



Review

Deep learning applications in single-cell genomics and transcriptomics data analysis



Nafiseh Erfanian^a, A. Ali Heydari^{b,c}, Adib Miraki Feriz^a, Pablo Iañez^d, Afshin Derakhshani^e, Mohammad Ghasemigol^f, Mohsen Farahpour^g, Seyyed Mohammad Razavi^g, Saeed Nasseri^h, Hossein Safarpour^{h,*}, Amirhossein Sahebkar^{i,j,k,**}

^a Student Research Committee, Birjand University of Medical Sciences, Birjand, Iran

^b Department of Applied Mathematics, University of California, Merced, CA, USA

^c Health Sciences Research Institute, University of California, Merced, CA, USA

^d Cellular Systems Genomics Group, Josep Carreras Research Institute, Barcelona, Spain

^e Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, AB, Canada

^f College of Engineering & Mines, University of North Dakota, ND, USA

^g Department of Electronics, Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran

^h Cellular and Molecular Research Center, Birjand University of Medical Sciences, Birjand, Iran

ⁱ Biotechnology Research Center, Pharmaceutical Technology Institute, Mashhad University of Medical Sciences, Mashhad, Iran

^j Applied Biomedical Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

^k Department of Biotechnology, School of Pharmacy, Mashhad University of Medical Sciences, Mashhad, Iran

ARTICLE INFO

Keywords:

Deep Learning
Single-cell omics
Genomics
Transcriptomics
Multi-omics integration

ABSTRACT

Traditional bulk sequencing methods are limited to measuring the average signal in a group of cells, potentially masking heterogeneity, and rare populations. The single-cell resolution, however, enhances our understanding of complex biological systems and diseases, such as cancer, the immune system, and chronic diseases. However, the single-cell technologies generate massive amounts of data that are often high-dimensional, sparse, and complex, thus making analysis with traditional computational approaches difficult and unfeasible. To tackle these challenges, many are turning to deep learning (DL) methods as potential alternatives to the conventional machine learning (ML) algorithms for single-cell studies. DL is a branch of ML capable of extracting high-level features from raw inputs in multiple stages. Compared to traditional ML, DL models have provided significant improvements across many domains and applications. In this work, we examine DL applications in genomics, transcriptomics, spatial transcriptomics, and multi-omics integration, and address whether DL techniques will prove to be advantageous or if the single-cell omics domain poses unique challenges. Through a systematic literature review, we have found that DL has not yet revolutionized the most pressing challenges of the single-cell omics field. However, using DL models for single-cell omics has shown promising results (in many cases outperforming the previous state-of-the-art models) in data preprocessing and downstream analysis. Although developments of DL algorithms for single-cell omics have generally been gradual, recent advances reveal that DL can offer valuable resources in fast-tracking and advancing research in single-cell.

Abbreviations: AEs, Autoencoders; ANNs, Artificial Neural Networks; BBKNN, batch balanced k -nearest neighbors; DL, Deep Learning; DNNs, Deep neural networks; DCA, Deep Count AE network; ERCC, External RNA Controls Consortium; HDST, high-definition spatial transcriptomics; FFNNs, Feed Forward Neural Networks; ML, machine learning; MMD, Maximum Mean Discrepancy; MNNs, mutual nearest neighbors; scRNA-seq, single-cell RNA sequencing; PCA, principal component analysis; RNN, Recurrent Neural Network; LSTMs, Long Short-Term memory; CNNs, Convolutional Neural Network; VAEs, Variational Autoencoders; GANs, Generative Adversarial Networks; QC, Quality Control; MAGIC, Markov Affinity-based Graph Imputation of Cells; SAVER, single-cell analysis via expression recovery; NGS, Next Generation Sequencing; TFs, transcription factors.

* Corresponding author.

** Corresponding author at: Biotechnology Research Center, Pharmaceutical Technology Institute, Mashhad University of Medical Sciences, Mashhad, Iran.

E-mail addresses: safarpour701@yahoo.com (H. Safarpour), amir_saheb2000@yahoo.com (A. Sahebkar).

<https://doi.org/10.1016/j.bioph.2023.115077>

Received 24 April 2023; Received in revised form 22 June 2023; Accepted 23 June 2023

Available online 1 July 2023

0753-3322/© 2023 The Author(s). Published by Elsevier Masson SAS. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since single-cell sequencing (sc-seq) was highlighted as the “Method of the Year” in 2013 [2], sequencing individual cells at the single-cell resolution has become the norm for studying cell-to-cell heterogeneity. RNA and DNA single-cell measurements, and recently, epigenetic and protein levels, stratify cells at the highest possible resolution. That is, single-cell RNA sequencing (scRNA-seq) makes it possible to measure transcriptome-wide gene expression at the single-cell level. Such resolution enables researchers to distinguish different cell types based on their characteristics [see [3–6]], organize cell populations, and identify cells transitioning between states. Such analyses provide a much better picture of tissues and the underlying dynamics in the development of an organism, which in turn allow for delineating intra-population heterogeneities that had previously been perceived as homogeneous by bulk RNA sequencing. Similarly, single-cell DNA sequencing (scDNA-seq) studies can reveal somatic clonal structures [e.g., in cancer, see [7,8]], thereby helping to monitor cell lineage development and providing insight into evolutionary mechanisms that function on somatic mutations.

The prospects resulting from sc-seq are tremendous: it is now possible to re-evaluate hypotheses on differences between predefined sample groups at a single-cell level regardless of samples being disease subtypes, treatment groups, or morphologically distinct cell types. As a result, in recent years, enthusiasm about the possibility of screening the basic units of life’s genetic material has continued to expand. Human Cell Atlas [9] is a prominent example: an effort to sequence the various cell types and cellular states that make up a human being. Encouraged by the great potential of single-cell investigation of DNA and RNA, there has been substantial growth in the development of related experimental technologies. In particular, the advent of microfluidics techniques and combinatorial indexing strategies [10,11] has resulted in routinely sequencing hundreds of thousands of cells in a single experiment. This growth has also allowed a recent publication to analyze millions of cells at once [12]. More large-scale sc-seq datasets are becoming accessible worldwide (with tens of thousands of cells), constituting a data explosion in single-cell analysis platforms. The continuous growth of the scale and quantity of available sc-seq data has raised significant questions: 1) How do we correctly interpret and analyze the increasing complexity of sc-seq datasets? 2) How can different types of data sets (mentioned above) provide a deeper understanding of the underlying biological dynamics for a specific condition? and 3) how can the acquired information be changed into practical applications in medicine, ranging from rapid and precise diagnoses to accurate medicine and targeted preventions. As we face the challenges of a rise in chronic diseases, aging populations, and limited resources, a transformation towards intelligent analysis, interpretation, and understanding of complex data is essential. In this paradigm shift, the rapidly emerging field of ML is central [13].

ML is the study of models which can learn from data without the need for an explicit set of instructions. Simultaneously with biomedical advancements of the past decade, there has been a surge in the development and application of ML algorithms, prominently led by advances in DL. Some of the earliest DL algorithms developed were intended to computationally model our brains’ learning process, therefore being called “Artificial Neural Networks” (ANN). DL models often consist of many processing layers (with many nodes in each layer), which enable them to learn a representation of data with several levels of abstraction. The recent improvements in computational hardware have made training DL models feasible, resulting in successful and revolutionary applications of such models across many domains.

This review aims to discuss DL applications in sc-seq analysis and elaborate on their instrumental role in improving sc-seq data processing and analysis. We first introduce some key concepts in DL, which enable us to discuss their applications in the sc-seq field. Because of the high technical noise and complexity of sc-seq data, we review appropriate techniques for analyzing such data, which ensure robustness and

reproducibility of results. Finally, we conclude this work by discussing possible developments, highlighting both emerging obstacles and opportunities to advance the rapidly evolving field of single-cell omics.

2. Essential concepts in deep learning

The field of DL has a rich history, beginning with the development of the MCP artificial neural model by McCulloch and Pitts in 1943 [14]. The concept of the perceptron was later introduced by Rosenblatt, building on the foundation of artificial neurons [15]. In 1974, the backpropagation algorithm was proposed by Werbos, enabling the use of multilayer neural networks [16]. However, it was not until 2006, with the introduction of Hinton’s algorithm [17], that deep learning technology truly reached its full potential by effectively addressing the issue of gradient disappearance in backpropagation.

DL is a subset of ML that uses deep neural networks (DNNs) to analyze large and complex datasets. DNNs are made up of layers of artificial neurons, which are inspired by the way human neurons work. Each neuron takes the weighted summation of all inputs and passes it through a non-linear activation function, such as sigmoid, rectifier (i.e., rectified linear unit [ReLU]), or hyperbolic tangent. A deep feed-forward neural network (DFNN) is the most basic deep architecture (Fig. 1a), which is created by stacking layers of neurons. The input information flows from the input layer through the hidden layers and the model generates an output at the last layer, the output layer. The large set of trainable weights of the neurons and the non-linear transformations enable DNNs to capture underlying complex patterns of the data. Training a DNN is the process of determining these trainable weights to optimize model performance.

DL models are trained through a process called backpropagation, which uses mathematical calculations to adjust the model’s parameters or weights based on the prediction errors. DL can be divided into three main categories: supervised learning, unsupervised learning, and semi-supervised learning. Supervised learning is commonly used in deep feed-forward neural networks (DFNNs), where the goal is to map input data to a representation that can be used for tasks such as classification or regression. Semi-supervised learning is used when only a small number of data points have labels, using the limited labels to inform the representation and labeling of the unlabeled data. Unsupervised learning is used to identify underlying patterns in data without any supervision, and is commonly used in sc-seq data analysis for dimensionality reduction and cell clustering.

In the following paragraphs, we provide an overview of several frequently-used DL models for single-cell applications.

