algorithm 1 shows the pseudo-code of the proposed imagificaiton approach for omics data such as scRNA-seq.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

Algorithm 1: Omics Imagification

Data: data set $\chi = \{x_1, x_2, ..., x_n\}$ with k classes $\zeta = \{c_1, c_2, ..., c_k\}$
data pre-processing;
Result: Converting numerical vector of Omics expression values to a 2D image
$\Upsilon = \{\upsilon_1, \upsilon_2, ..., \upsilon_n\}$.
begin
  set pixel size;
  Split data to train and test sets;
  Perform class-wise partitioning of the training data;
  for *each class-wise partition* **do**
    compute Cartesian coordinates;
      -$[X_p, Y_p] \longleftarrow$ **t-SNE**;
      -computing convex hull;
      -rotate convex hull plane;
  end
  for $x \in \chi$ **do**
    for $c \in \zeta$ **do**
      Compute RGB images across Cartesian coordinates of class c;
        -map features into rotated plane;
        -compute discrete Fourier transform (DFT);
        -positive DFT;
        -compute channel intensity with real and imaginary DFTs;
        -concatenate channels;
        -$I \longleftarrow$ scale converted 2D images;
    end
    create a block diagonal image from provided images
      $\upsilon =$ block.diag$\{I_1, I_2, ..., I_k\}$;
  end
end

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

Transformed RGB images convey more information for every single sample (i.e., cell-types) for training of the CNN model. Not only does imagification imply reorganization of the data to provide a structure with which the convolutional filters can exploit local knowledge patterns, but it also overcomes the curse of dimensionality inherent in genomics data. Training a CNN architecture or, more generally, a deep learning model, with a proper set of samples composed of a large number of features can allow the model to achieve significant performance. Correspondingly, the proposed image-form approach, omics imagification, provides a simple CNN model (see Fig. 1) with enough information to exploit local patterns in the corresponding sample.

Normalization is the key approach to preserving the topology of the functions. The minimum value of each function needs adjusting. For this reason, expression values across samples are normalized between 0 and 1 because we assume that the features are mutually independent, as a feature is normalized by its extrema values. The first channel, the expression value, is normalized as Eq. 5.

$$x(j,:) = \frac{x(j,:) - \min(x_{train})_j}{\max(x_{train})_j - \min(x_{train})_j} \tag{5}$$

where $j = 1, 2, ..., m$ represents the number of features (genes). Note that the normalization is processed over the training set so validation and test data are adjusted by training extrema. To concatenate the second and third channels of the RGB images, the DFT form of the gene expression over a sample (cell-type), $Y = [y_1, y_2, ..., y_n]$, is used in a logarithmic scale to position the feature values between 0 and 1. The logarithmic scale in a complex sinusoid is formulated as per Eq. 6.

$$y(:,k) = \frac{log(y(:,k) - \min(y(:,k)) + 1)}{\max(log(y(:,k) - \min(y(:,k)) + 1))} \tag{6}$$

where $k = 1, 2, ..., n$ is the number of samples (cell-types).

Data Collection

A total of six datasets were obtained from a curated publicly available single-cell RNA sequencing (scRNA-seq) collection maintained by the Hemberg Lab [29]. Each dataset includes labels for predicted cell types, which serve as the ground truth for evaluating classifier accuracy. The selected datasets encompass a wide variety of protocols, tissue types, and dataset sizes. They include different species (human and mouse), tissues (colorectal tumor, embryo, heart, colon, liver), and sequencing protocols (SMARTer, Smart-Seq2, Chromium).

In identifying different cell types, transcriptomic reads and unique molecular identifiers (UMIs) counts play an important role in compensating for variations in sequencing depth across cells. Accordingly, two different read types—read counts and UMIs—are considered in this study.