2.1. Recurrent Neural Network (RNNs)

RNNs [18] are used for processing sequential data, including natural language and time series. RNNs process sequential inputs one at a time and implicitly maintain a history of previous elements in the input sequence. A typical architecture for RNNs is presented in Fig. 1c. Similar to FFNNs, RNNs learn by propagating the gradients of each hidden state’s inputs at discrete times. This process becomes more intuitive if we consider the outputs of the hidden units at various time iterations as if they were the outputs of different neurons in a deep multilayer network. However, deep RNNs are notoriously hard to train due to vanishing or exploding gradients. Moreover, RNNs are not parallelizable (since each layer depends sequentially on the output of the previous modules) and therefore expensive to train. These issues have paved the way for newer sequence-to-sequence models which have tried to alleviate RNN’s training challenges. Among these models, Long Short-Term memory (LSTMs) [19] have been able to somewhat address the gradient issues, while Transformers [20] have enabled deeper and more stable parallelizable sequence-to-sequence models. Though these models remain underexplored in SC omics, some studies have already shown promising application for SC omics analysis, including [21,22] [using

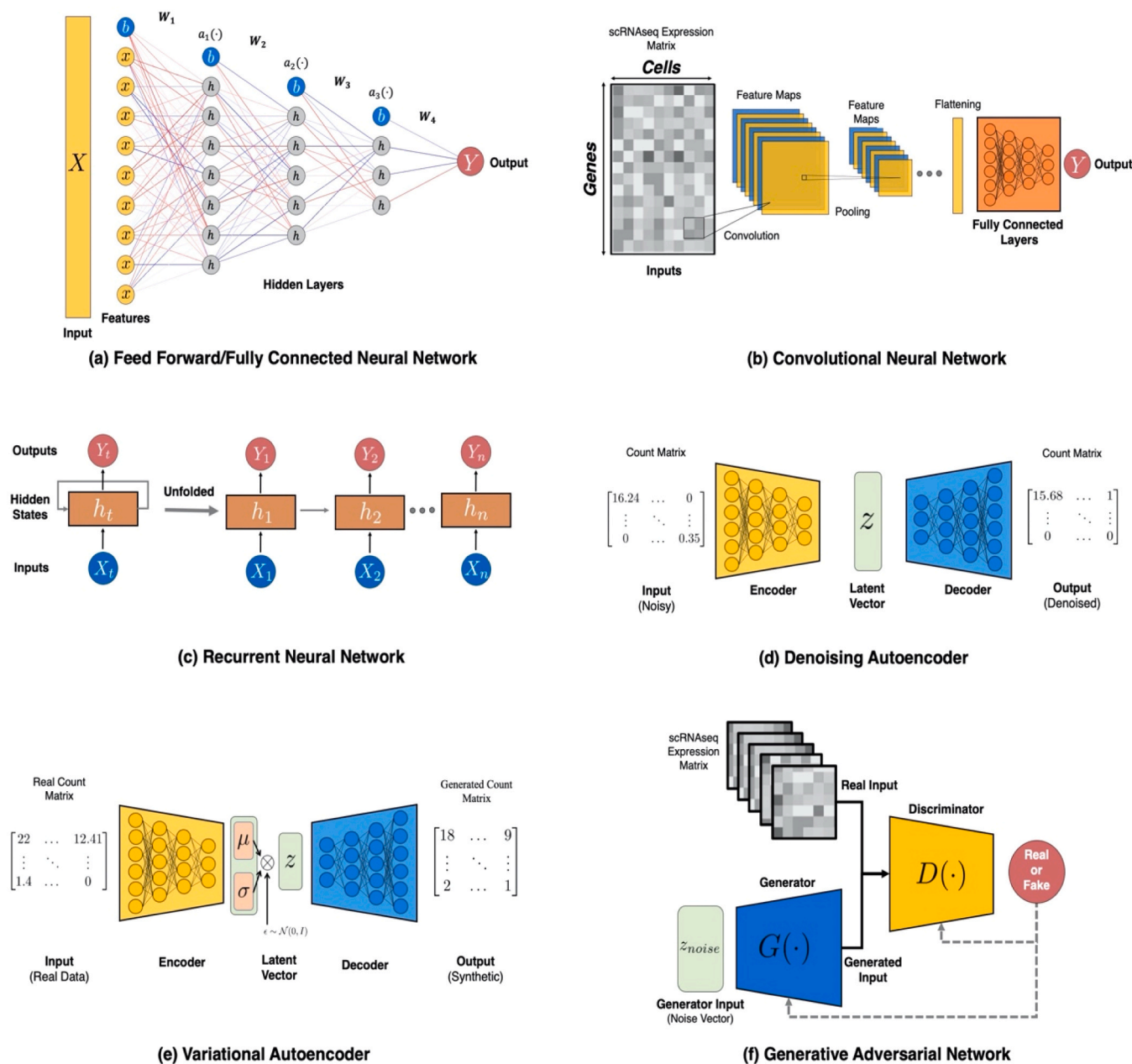


Fig. 1. (a) An example of an FFNN architecture. In this network, the raw input is shown as X , the weight tensors are denoted W_i , activation layers are α_i and the bias nodes are shown in blue nodes marked by b . The positive and negative weights are shown as blue and red edges, respectively, where a smaller numerical value corresponds to more transparent edges and vice versa. (b) An example of CNN architecture for classification. In this hypothetical architecture, the model feeds the inputs through the three stages of a CNN (with non-linear activation not depicted) to extract features. As is often the case, the output of the CNN is inputted to a fully connected neural network for classification. (c) An RNN depicted with its training flow, and an unrolled version showing the timestep-dependent inputs, hidden state, and outputs are marked. (d) An illustration of a denoising autoencoder. Autoencoders (AE) have an encoder network, which tries to extract the most critical features and map them to a latent space (often of a much smaller dimension than the input data). The decoder network takes in the latent vector produced by the encoder (denoted as z) and maps this representation back to the original input dimension. (e) The architecture of a Variational AE (VAE) that aims at generating synthetic scRNA-seq data using the gene expression values. (f) A depiction of generative adversarial network for generating in-silico scRNA-seq data. In this illustration, the scRNA-seq count matrix is converted to images that are used as real samples.

LSTMs to predict cell types and cell motility] and [23] [using Transformers to generate meaningful embeddings from sparse scRNAseq data].

2.2. Convolutional Neural Network (CNNs)

CNNs [24] are specialized types of networks that use convolution (the mathematical operation) instead of tensor multiplication (which is done in FFNNs) in at least one of their layers (Fig. 1b). This convolution

operation makes CNNs ideal for processing data with a grid-like topology (images are the quintessential example of such datasets). Compared to other ANNs, CNNs have three key benefits: (i) sparse interactions, (ii) shared weights, and (iii) equivariant representations [24]. CNNs have been effectively used in many applications in computer vision and time-series analysis but are not utilized as frequently for sc-seq applications (since sc-seq datasets do not have a grid-like structure). However, some studies, such as Xu et al. [25] have used CNNs after converting sc-seq data to images, which have shown promising results.

2.3. Autoencoders (AEs)

AEs are neural networks that aim to reconstruct (or copy) the original input via a non-trivial mapping. Conventional AEs have an “hourglass” architecture with two mirroring networks: an encoder and a decoder. The encoder’s task is to map the input data to a latent space, often of a much smaller dimension than the original input space. The encoder is responsible for data compression and feature extraction, forming the narrowing part of the hourglass architecture (Fig. 1d). The output of the encoder network (latent vector) will contain the most important features present in the data in a compressed form. Conversely, the decoder is tasked with mapping the latent vector back to the original input dimension and reconstructing the original data. In the ideal case, the decoder’s output will be an exact copy of the training sample.

Fig. 1(d) depicts an example of a denoising AE. AEs have been used extensively for scRNAseq analysis, such as batch effect removal [26], data imputation [27] and multi-omics integration [28]. We describe the applications of AEs in SC analysis in the following sections.

2.4. Variational Autoencoders (VAEs)

VAEs [29,30] are generative models that learn latent-variable and inference models simultaneously, i.e., they are made up of generative and inference models. VAEs are AEs that employ variational inference to recreate the original data, allowing them to produce new (or random) data that is “similar” to that which already exists in a dataset (Fig. 1e). VAEs have better mathematical properties and training stability than Generative Adversarial Networks (GANs), but they suffer from two major weaknesses: classic VAEs create “blurry” samples (those that adhere to an average of the data points), rather than the sharp samples that GANs generate due to adversarial training. Introspective VAEs (IntroVAEs) Huang et al. have generally solved this issue by specifying an adversarial training between the encoder and the decoder. IntroVAEs are single-stream generative algorithms that assess the quality of the images they generate. They have largely been employed in computer vision, where they have outperformed their GAN counterparts in applications like synthetic image generation [31] and single-image super-resolution [32].

The other major issue with VAEs is *posterior collapse*: when the variational posterior and actual posterior are nearly identical to the prior (or collapse to the prior), which results in poor data generation quality [33]. Posterior collapse has been attributed to VAE’s distribution regularization term in the objective function [34], i.e. when the prior and posterior divergence is close to zero. Studies aimed at reducing posterior collapse can be divided into two categories: (i) solutions aimed at weakening the generative model [35,36], (ii) alterations to the training purpose [35,37,38], and (iii) alterations to the training procedure [33,39]. If both issues mentioned above can be addressed, VAEs have shown to perform comparable (or similar) to GANs while training faster due to the simpler training procedure. VAEs have been utilized extensively for SC omics analysis: Since VAEs’ latent space is well regularized (as opposed to AEs), the latent representation can be used for clustering and dimensionality reduction [40]. Given that VAEs are generative models, they can also be used to impute missing data [41] or generate realistic sc-seq for data augmentation, further improving downstream SC omic analyses [42].