Specifically, human colorectal tumor and mouse embryo datasets were obtained from the following accession numbers: GEO GSE81861 (SMARTer), GEO GSE45719 (Smart-Seq2), and the Tabula Muris datasets sequenced with SMART-Seq2 and Chromium, as summarized in Table 1.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

	Dataset	Species	Tissue Type	Protocol	Genes	Cells	Classes	Read Type
1	li	Human	colorectal tumor	SMARTer	55186	561	9	Read
2	deng-reads	Mouse	Embryo	SMART-Seq2	22431	268	6	Read
3	TabulaMuris Heart 10X	Mouse	Heart	Chromium	23433	654	6	UMI
4	TabulaMuris Heart FACS	Mouse	Heart	SMART-Seq2	23433	7115	9	Read
5	TabulaMuris Colon FACS	Mouse	Colon	SMART-Seq2	23433	4149	6	Read
6	TabulaMuris Liver FACS	Mouse	Liver	SMART-Seq2	23433	981	6	Read

Table 1: Single Cell Data usded for Omics Imagification

Results and Discussion

The proposed method, omics imagification, is a transformation technique that rearranges one-dimensional data into two-dimensional RGB images. The method organizes spatially coherent pixels in sample spaces by applying a dimensionality reduction approach (e.g., t-SNE) along the sample dimension, combined with domain-intensity knowledge from the omics profile. In this study, six datasets provided by the Hemberg Lab [29] were used to evaluate the performance of machine learning methods after applying imagification to the data.

Data normalization is essential for scRNA-seq datasets due to various biases and noise introduced during the sequencing process. When expression quantification exhibits significant biases, adjusting expression measures is critical to reduce batch effects. According to the literature, a wide range of normalization methods have been developed, including those adapted from bulk RNA-seq or microarray data as well as novel approaches specifically designed for single-cell studies [30] [31]. However, many of these methods overlook key characteristics of scRNA-seq data, such as "zero inflation", the artefactual zero-read counts observed in several single-cell protocols (e.g., SMART-seq) [32] [33]. In this paper, both raw data and one of the most popular normalization methods, Counts Per Million (CPM) [34], are considered as global scaling preprocessing steps.

To illustrate the transformed raw data as RGB images, one sample from each dataset is shown in Fig. 2. This approach generates meaningful local structures by rearranging elements and captures feature dissimilarities through feature extraction and CNN classification. In this transformation, each pixel may have three possible intensity cases: a single gene (g_m) at position (x_{p_1}, y_{p_1}), multiple genes $(g_1, g_2, \ldots, g_j)$ at (x_{p_2}, y_{p_2}), or no genes/features at (x_{p_3}, y_{p_3}), where the pixel value is set to zero. The density representations of these images (see Fig. 2) clearly demonstrate the density of different cell types per pixel, providing a holistic view of the samples within each dataset.

To evaluate the performance of the proposed method, downstream analysis using supervised machine learning methods was used to predict the true labels with the corresponding cell populations. Three popular classifiers, Random Forest (RF), K-Nearest Neighbors (KNN), and Support Vector Machine (SVM)), were applied to six diverse datasets (Table 1). Additionally, a CNN model (Fig. 1) was trained with image-formed data.

The datasets were first divided into training (80% of all samples), validation (10% of the training set), and test (20% of all samples) sets. The training set was used for model fitting, and the validation set was employed to assess model fitness and select

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

Fig. 2: Illustration of sample imagification of databases and the density of scRNA-seqs per pixel

hyperparameters yielding the lowest validation error. The test set was strictly reserved for unbiased evaluation and was not used during model training or tuning. Classification accuracy on the test set was calculated as the percentage of correctly classified samples, using 5-fold cross-validation.

Figure 3 shows the performance of RF, KNN, and SVM to raw and CPM normalized data without imagification; whereas, the CNN model processed the image-formed data. The CNN architecture, shown in Fig. 1, consists of three blocks, each including a 2D convolutional layer, a ReLU activation layer, and a max-pooling layer. The output of the third block is flattened and passed through a fully connected layer. Dropout layers are incorporated to prevent overfitting during training. Finally, a sigmoid activation function produces the output class labels. The pixel frame size was set to $(P_x, P_y) = (100, 100)$. The method was implemented in Python on a computer with a 1.6 GHz dual-core Intel Core i5 processor and 16 GB of memory.

Experimental results demonstrate that the proposed imagification of cellular transcriptomics data combined with a simple CNN model can accuraltel classify different cell-types, achieving an average test accuracy of (94%) on raw data and (95%) on CPM-normalized data. In comparison, the average accuracies of the commonly used RF, KNN, and SVM classifiers on numerical data were 86%, 41%, and 46% for raw data, and 85%, 37%, and 39% for CPM data, respectively.

The boxplot in Fig. 3 illustrates that the CNN model using the imagification method outperforms the other methods and is robust to normalization techniques. Additionally, the comparatively narrow CNN boxplot indicates consistent high accuracy across diverse datasets and normalization methods. In contrast, KNN and SVM performed poorly in identifying cell types across datasets.

To further illustrate the performance of the CNN classifier, a confusion matrix is presented showing the classification results on a test dataset with known true labels. As shown in Fig. 4, the CNN model using the omics imagification method significantly outperformed the other classifiers on scRNA-seq data.

The proposed omics imagification method converts a $1D$ vector into a $2D$ RGB image format. Combined with a CNN model, this approach leverages local pattern feature learning from the omics image data. It can be extended to identify important

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

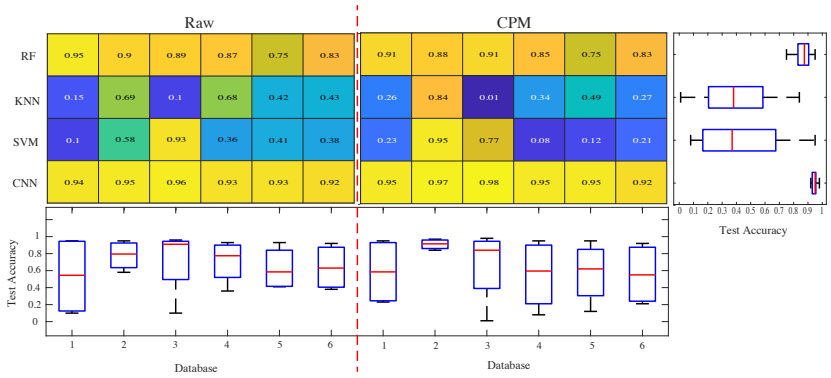

Fig. 3: Testing accuracy of three common machine learning models (RF, KNN, SVM) on the data matrix and of the CNN model on imagification across six different databases with omics data pre-processed in two different ways (raw data and counts per million, CPM)

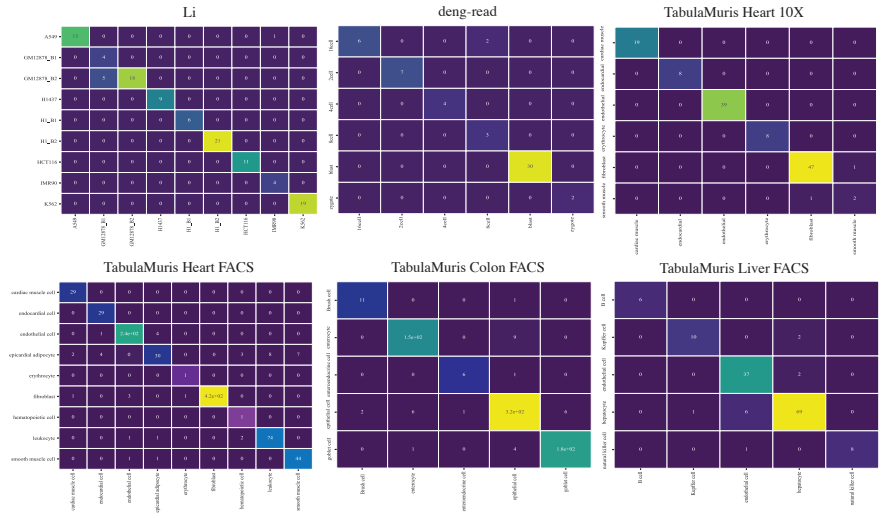

Fig. 4: Confusion matrix over transformed image-form test data for the CNN model

features (genes) as biomarkers across samples using image pattern recognition techniques. Additionally, different types of data (e.g., clinical and non-clinical) can be concatenated into images for integrated analysis and disease diagnosis. This technique is especially advantageous in applications where the original data is not naturally in image form.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

Conclusion

Depending on the underlying technology, various omics profiles can include measurements of hundreds to thousands of molecules. In this paper, we propose the term *omics imagification* to describe the process of converting a numerical measurement vector into an image with a one-to-one correspondence to the original sample. The proposed imagification approach transforms gene expression vectors into two-dimensional RGB images, enabling phenotype classification through automated image-recognition techniques. These image-formed data convey richer information about cell types to the classifier, including gene expression levels and intensities.