2.5. Generative Adversarial Networks (GANs)

GANs [43] can generate realistic synthetic data and have been effectively utilized in a variety of computer vision tasks [44–46], natural language processing [47,48], time series synthesis [49,50], and bioinformatics [51]. GANs are made up of a generator network (G) and a discriminator network (D) that compete in a zero-sum game; we present the architecture of GANs in Fig. 1f. The goal of the G network is to generate fake samples that resemble the distribution of the real data,

“fooling” the D network into believing that these fake samples are real. Conversely, D trains to learn the difference between real and synthetic samples and “discriminate” between them. In each GAN training iteration, the entire system is re-adjusted to update both G and D parameters. In this process and through many iterations, the generator learns to make more realistic samples which deceive the discriminator as real data. At the same time, the discriminator is learning the distinction between real and generated data (from the G network). GANs’ ability to produce realistic samples is attributed to the adversarial training between G and D networks. Compared to other generative models, GANs have several advantages, such as the flexibility to learn any distribution, requiring no assumptions on the prior distribution, and no limitations on the size of the latent space.

Despite these advantages, GANs are notoriously difficult to train since achieving Nash equilibrium for G and D is extremely difficult [52]. Another drawback of GANs is vanishing gradients, which occurs if D learns the distinction between real and generated data well too quickly, prohibiting G from training properly [53]. Another problem with GANs is “mode collapse,” which occurs when G produces only a small number of outputs that potentially trick D. This happens when G has trained to map many noise vectors to the same output that D recognizes as real data. Quantifying how much GANs have learned the distribution of real data is often difficult, hence one of the most common methods of assessing GANs is to evaluate the output directly [54], which could be laborious. Even though certain GAN variations have been proposed to reduce vanishing gradients and mode collapse, (e.g., Wasserstein-GANs (WGANs) [53] and Unrolled-GANs [55]), the convergence of GANs is still a big issue. Due to their ability to generate realistic samples, we anticipate GANs to be increasingly utilized for SC omic analysis. Indeed, GANs have shown promising results in generating realistic ta [51], imputing and augmenting existing datasets [25] and predicting the molecular progress of Alzheimer’s disease [56].

3. Essential concepts in single-cell omics

A cell’s state is defined and characterized mainly by the interaction between its genome, epigenome, transcriptome, proteome, and metabolome. The data obtained from these layers is usually referred to as “omics”. Until recently, attaining various omics data at a single-cell resolution was not possible, due to many obstacles concerning single-cell isolation and the analysis of minute quantities of molecules present in each cell. However, recent technological and engineering advances have enabled many of such “omic” layers to be read and sequenced at single-cell resolution. The advent of SC omics has enabled the exploration of heterogeneity between cells and within populations, which previously could have been masked and thought to be homogeneous.

3.1. Transcriptomics

Transcriptomics refers to the study of the transcriptome (the entire collection of RNA transcripts generated by the genome under specific conditions or in a particular cell) utilizing high-throughput technologies such as sc-seq. Comparison of transcriptomes allows for the identification of genes that are differentially expressed in various cell types or in response to different stimuli. SC transcriptomics can provide unique insights into the differentiation dynamics, cellular responses to stimulation, and transcription’s stochastic nature. The recent technological advancements allow SC transcriptomics to be used for constructing single-cell resolution reference maps of healthy human tissues, organs, and systems, which will be crucial in many biological disciplines. [57].

3.2. Genomics

As a discipline, genomics is the study of genome structure, function, evolution, mapping, and editing. SC genomics is the use of omics

methods and technologies to examine the uniqueness of cells. Such analyses provide a new understanding of complex biological events and systems. Multiple kinds of mutations were studied at the single-cell level, revealing unexpected insights into the genetic heterogeneity that may occur throughout the human life cycle [58].

3.3. Epigenomics

Epigenomics is the study of the physical modifications, including DNA methylation, histone modification, chromatin accessibility, and chromosome conformation, which empower cells with identical

genotypes to have distinct gene expression profiles. The advancement of experimental techniques for characterizing epigenomes at the single-cell level has given rise to new possibilities, providing a better understanding of epigenetic processes and their complex interactions with gene regulation. Recent technological advances have made it possible to also evaluate and identify the dynamics of epigenetic maintenance and reprogramming at the single-cell level [59].

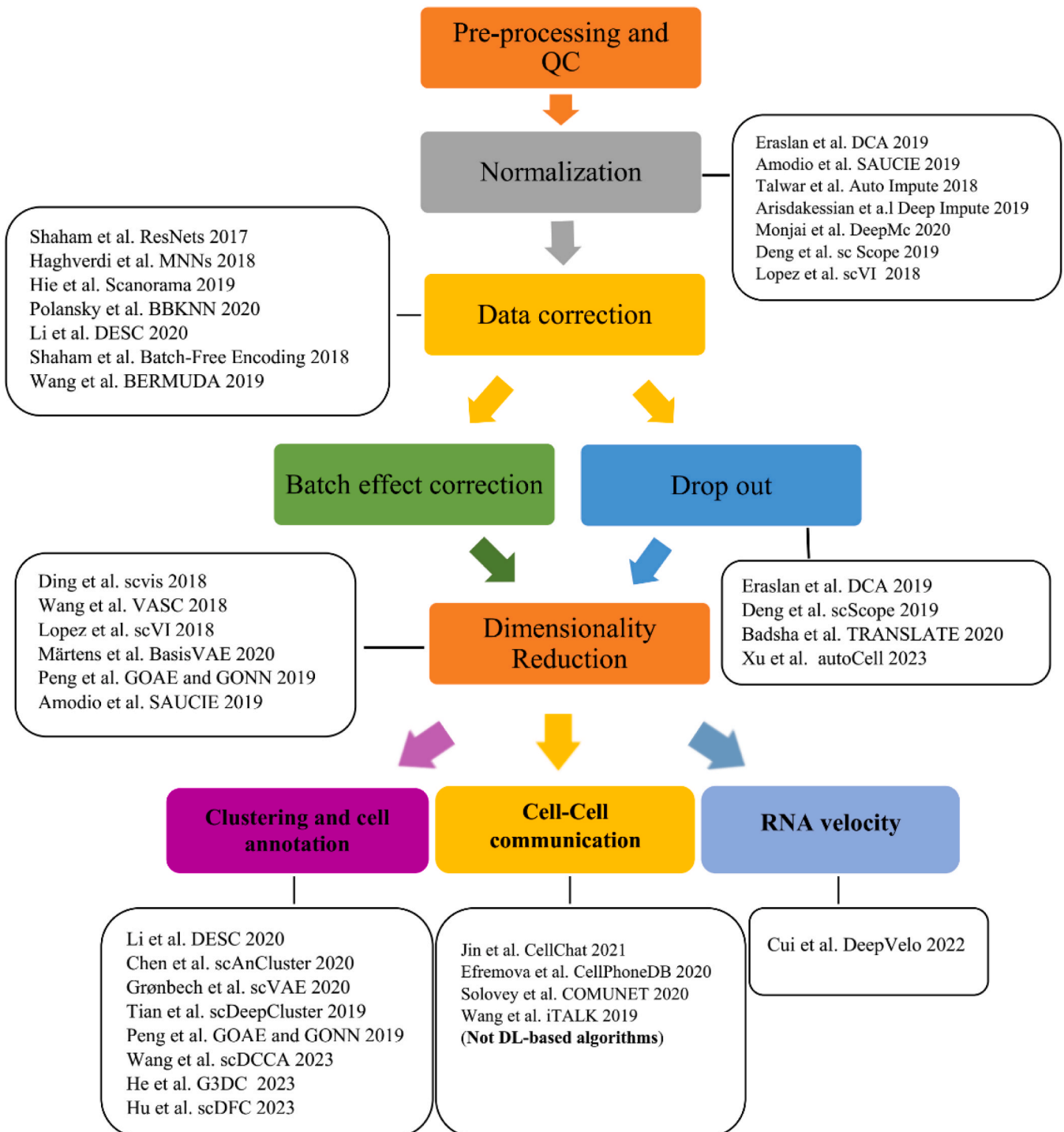


Fig. 2. Workflow of RNA-seq data analysis.

4. Deep Learning applications in single-cell omics

4.1. Deep learning in single-cell transcriptomics

Single-cell RNA sequencing (scRNA-seq) has improved our

understanding of biological processes substantially in recent years. Researchers have shown the potentials of single-cell (SC) transcriptomics through studying the cellular heterogeneity of many organisms, such as humans, mice, zebrafish, frogs, and planaria [5,6], and uncovering previously unknown cell populations [60,61]. Moreover, these studies

Table 1
Computational approaches for the various stages of scRNA-seq analysis.