The method was evaluated on single-cell RNA sequencing (scRNA-seq) datasets, and experimental results demonstrate that a simple convolutional neural network (CNN) architecture significantly outperforms commonly used machine learning methods such as Random Forest (RF), K-Nearest Neighbors (KNN), and Support Vector Machines (SVM).

References

1. Konrad J Karczewski and Michael P Snyder. Integrative omics for health and disease. *Nature Reviews Genetics*, 19(5):299, 2018.
2. Stephan C Schuster. Next-generation sequencing transforms today's biology. *Nature methods*, 5(1):16–18, 2008.
3. Kenneth J Livak and Thomas D Schmittgen. Analysis of relative gene expression data using real-time quantitative pcr and the 2- δ δ ct method. *methods*, 25(4):402–408, 2001.
4. Aviv Regev, Sarah A Teichmann, Eric S Lander, Ido Amit, Christophe Benoist, Ewan Birney, Bernd Bodenmiller, Peter Campbell, Piero Carninci, Menna Clatworthy, et al. Science forum: the human cell atlas. *Elife*, 6:e27041, 2017.
5. Lei S Qi, Matthew H Larson, Luke A Gilbert, Jennifer A Doudna, Jonathan S Weissman, Adam P Arkin, and Wendell A Lim. Repurposing crispr as an rna-guided platform for sequence-specific control of gene expression. *Cell*, 152(5):1173–1183, 2013.
6. Paul Datlinger, André F Rendeiro, Christian Schmidl, Thomas Krausgruber, Peter Traxler, Johanna Klughammer, Linda C Schuster, Amelie Kuchler, Donat Alpar, and Christoph Bock. Pooled crispr screening with single-cell transcriptome readout. *Nature methods*, 14(3):297–301, 2017.
7. Gioele La Manno, Ruslan Soldatov, Amit Zeisel, Emelie Braun, Hannah Hochgerner, Viktor Petukhov, Katja Lidschreiber, Maria E Kastriti, Peter Lönnerberg, Alessandro Furlan, et al. Rna velocity of single cells. *Nature*, 560(7719):494–498, 2018.
8. Yueli Cui, Yuxuan Zheng, Xixi Liu, Liying Yan, Xiaoying Fan, Jun Yong, Yuqiong Hu, Ji Dong, Qingqing Li, Xinglong Wu, et al. Single-cell transcriptome analysis maps the developmental track of the human heart. *Cell reports*, 26(7):1934–1950, 2019.
9. Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003.
10. Terry M Therneau and Patricia M Grambsch. The cox model. In *Modeling survival data: extending the Cox model*, pages 39–77. Springer, 2000.
11. Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, 67(2):301–320, 2005.
12. Hemant Ishwaran, Thomas A Gerds, Udaya B Kogalur, Richard D Moore, Stephen J Gange, and Bryan M Lau. Random survival forests for competing risks. *Biostatistics*, 15(4):757–773, 2014.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