Stage	Algorithm	Advantage	Disadvantage	REF
Normalization	DCA	extremely scalable to datasets with millions of cells or more	For downstream analyses, the hyperparameters were not fine-tuned.	[62]
	SAUCIE	execute a number of critical tasks on single-cell datasets in a highly scalable manner (by using deep learning's parallelizability with GPUs) without requiring the use of other algorithms or processing methods	Covariance may not adequately represent nonlinear variation.	[63]
	Auto Impute	scales adequately to massive datasets including thousands of transcriptomes	Variance stabilization may not be a meaningful criterion for evaluating scRNA-seq data.	[64]
	Deep Impute	Fast speed and high scalability	Not reported	[65]
	DeepMc	High accuracy	does not make any assumptions about gene expression distribution	[66]
	Sc Scope	Fast speed and high scalability	not fare well with a low number of components	[67]
Data correction	scVI	High accuracy	GPU Needed	[68]
	ResNets	Fast speed and low memory peak	not bringing any cell clusters of the same type closer together	[69]
	MNNs	Fast speed	Memory usage	[69]
	Scanorama	Low memory usage	Time usage	[69]
	BBKNN	Lowest memory usage	Low cell number	[69]
	DESC	Handling full of features	Low scalability	[70]
	Batch-Free Encoding	Not reported	Not reported	-
	BERMUDA	Compatible with R and Python	High memory usage	[71]
	SCIBER	Better accuracy in removing batch effects on real datasets	Not reported	[136]
	Dropout	DCA	High scalability	Low accuracy
Sc Scope		Fast speed and high scalability	not fare well with a low number of components	[67]
TRANSLATE		Combines autoencoder with transfer learning	Not reported	[73]
autoCell		Analyze high-dimensional, sparse scRNA-seq data and strong preservation of large-scale distances	Cannot handle batch effects in scRNA-seq data	[74]
Dimensionality reduction	scvis	broad dataset compatibility	Low accuracy	[75]
	VASC	Low memory usage	Not reported	[67]
	scVI	flexible and scalable with huge datasets	Poor in integrating strong batch effects	[70]
	BasisVAE	High accuracy	Doesn't provide sufficient insights	[76]
	GOAE and GONN	Low memory usage	Not reported	[77]
	SAUCIE	Handling full of features	Low accuracy	[70]
Clustering and cell annotation	DESC	High Accuracy	Low scalability	[70]
	scAnCluster	Not reported	Low speed	[78]
	scVAE	High accuracy in huge datasets	Not reported	[79]
	scDeepCluster	High accuracy	Low accuracy in small datasets	[79]
	GOAE and GONN	significantly improves clustering	Not reported	[77]
	scDCCA	Separate different cell types most effectively with only a few outliers	Not reported	[157]
	G3DC	More efficient to fuse attributes, structure, and attention information	Not reported	[80]
	scDFC	Fast speed	Not reported	[159]
	Cell-Cell communication	CellChat	Fast speed, High accuracy	Low accuracy
CellPhoneDB		Not reported	Low enrichment score	[81]
COMUNET		Not reported	Not reported	[81]
CellCall		Not reported	Memory Usage	[81]
CytoTalk		Not reported	Time Usage	[81]
SingleCellSignalR		Not reported	Memory Usage	[81]
ICELLNET		Fast Speed	Not reported	[81]
GCNG		Low memory usage	Not reported	[81]
iTALK		High Accuracy	Low accuracy	[81]
RNA velocity Multi-Modal Integration		DeepVelo	flexible to encompass diverse joint-learning strategy	Low speed
	scMVAE		No guiding principles are provided with respect to how to pick a specific learning strategy for the specific dataset	[83]
	DCCA	Able to generate biologically meaningful missing omics data based on the learned latent representation of another omics data	Performance is not robust against high levels of noise	[84]
	totalVI	Computationally scalable and flexible	Not reported	[85]
	BABEL	efficient interoperable design, resulting in efficient cross-modality prediction	is limited by the amount of mutual information shared between the input data modalities	[86]
	DeepMAPS	Learns interpretable cell type-specific biological networks based on data modality	Computational cost does not scale efficiently to super-larger datasets. Reproducibility could be dependent on the specific GPU model.	[87]
	scArches	Facilitates the integration of query datasets onto a reference dataset through data integration	Not reported	[88]
	ExpiMap	Enables the interpretable mapping of SC transcriptome data sets across cohorts, disease states and other perturbations.	Not reported	[89]

have also highlighted the need for better computational methods which can facilitate the analysis of large and complex scRNA-seq datasets. Many scRNA-seq data are publicly available for download, with most hosted on popular databases such as the Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>) and Human Cell Atlas (<https://www.humancellatlas.org/>) and 10x Genomics (<https://www.10xgenomics.com/resources/datasets>). In the following sections, we provide an overview of existing computational approaches for the various stages of scRNA-seq analysis (summarized in Fig. 2) (Table 1).

4.1.1. Pre-processing and quality control

Cell barcodes are designed to delineate various cell populations present within a sequenced sample. However, barcodes can mistakenly tag several cells (doublet) or not tag any cells at all (empty droplet/well), which prompts the quality control (QC) step in scRNA-seq analysis. Many raw data pre-processing pipelines, including Cell Ranger [94], indrops [95], SEQC [96], and z-unique molecular identifiers (zUMIs) [97], can perform QC. The dimension of the count matrix generated by sequencing technologies depends on the number of barcodes and transcripts. Though the noise rate in measurements varies across reads and count data, standard research pipelines very often employ the same processing techniques [98].

Despite scRNA-seq data's richness which can offer significant and deeper insights, the data's complexity and noise are far higher than traditional bulk RNA-seq, making it challenging to process the raw data for downstream analysis. Unwanted variations such as biases, artifacts, etc., require extensive QC and normalization efforts [99]. The number of counts per barcode (count depth), the number of genes per barcode, and the fraction of counts from mitochondrial genes per barcode are three QC covariates widely used in the QC step [100]. On the other hand, other experimental factors (such as damaging the sample during dissociation) could result in low-quality scRNA-seq libraries, which can yield erroneous findings in downstream analyses. Currently, there is an unmet need for developing more efficient and accurate methods for filtering low-quality cells during the library preparation. Given DL algorithms' multi-tasking and scaling capabilities, DL models for end-to-end analyses will be of utmost importance and utility to SC omics. DL models are uniquely positioned in this space since they can be trained to perform QC, data correction and downstream analyses all together, with each task informing and improving the performance of the others. We anticipate such models to outperform the traditional SC analysis pipelines, which use separate frameworks to perform various tasks independently. Though such end-to-end DL models remain nascent, preliminary works in this area (such as the scVI framework [40]) have shown promising results, demonstrating that DL models hold a great potential in the development of end-to-end SC pipelines.

Due to the limited number of studies regarding pre-processing and quality control, we will focus on the DL applications that are most related to normalization, data correction, and downstream analysis.

4.1.2. Normalization

Normalization is a crucial first step in pre-processing scRNA-seq expression data to address the constraints caused by low input content, or the different types of systematic measuring biases [101]. Normalization aims to detect and remove changes in measurements between samples and features (e.g., genes) caused by technical artifacts or unintended biological effects (e.g., batch effects) [102]. Methods designed for normalizing bulk RNA-seq and microarray data are often used to normalize scRNA-seq data. However, these techniques often ignore essential aspects of scRNA-seq results [102]. For scRNA-seq data, a few families of normalization techniques have been developed, such as scaling techniques [103], regression-based techniques for identified nuisance factors [101,104], and techniques based on spike-in sequences from the External RNA Controls Consortium (ERCC) [105,106]. However, these methods are specific to certain experiments and can not be

applied to all research designs and experimental protocols. Although some DL-based methods have been proposed to generalize the normalization stage, accounting for technical noise of scRNA-seq data still remains a challenge and an active area of research within the field [13].

4.1.3. Data correction

Although normalization aims to address the noise and bias in the data, normalized data can still contain unexpected variability. These additional technical and biological variables, such as batch, dropout, and cell cycle effects are accounted for during the "data correction" stage, which depends on the downstream analysis [107]. In addition, it is a recommended practice to address biological and technical covariates separately [107], since they serve different purposes. Given the mentioned subtleties, designing DL models that can address most of these challenges is difficult. Therefore, there are currently no DL models that are widely used for data correction within the field.

4.1.3.1. Dropout. Compared to bulk RNA-seq, scRNA-seq datasets are noisy and sparse and pose unique nuances such as "dropout" which is one of the most significant issues in this field [108,109]. Dropout occurs when a gene is observed at a moderate or high expression level in one cell but is not detected in another cell [110]. Dropout events can occur during library preparation (e.g. extremely low levels of mRNA in single cells) or due to biological properties (the stochastic aspect of gene expression in multiple cells) [111]. Additionally, shorter genes have lower counts and a greater dropout rate [112]. Overall, a low RNA capture rate results in the inability of detecting an expressed gene, leading to a "false" zero, known as a dropout event. Furthermore, it has been suggested that sometimes near-zero expression measurements can also be dropouts [113]. Dropout events will introduce technical variability and noise, adding an extra layer of difficulty in analyzing scRNA-seq [114], and downstream analyses such as clustering and pseudo-time reconstruction [65].

It is essential to understand the difference between "false" and "true" zero counts. True zero counts mean that a gene is not expressed in a particular cell type, indicating true cell-type-specific expression [26]. Hence, it is important to note that zeros in scRNA-seq data do not necessarily translate to missing values and must remain in the data. However, the false zeros (missing values) must be imputed to further improve the analysis. The missing values are replaced with either random values or by an imputation method [26]. Imputation approaches designed for bulk RNA-seq data may not be suitable for scRNA-seq data for multiple reasons, mainly due to scRNA-seq's heterogeneity and dropouts. ScRNA-seq has much higher cell-level heterogeneity than bulk RNA-seq data; scRNA-seq has cell-level gene expression data while bulk RNA-seq data represents the averaged gene expression of the cell population. Additionally, the number of missing values in bulk RNA-seq data is much lower compared to scRNA-seq [108]. Given these factors and the non-trivial difference between true and false zero counts, classic imputation approaches with specified missing values are often not appropriate for scRNA-seq data, and scRNA-seq-specific dropout imputation methods are required.

Current scRNA-seq imputation methods can be divided into two groups: (i) those that change all gene expression levels, such as Markov Affinity-based Graph Imputation of Cells (MAGIC) [115] and single-cell analysis via expression recovery (SAVER) [116], and (ii) methods that impute drop-out events (zero or near-zero counts) alone, such as scImpute [117], DrImpute [108], and LSImpute [118]. These techniques can fail to account for the non-linearity of the data's count structure. Moreover, as larger scRNA-seq datasets become available and common, imputation methods should scale to millions of cells. However, many of the earlier models are either incapable of or very slow at processing datasets of larger size (tens of thousands or more) [26]. As a result, many have resorted to designing DL-based approaches to combat these challenges, both on the technical and efficiency fronts.

Most of DL algorithms for imputing drop-out events are based on AEs. For example, in 2018, Talwar et al. proposed AutoImpute, a technique for retrieving the whole gene expression matrix using over-complete AEs to impute the dropouts. AutoImpute learns the underlying distribution of the input scRNA-seq data and imputes missing values based on the learned distribution, with minor modifications to the biologically silent gene expression values. Through expanding the expression profiles into a high-dimensional latent space, AutoImpute learns the underlying distribution and patterns of gene expression in single cells and reconstructs an imputed model of the expression matrix. At the time, Talwar et al. claimed that their system was the only model that could perform imputation on the largest of the nine datasets they studied (68 K PBMC, which contains ~68,000 cells), without running out of memory [119].

In another study, Eraslan et al. proposed the Deep Count AE network (DCA). DCA uses a negative binomial noise model both with and without zero-inflation to account for the count distribution and over dispersion while capturing nonlinear gene-gene dependencies. Since their approach scales linearly with the number of cells, DCA can be used on datasets with millions of cells. DCA also depends on gene similarities; using simulated and true datasets, DCA denoising enhances several traditional scRNA-seq data analyses. One of the key benefits of DCA is that it only requires the user to define the noise model. Current scRNA-seq approaches depend on a variety of hypotheses and often use standard count distributions, such as zero-inflated negative binomial. DCA increases biological exploration by outperforming current data imputation approaches in terms of quality and time. Overall, DCA calculates the "dropout probability" of a zero-expression value due to scRNA-seq dropout and imputes the zeros only when the probability is high. Consequently, while DCA effectively detects true zeros, it can be biased when dealing with nonzero values [26].

Badsha et al. propose TRANSfer learning with LATE (TRANSLATE) [27], a DL model for computing zeros in scRNA-seq datasets which are extremely sparse. Their nonparametric approach is based on AEs and builds on their previous method, Learning with AuToEncoder (LATE). The key presumption in LATE and TRANSLATE is that all zeros in the scRNA-seq data are missing values. In most cases, their approach achieves lower mean squared error, restores nonlinear gene-gene interactions, and allows for improved cell type separation. Both LATE and TRANSLATE are also very scalable, and when using a GPU, they can train on over a million cells in a few hours. TRANSLATE has shown better performance on inferring technical zeros than other techniques, while DCA is better at inferring biological zeros than TRANSLATE.

Sparse Autoencoder for Unsupervised Clustering, Imputation, and Embedding (SAUCIE) [120] is a regularized AE that denoises and imputes data using the reconstructed signal from the AE. Despite the noise in the input data, SAUCIE can restore the significant relationships across genes, leading to better expression profiles which can improve downstream analyses such as differential gene expression [120].

ScScope [121] is a recurrent AE network that iteratively handles imputation by employing a recurrent network layer; taking the time recurrence of ScScope to one (i.e. $T = 1$) will reduce the model to a traditional AE. Given that ScScope is a modification of traditional AEs, its runtime is similar to other AE-based models [121].

A few non-AE-based models have also been developed for imputation and denoising of scRNA-seq data. DeepImpute [65] uses several sub-neural networks to impute groups of target genes using signals (genes) that are strongly associated with the target genes. Arisdakessian et al. demonstrate that DeepImpute has a better performance than DCA, contributing the advantages to their divide-and-conquer approach [65].

Mongia et al. [66] introduced Deep Matrix Completion (deepMC), an imputation method based on deep matrix factorization for missing values in scRNA-seq data that utilizes a feed backward neural network. In most of their experiments, deepMC outperformed other existing imputation methods while not requiring any assumption on the prior distribution for the gene expression. We predict that deepMC will be the

preferred initial approach for imputing scRNA-seq data, given the superior performance and simplicity of the model.

Single-cell variational inference (scVI) is another DNN algorithm introduced by Lopez et al. [40]. ScVI is based on a hierarchical Bayesian model and uses a DNN to define the conditional probabilities, assuming either a negative binomial or a zero-inflated negative binomial distribution [40]. Lopez et al. show that scVI can accurately recover gene expression signals and impute the zero-valued entries, potentially enhancing the downstream analyses without adding any artifacts or false signals.

Recently, Patrino et al. [122] compared 19 denoising and imputation methods, based on numerous experimental scenarios such as recovery of true expression profiles, characterization of cell similarity, identification of differentially expressed genes, and computation time. Their results showed that ENHANCE (Expression deNoising Heuristic using Aggregation of Neighbors and principal Component Extraction), MAGIC, SAVER, and SAVER-X offer the best overall results when considering efficiency, accuracy and robustness for the investigated tasks [122]. Recently in 2023 Xu et al. presented autoCell, a DL-based method for scRNA-seq dropout imputation and feature extraction. It uses a VAE network with graph embedding and a probabilistic depth Gaussian mixture model to analyze high-dimensional, sparse scRNA-seq data. autoCell offers a comprehensive toolbox for visualizing, clustering, imputing, and identifying disease-specific gene networks in scRNA-seq data [74].

It is important to note that traditional methods, despite their current success, are not well-suited for large-scale scRNA-seq studies. As larger scRNA-seq datasets become the norm, we anticipate that DL-based models will prove to be advantageous. Therefore, more work is required to build upon the existing DL methods for imputing dropout effects and better managing technical zeros while retaining biological zeros.

4.1.3.2. Batch effects correction. When samples are conducted in separate batches, the term 'batch effect' is used to describe the variation caused by technical effects. Different types of sequencing machines or experimental platforms, laboratory environments, different sample sources, and even technicians who perform the experiments can cause batch effects [123]. Removing and accounting for batch effects is often helpful and recommended, however, the success varies significantly across different studies. For example, batch effect removal on bulk RNA-seq data from Encyclopedia of DNA Elements (ENCODE) human and mouse tissues [124] is a recommended standard data preparation step. Batch effect correction has been an active area of research since the microarray time. Johnson et al. suggested parametric and non-parametric empirical Bayes frameworks for adjusting the data for batch effects removal [125]. In recent years and with an increased level of complexity in sequencing datasets, more involved batch effect correction methods have been proposed and used [123]. However, a majority of the existing approaches require biological group expertise for each observation and were originally designed for bulk or microarray RNA-seq data. Given the heterogeneity present within scRNA-seq data, these earlier techniques are not well suited for single-cell analysis in certain cases [126]. Batch effects in scRNA-seq data may have a substantial impact on downstream data analysis, impacting the accuracy of biological measurements and ultimately contributing to erroneous conclusions [127]. Therefore, alternative batch effect correction techniques for scRNA-seq data have been developed to address the specific needs of single-cell datasets.

Several statistical methods, including linear regression models like ComBat [125], and nonlinear models like Seurat's canonical correlation analysis (CCA) [128] or scBatch [123], have been designed to eliminate or minimize scRNA-seq batch effects while aiming to maintain biological heterogeneity of scRNA-seq data. Additionally, some differential testing frameworks such as Linear Models for Microarray (limma) [129],

Model-based Analysis of Single-cell Transcriptomics (MAST) [130], and DESeq2 [131] already integrate the batch effect as a covariate in model design.

Haghverdi et al. [1] developed a novel and efficient batch correction

method for single-cell data which detects cell mappings between datasets, and subsequently reconstructs the data in a shared space (Fig. 3). To create relations between two datasets, the algorithm first identifies mutual nearest neighbors (MNNs). The translation vector is then

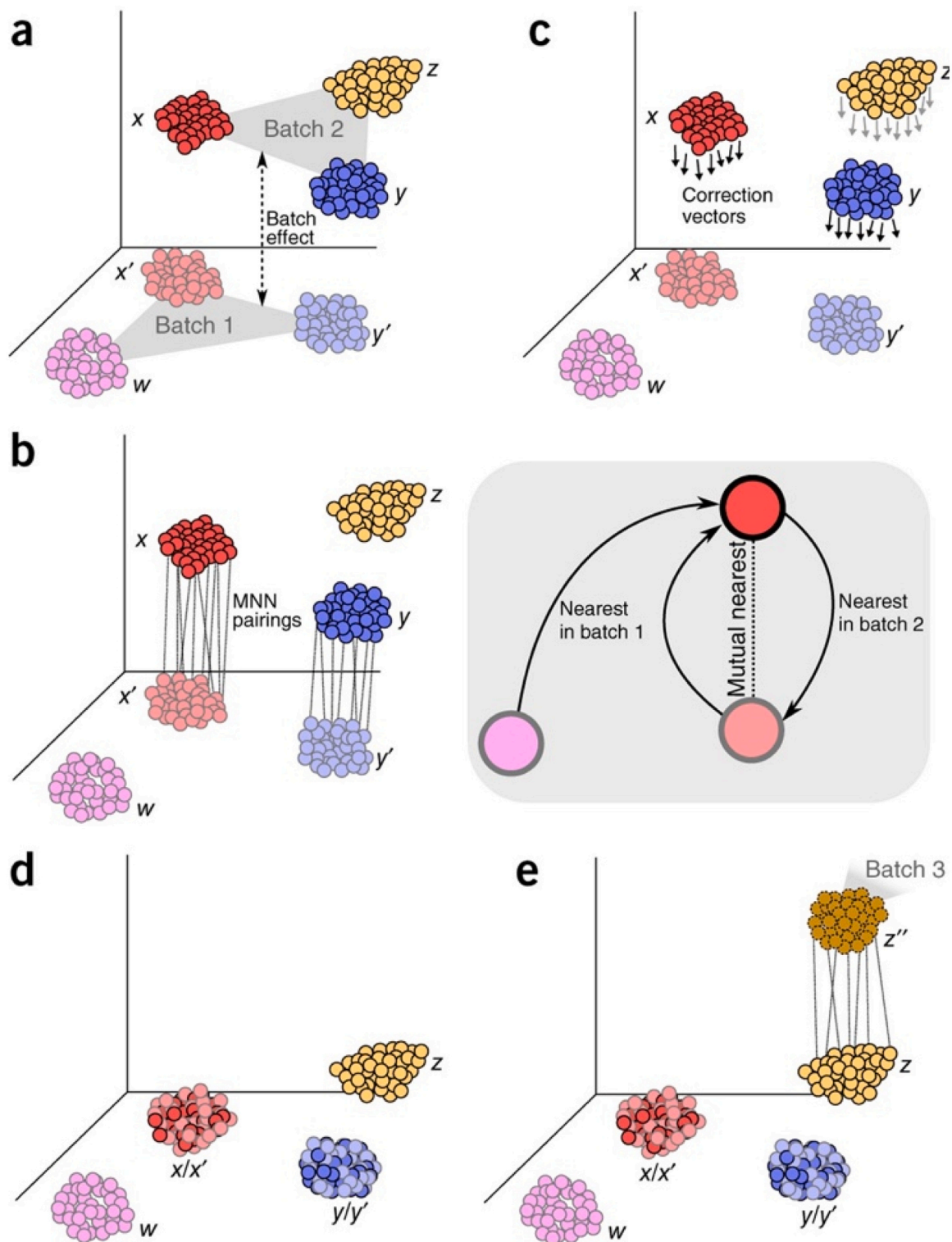


Fig. 3. Batch-effect correction via MNN. (a) Batch 1 and batch 2 in high dimensions, with a batch effect variation that is almost orthogonal. (b) By identifying MNN pairs of cells, the algorithm recognizes matching cell types (gray box). (c) Between the MNN pairs, batch-correction vectors are measured. (d) Batch 1 is considered the reference, and batch 2 is combined into it by subtracting correction vectors. (e) The integrated data is used as a reference, and the process is repeated with each new batch of data.