13. Vanessa M Strike, Jijun Gao, and Pratima Bansal. Being good while being bad: Social responsibility and the international diversification of us firms. *Journal of International Business Studies*, 37(6):850–862, 2006.
14. Anny Xiang, Pablo Lapuerta, Alex Ryutov, Jonathan Buckley, and Stanley Azen. Comparison of the performance of neural network methods and cox regression for censored survival data. *Computational statistics & data analysis*, 34(2):243–257, 2000.
15. Boyu Lyu and Anamul Haque. Deep learning based tumor type classification using gene expression data. In *Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics*, pages 89–96, 2018.
16. Shiyong Ma and Zhen Zhang. Omicsmapnet: Transforming omics data to take advantage of deep convolutional neural network for discovery. *arXiv preprint arXiv:1804.05283*, 2018.
17. Anne E Carpenter, Thouis R Jones, Michael R Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David A Guertin, Joo Han Chang, Robert A Lindquist, Jason Moffat, et al. Cellprofiler: image analysis software for identifying and quantifying cell phenotypes. *Genome biology*, 7:1–11, 2006.
18. Vebjorn Ljosa, Peter D Caie, Rob Ter Horst, Katherine L Sokolnicki, Emma L Jenkins, Sandeep Daya, Mark E Roberts, Thouis R Jones, Shantanu Singh, Auguste Genovesio, et al. Comparison of methods for image-based profiling of cellular morphological responses to small-molecule treatment. *Journal of biomolecular screening*, 18(10):1321–1329, 2013.
19. Seid Miad Zandavi, Derong Liu, Vera Chung, Ali Anaissi, and Fatemeh Vafaee. Fotomics: fourier transform-based omics imagification for deep learning-based cell-identity mapping using single-cell omics profiles. *Artificial Intelligence Review*, 56(7):7263–7278, 2023.
20. Boris Kovalerchuk, Bedant Agarwal, and Divya Chandrika Kall. Solving non-image learning problems by mapping to images. In *2020 24th International Conference Information Visualisation (IV)*, pages 264–269. IEEE, 2020.
21. Simon Anders. Visualization of genomic data with the hilbert curve. *Bioinformatics*, 25(10):1231–1235, 2009.
22. Bongki Moon, Hosagrahar V Jagadish, Christos Faloutsos, and Joel H. Saltz. Analysis of the clustering properties of the hilbert space-filling curve. *IEEE Transactions on knowledge and data engineering*, 13(1):124–141, 2001.
23. Alok Sharma, Edwin Vans, Daichi Shigemizu, Keith A Boroevich, and Tatsuhiko Tsunoda. Deepinsight: A methodology to transform a non-image data to an image for convolution neural network architecture. *Scientific reports*, 9(1):1–7, 2019.
24. Guillermo López-García, José M Jerez, Leonardo Franco, and Francisco J Veredas. Transfer learning with convolutional neural networks for cancer survival prediction using gene-expression data. *PloS one*, 15(3):e0230536, 2020.
25. Minoru Kanehisa, Yoko Sato, Miho Furumichi, Kanae Morishima, and Mao Tanabe. New approach for understanding genome variations in kegg. *Nucleic acids research*, 47(D1):D590–D595, 2019.
26. Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.
27. David Avis, David Bremner, and Raimund Seidel. How good are convex hull algorithms? *Computational Geometry*, 7(5-6):265–301, 1997.
28. Shmuel Winograd. On computing the discrete fourier transform. *Mathematics of computation*, 32(141):175–199, 1978.
29. Vladimir Yu Kiselev, Andrew Yiu, and Martin Hemberg. scmap: projection of single-cell rna-seq data across data sets. *Nature methods*, 15(5):359–362, 2018.
30. Rhonda Bacher and Christina Kendziorski. Design and computational analysis of single-cell rna-sequencing experiments. *Genome biology*, 17(1):1–14, 2016.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23

31. Catalina A Vallejos, Davide Risso, Antonio Scialdone, Sandrine Dudoit, and John C Marioni. Normalizing single-cell rna sequencing data: challenges and opportunities. *Nature methods*, 14(6):565, 2017.
32. Peter V Kharchenko, Lev Silberstein, and David T Scadden. Bayesian approach to single-cell differential expression analysis. *Nature methods*, 11(7):740–742, 2014.
33. Greg Finak, Andrew McDavid, Masanao Yajima, Jingyuan Deng, Vivian Gersuk, Alex K Shalek, Chloe K Slichter, Hannah W Miller, M Juliana McElrath, Martin Prlic, et al. Mast: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell rna sequencing data. *Genome biology*, 16(1):1–13, 2015.
34. Mark D Robinson, Davis J McCarthy, and Gordon K Smyth. edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140, 2010.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_23