This figure has been reused with permission from authors [1].

computed from the resulting list of paired cells (or MNNs) to align the data sets into a shared space. The benefit of this method is that it produces a normalized gene expression matrix, which can be used in downstream analysis and offer an effective correction in the face of compositional variations between batches [1]. Scanorama [132] and batch balanced k -nearest neighbors (BBKNN) [133] are two other approaches that look for MNNs in reduced-dimension spaces, and use them in a similarity-weighted way to direct batch integration.

Hie et al. [132] proposed Scanorama which can combine and remove batch effects from heterogeneous scRNA-seq studies by identifying and merging common cell types across all pairs in a dataset. Using a variety of existing tools, Scanorama batch-corrected output can be used for downstream tasks, such as classify cluster-specific marker genes in differential expression analysis. Scanorama outperforms current methods for integrating heterogeneous datasets and it scales to millions of cells, allowing the identification of rare or new cell states through a variety of diseases and biological processes [132].

Polański et al. [133] developed BBKNN, a fast graph-based algorithm that removes batch effects through linking analogous cells in different batches. BBKNN is a simple, rapid, and lightweight batch alignment tool, and its output can be directly use for dimensionality reduction. BBKNN's default approximate neighbor mode scales linearly with the size of datasets and remains consistently faster (by one or two orders of magnitude) when compared to other existing techniques [133].

Recently, there has been considerable progress in using DL for batch effect corrections. Residual Neural Networks (ResNets)[134] and AEs are two of the most commonly used DL-based batch correction approach in scRNA-seq analysis. ResNets are a form of deep neural network that make a direct connection between the input of a layer (or network) and the outputs, often through an addition operation. Shaham et al. [135] suggested a non-linear batch effect correction approach based on a distribution-matching ResNet. Their approach focuses on reducing the Maximum Mean Discrepancy (MMD) between two multivariate replication distributions that were measured in separate batches. Shaham et al. applied their methodology to batch correction of scRNA-seq and mass cytometry datasets, finding that their model can overcome batch effects without altering the biological properties of each sample [135].

Li et al. [136] presented deep embedding algorithm for single-cell clustering (DESC), an unsupervised DL algorithm for "soft" single-cell clustering which can also remove batch effects. DESC learns a non-linear mapping function from the initial scRNA-seq data space to a low-dimensional feature space using a DNN, iteratively optimizing a clustering objective function. This sequential process transfers each cell to the closest cluster and attempts to account for biological and technical variability across different clusters. Li et al. demonstrated that DESC can eliminate the technical batch effect more accurately than MNN-based methods while better preserving true biological differences within closely related immune cells [136].

In a prior study, Shaham [137] proposed batch effect correction through batch-free encoding using an adversarial VAE. Shaham utilized the adversarial training to achieve data encoding that corresponded exclusively to a subject's intrinsic biological state, as well as to enforce accurate reconstruction of the input data. This approach results in maintaining the true biological patterns expressed in the data and minimizing the significant biological information loss [137].

Wang et al. introduced Batch Effect ReMoval Using Deep Autoencoders (BERMUDA) [138], an unsupervised framework for correcting batch effect in scRNA-seq data across different batches. BERMUDA combines separate batches of scRNA-seq data with completely different cell population compositions and amplifies biological signals by passing information between batches. Most nearest neighbor-based models can manage variations in cell populations between batches when such differences are significant. However, BERMUDA was developed with an emphasis on scRNA-seq data with distinct cell populations in mind, focusing on the similarities between cell clusters rather than significant variations. In a more recent work in 2023, Gan et al. presented

Single-Cell Integrator and Batch Effect Remover (SCIBER), a new method for removing batch effects in single-cell data. SCIBER matches cell clusters across batches using differentially expressed genes and achieves comparable or better accuracy than complex state-of-the-art methods. It follows a reference-based approach, making it suitable for integrating user-generated datasets with standard reference data like the Human Cell Atlas [139]. Altogether, a rapidly expanding number of general DL methods for batch effects correction in biological datasets represent new ways for eliminating batch effects in biological datasets.

4.1.3.3. Dimensionality reduction. Dimensionality reduction is a crucial step in visualizing scRNA-seq data, since typical datasets contain thousands of genes as features (dimensions) [140]. The most common dimensionality reduction techniques used for scRNA-seq are principal component analysis (PCA) [141], t-Distributed Stochastic Neighbor embedding (t-SNE) [142], diffusion map [143], Gaussian Process Latent Variable Models (GPLVM) [144,145], Single-cell Interpretation via Multi-kernel Learning (SIMLR), and Uniform Manifold Approximation and Projection (UMAP) [75].

In low-dimensional spaces, linear projection methods like PCA traditionally cannot depict the complex structures of single-cell data. On the other hand, nonlinear dimension reduction techniques like t-SNE and UMAP, have been shown to be effective in a variety of applications and are commonly used in single-cell data processing [146]. These methods also have some drawbacks, such as lacking robustness to random sampling, inability to capture global structures while concentrating on local data structures, parameter sensitivity, and high computational cost [13]. Several DL techniques for reducing the dimensionality of scRNA-seq data have recently been developed. Here we focus on the ones that are based on VAEs or AEs, which are more commonly used in the field.

Ding et al. [146] proposed a VAE-based model (called scvis) to learn a parametric transformation from a high-dimensional space to a low-dimensional embedding, ultimately learning the estimated posterior distributions of low-dimensional latent variables. Compared to common techniques (e.g. t-SNE), scvis can (i) better obtain the global structure of the data, (ii) provide greater interpretability, and (iii) be more robust to noise or unclear measurements. Ding et al. demonstrated that scvis is a promising tool for studying large-scale and high-resolution single cell populations [146]. However, according to Becht et al. [75], the runtime of scvis is long, particularly for dimensionality reduction, and it appears to be less effective at separating cell populations.

In another work, Wang et al. [140] proposed a method for unsupervised dimensionality reduction and visualization of scRNA-seq using deep VAEs, called VAE for scRNA-seq data (VASC). VASC's architecture consists of the traditional encoder and decoder network of a VAE, with an addition of a zero-inflated layer that simulates dropout events. In comparison to current methods such as PCA, t-SNE, and Zero Inflated Factor Analysis (ZIFA) [147], VASC can identify nonlinear patterns present within the data, and has broader compatibility as well as better accuracy, particularly when sample sizes are larger [140].

In 2020, Märtens et al. [76] proposed BasisVAE as a general-purpose approach for joint dimensionality reduction and clustering of features using a VAE. BasisVAE modified the traditional VAE decoder to incorporate a hierarchical Bayesian clustering prior, demonstrating how collapsed variational inference can identify sparse solutions when over-specifying K [76].

Peng et al. [120] proposed an AE-based model that combines gene ontology (GO) and DNNs to achieve a low-dimensional representation of scRNA-seq data. Based on this idea, they proposed two innovative approaches for dimensionality reduction and clustering: an unsupervised technique called "Gene Ontology AutoEncoder" (GOAE) and a supervised technique called "Gene Ontology Neural Network" (GONN) for training their AE model and extracting the latent layer as low dimensional representation. Their findings show that by integrating prior

information from GO, neural network clustering and interpretability can be enhanced and that they outperform the state-of-the-art dimensionality reduction approaches for scRNA-seq [77].

In a study by Armaki [148], the dimensionality reduction capabilities of VAE- and AE-based models were evaluated and benchmarked against principal component analysis. They found that the best approach for reducing the dimensionality of single-cell data was using AE-based models, while the more efficient VAEs performed worse in some respects than the linear PCA. One possible hypothesis could be that the prior used for modeling the latent space, which was Gaussian distribution, is not a good fit for single-cell data. A prior more befitting single-cell data (such as negative binomial distribution) could improve the performance of the VAE based model [148].

Finally, there is always the endeavor of optimizing algorithms. As mentioned earlier, developing an efficient DL method for dimensional reduction of data is a necessary next step, since it can potentially improve the quality of lower-dimensional representations.

4.1.3.4. In-silico generation and augmentation. Given limitations on scRNA-seq data availability and the importance of adequate sample sizes, in-silico data generation and augmentation offer a fast, reliable, and cheap solution. Synthetic data augmentation is a standard practice in various ML areas, such as text and image classification [149]. With the advent of DL, traditional data augmentation techniques (such as geometric transformations or noise injection) are now being replaced with deep-learned generative models, primarily VAEs [29] and GANs [43]. In computational genomics, both GAN- and VAE-based models have shown promising results in generating omics data. Here, we focus on the recent methods introduced for generating realistic in-silico scRNA-seq.

Marouf et al. [51] introduced two GAN-based models for scRNA-seq generation and augmentation called single-cell GAN (scGAN) and conditional scGAN (cscGAN); we collectively refer to these models as scGAN. At the time, scGAN outperformed all other state-of-the-art methods for generating and augmenting scRNA-seq data [51]. The success of scGAN was attributed to the Wasserstein-GAN [53] that learns the underlying manifold of scRNA-seq data, subsequently producing realistic never-seen-before samples. Marouf et al. showcase the power of scGAN by generating specific cell types that are almost indistinguishable from the real data and augmenting the dataset with the synthetic samples improved the classification of rare cell populations [51].

In a related work, Heydari et al. proposed a VAE-based in-silico scRNA-seq model that aimed at improving Marouf et al.'s training time, stability, and generation quality using only one framework (as opposed to two separate models). Heydari et al. proposed *ACTIVA* (Automated Cell-Type-informed Introspective Variational Autoencoder), which employs a single-stream adversarial VAE conditioned with cell-type information. The cell-type conditioning encourages *ACTIVA* to learn the distribution of all cell types in the dataset (including rare populations), which allows the model to generate specific cell types on demand. Heydari et al. showed that *ACTIVA* performs better or comparable to scGAN while training up to 17 times faster due to the design choices. Data generation and augmentation with both *ACTIVA* and scGAN can enhance scRNA-seq pipelines and analysis, such as benchmarking new algorithms, studying the accuracy of classifiers, and detecting marker genes. Both generative models will facilitate the analysis of smaller datasets, potentially reducing the number of patients and animals necessary in initial studies [42].

4.1.4. Downstream analysis

Following pre-processing, downstream analysis methods are used to derive biological understandings and identify the underlying biological mechanism. For example, cell-type clusters are made up of cells with similar gene expression profiles; minor differences in gene expression between similar cells indicate continuous (differentiation) trajectories;

or genes which expression profiles are correlated, signaling co-regulation [150].

4.1.4.1. Clustering and cell annotation. A significant phase in the scRNA-seq study is to classify cell subpopulations and cluster them into biologically relevant entities [151]. The creation of several atlas projects such as Mouse Cell Atlas [152], Aging Drosophila Brain Atlas [153], and Human Cell Atlas [154] has been initiated by advances in single-cell clustering. Different clustering approaches have emerged in recent years to help characterize different cell populations within scRNA-seq data [155]. Given that DL-based algorithms outperform traditional ML models in clustering tasks when applied to image and text datasets [156], many have turned to design supervised and unsupervised DL-based clustering techniques for scRNA-seq.

Li et al. presented DESC, an AE-based method for clustering scRNA-seq data with a self-training target distribution that can also denoise and potentially remove batch effects. In experiments conducted by Li et al. [136], DESC attained a high clustering accuracy across the tested datasets compared to several existing methods. DESC also showed consistent performance in a variety of scenarios, and did not directly need the batch definition for batch effect correction. Given that Li et al. use a deep AE to reconstruct the input data, the latent space is not regularized with additional properties that could additionally help with clustering [157]. As with most DL models, DESC can be trained on CPUs or GPUs.

Chen et al. proposed scAnCluster (Single-Cell Annotation and Clustering), an end-to-end supervised clustering and cell annotation framework that is built upon their previous unsupervised clustering work, namely scDMFK and scziDesk. ScDMFK algorithm (Single-Cell Data Clustering through Multinomial Modeling and Fuzzy K-Means Algorithm) combined deep AEs with statistical modeling. It proposed an adaptive fuzzy k-means algorithm to handle soft clustering while using multinomial distribution to describe the data structure and relying on a neural network to facilitate model parameter estimation [157]. More specifically, the AE in scAnCluster learns a low-dimensional representation of the data, and an adaptive fuzzy k-means algorithm with entropy regularization for soft clustering. On the other hand, ScziDesk aimed to learn a "cluster-friendly" lower-dimensional representation of the data. Many existing DL-based clustering algorithms for scRNA-seq do not consider distance and affinity constraints between similar cells. This prohibits such models to learn cluster friendly lower-dimensional representations of the data [157]. For scAnCluster, Chen et al. use the available cell marker information to construct a new DL model that integrates single-cell clustering and annotation. ScAnCluster can do both intra-dataset and inter-dataset cell clustering and annotation, and it also reveals a clear discriminatory effect in the detection of new cell types not found in the reference [157].

Tian et al. created the Single-Cell model-based Deep embedded Clustering (scDeepCluster) technique, which uses a nonlinear approach to combine DCA modeling and the DESC clustering algorithm. Their approach sought to improve clustering while reducing dimensions directly. ScDeepCluster outperformed state-of-the-art approaches on a variety of clustering efficiency metrics, showing runtime increasing linearly with sample size. In contrast to similar methods, scDeepCluster requires less memory and is scalable to larger datasets [158]. On the other hand, ScDeepCluster lacks the pairwise distance of associated cells and ignores the affinity constraints of similar cells. ScDeepCluster does not pre-select informative genes as input data, making the model more computationally efficient, but decreasing the clustering accuracy [157].

Peng et al. [77] developed a strategy for optimizing cell clustering based on global transcriptome profiles of genes. They used a combination of DNN and GO to reduce the dimensions of scRNA-seq data and improve clustering. Their supervised approach was based on a conventional neural network, while the unsupervised model utilized an AE. Their model consisted primarily of two main components: the choosing

of important GO factors and the combination of GO terms with the DNN-based model [77]. Single-Cell Variational Auto-Encoders (scVAE) [159] is another VAE-based model for downstream analysis scRNA-seq data. However, compared to other models, scVAE has the advantage of not requiring many of the traditional pre-processing steps since it uses the raw count data as input. ScVAE can accurately predict expected gene expression levels and a latent representation for each cell and is flexible to use the known scRNA-seq count distributions (such as Poisson or Negative Binomial) as its model assumption [159]. Wang et al. introduced a novel deep clustering algorithm called scDCCA, combining a denoising AE and a dual contrastive learning module within a deep clustering framework. scDCCA incorporates a denoising Zero-Inflated Negative Binomial model-based AE to extract low-dimensional features, enhancing data representation and learning robustness. Notably, scDCCA integrates feature learning and clustering into a seamless end-to-end process, enabling comprehensive representation learning and cell clustering simultaneously [160]. Moreover, He et al. developed G3DC, a deep embedded clustering method that combines a graph loss based on gene networks and a reconstruction loss. This approach achieves both discriminative and informative embedding, improving clustering performance and identifying functionally coherent genes. G3DC selects biologically relevant genes to provide insights into the biological functionality that differentiates cell groups [80].

More recently in 2023, Hu et al. proposed a single-cell deep fusion clustering (scDFC) algorithm based on AE. It includes an attributed feature clustering module and a structure-attention feature clustering module, capable of handling different feature types. Experimental results confirm the algorithm's efficiency in fusing attributes, structure, and attention information for scRNA-seq data [161].

Clustering challenges in scRNA-seq research are primarily reflected in an unknown number of SC clusters, unfixed cell types, and poor scalability. The quantity of cells used in scRNA-seq investigations has increased by orders of magnitude in recent years. Despite the fact that academics have created a number of tools, they are not user-friendly, employ various programming languages and need distinct input data formats [162].

Given the high noise level in scRNA seq studies, one must question if each cluster correlates to a meaningful biological effect or emerged as a result of technological artifacts, such as droplets includes two cells (doublets). Because they have a phenotype that is intermediate between the two originating cell types, doublets developing from cells of two separate cell types are often confused with uncommon transitional cells [163].

Also, Estimating technical noise in scRNA seq data is difficult since each cell is a biological, rather than a technological, duplicate. Many noise models, however, have been established using endogenous spike in RNA. By adding generated noise to data sets and repeating the clustering method, these may be used to determine the resilience of clusters [163].

The use of DL for cell clustering based on scRNA-seq data have shown improvement compared to traditional non-DL approaches. However, it is not yet clear if one approach is superior to the other in terms of downstream analysis. This is an area that could be explored in future research. Additionally, the optimal number of clusters and similarity between cell types has not been fully studied.

4.1.4.2. Cell-Cell communication analysis. In recent years, scRNA-seq has become a powerful tool for analyzing cell-cell communication in tissues. Intercellular communication controlled by ligand-receptor complexes is essential for coordinating a wide range of biological processes, including development, differentiation, and inflammation. Several algorithms have been proposed to carry out these analyses. These algorithms begin with a database of interacting molecular partners (such as ligand and receptor pairs) and predict a list of possible signaling pathways between types of cells based on their expression

patterns.

In the case of Cell-Cell communication, there are different methods which are typically used. DeepCCI [164] is the first method for investigating intercellular communications using scRNA-seq data using a DL approach. DeepCCI's significant advances is including globally propagated topological properties between L-R pairings via Graph Convolutional Networks (GCNs), as well as incorporating gene expression of cell pairs in the prediction process for intercellular communications. GCNG [165] encodes spatial data as a graph and uses supervised training to integrate it with expression levels. GCNG improves on previous approaches for analyzing spatial transcriptomics data and can identify new pairs of extracellularly interacting genes.

Cellular crosstalk based on Cell-Cell Interactions (CCIs) is the foundation of many biological processes. The improved throughput of scRNA-seq technologies allows computational inferring CCIs based on ligand-receptor information. Although several CCI techniques have been established, their findings are very dynamic, and no ground truth is supplied to measure accuracy. Evaluation of these interactions by loss-of-function assays or high-throughput screening methods is labor-intensive and time-consuming. There is a critical need for dependable and cost-effective standards for assessing CCIs [81].

4.1.4.3. RNA velocity. RNA velocity has created new ways to research cellular differentiation in scRNA-seq studies. RNA velocity represents the rate of change in gene expression for a single gene at a given time point depending on the spliced ratio to unspliced mRNA. In other words, RNA velocity is a vector that forecasts the possible future of individual cells on a timescale of hours, which proves new insights into cellular differentiation that have challenged conventional and long-standing views. Existing experimental models for tracking cell fate and reconstructing cell lineages, such as genetic methods or time-lapse imaging have limited power as they do not show the trajectory of differentiation or reveal the molecular identity of intermediate states. But recently, with the advent of scRNA-seq, a variety of algorithms such as VeloViz [166] or scVelo [167] can visualize the velocity estimates in low dimensions. Although such models show promising results (particularly in well-characterized systems), we are far from a complete understanding of cellular differentiation and cell fate decisions. By this means, recently Cui et al. presented DeepVelo, a DNN-based approach that is empowered by deep GCN and reduced the restrictions of existing methods. DeepVelo can predict RNA cellular velocities without pre-defined kinetic patterns, as well as RNA velocities for high-complexity dynamics, especially for cell populations with many lineages and heterogeneous cell types [168].

Current RNA velocity techniques include size factors proportional to number depth per cell, as well as variants on this theme. Nevertheless, cell size also reflects the natural expansion of the RNA transcription reservoir. It is unclear how to appropriately account for cell count depth, whether to normalize intronic and exonic matrices to matrix-specific variables, shared factors, or not at all. More broadly, we should look at how changes in global cellular factors, such as splicing efficiency or RNA polymerase abundance, impact the kinetic models. Normalization by cell size is a straightforward method for removing the impacts of count sampling, but it can distort these effects in a significant way. Accurate velocity estimations need adequate preprocessing and, ideally, the incorporation of these effects into the model [169]. We foresee the development of more DL-based models, to resolve long-standing concerns regarding cell fate choices and lineage specification using RNA velocity.

4.2. Deep learning in single-cell genomics

Traditional sequencing methods are limited to measuring the average signal in a group of cells, potentially masking heterogeneity and rare populations [170]. On the other hand, scRNA-seq technologies provide a tremendous advantage for investigating cellular heterogeneity

and recognizing new molecular features correlated with clinical outcomes, resulting in the transformation of many biological research domains. Similar to scRNA-seq, single-cell (SC) genomics is being used in many areas, such as predicting the sequence specificity of DNA- and RNA-binding proteins, enhancer and cis-regulatory regions, methylation status, gene expression, control splicing, and searching for associations between genotype and phenotype. However, SC genomics data are often too large and complex to be analyzed only through visual investigation of pairwise correlations. As a result, a growing number of studies have leveraged DL techniques to process and analyze these large datasets. In addition to the scalability of DL algorithms, another advantage of DL techniques is the learned representations from raw input data, which are beneficial in specific SC genomics and epigenomics applications. These application include cell-type identification, DNA methylation, chromatin accessibility, TF-gene relationship prediction, and histone modifications. In the following sections, we review some applications of DL models for analyzing scDNA-seq data (summarized in Fig. 4). Many genomics data can be accessed from public repositories, including NCBI (<https://www.ncbi.nlm.nih.gov/genome>), Ensembl (<https://www.ensembl.org/>), and DDBJ (<http://www.ddbj.nig.ac.jp>).

4.2.1. Cell type identification in CyTOF

An essential and challenging task in genomics research is to accurately identify and cluster individual cells into distinct groups of cell types. Li et al. [171] describe AE methods (stacked AE and multi-AE) as a

gating strategy for mass cytometry (CyTOF). CyTOF is a recent technology for high-dimensional multiparameter SC analysis. They introduced DeepCyTOF as a standardization procedure focused on a multi-AE neural network. DeepCyTOF focuses on domain adaptation principles and is a generalization of previous work that helps users in calibrating between a source domain distribution (reference sample) and several target domain distributions (target samples) in a supervised manner. Moreover, DeepCyTOF requires labelled cells from only a single sample. DeepCyTOF was applied to two CyTOF datasets produced from primary immune blood cells: (a) cases with a history of West Nile virus (WNV) infection and (b) normal cases of various ages. They manually gated a single baseline reference sample in each of these datasets to automatically gate the remaining uncalibrated samples.

Li et al. revealed that DeepCyTOF's cell classification was very consistent with classifications done by individual manual gateings, with over 99% concordance. Additionally, they used a stacked AE (which is one of DeepCyTOF's key components) to tackle the semi-automated gating challenge of the FlowCAP-I competition. Li et al. found that their model outperformed other existing gating approaches benchmarked on the fourth challenge of the competition. Overall, stacked AEs combined with a domain adaptation technique suggest promising results for CyTOF semi-automated gating and flow cytometry data, requiring manual gating of one reference sample to precisely gate the remaining samples.

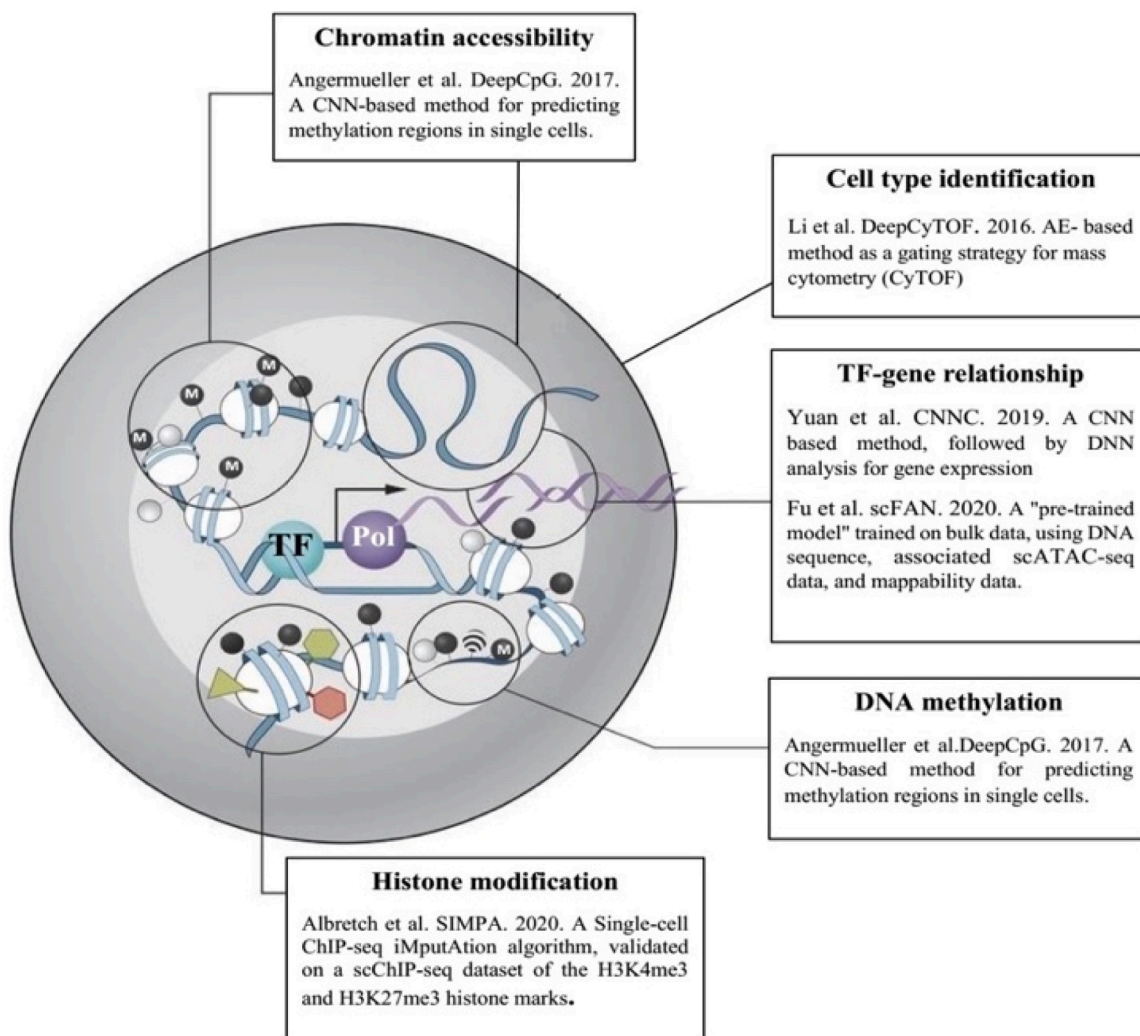


Fig. 4. Overview of algorithms that are used in different parts of single-cell genomics analysis.

4.2.2. DNA methylation

Recent technological advances have made it possible to assay DNA methylation at the single-cell resolution. Angermueller et al. propose DeepCpG [171], a CNN-based computational method for predicting methylation regions. DeepCpG is made up of three modules: a DNA module that extracts features from the DNA sequence, a CpG module that extracts features from the CpG neighborhood of all cells, and a multi-task joint module that integrates evidence from both modules for predicting the methylation regions of target CpG sites for different cells. The trained DeepCpG model can be used in various downstream studies, such as inferring low-coverage methylation profiles for groups of cells and recognizing DNA sequence motifs linked to methylation states and cell-to-cell heterogeneity. Angermueller et al. apply their model to both mouse and human cells, which achieves significantly better predictions than previous techniques [171]. An overview of the DeepCpG model is shown in (Fig. 5).

Interestingly, DeepCpG can be used for differentiating human induced pluripotent stem cells in parallel with transcriptome sequencing to specify splicing variation (exon skipping) and its determinants. Linker et al. [172] presented that variation in SC splicing can be precisely predicted based on local sequence composition and genomic features. DeepCpG, which is used for DNA methylation profiles, imputes unobserved methylation regions of individual CpG sites. The

cell-type-specific models were made using CpG and genomic information according to DeepCpG's setup of a joint model. Finally, during cell differentiation, Linker et al. identified and characterized associations between DNA methylation and splicing changes, which led to indicating novel insights into alternative splicing at the SC level.

Another method for studying chromatin accessibility at single-cell resolution, called Assay of Transposase Accessible Chromatin sequencing (scATAC-seq), has recently gained considerable popularity. In scATAC-seq, mutation-induced hyperactive Tn5 transposase tags and fragments regions in open chromatin sites in the DNA sequence, which is later sequenced using paired-end Next Generation Sequencing (NGS) technologies [173]. The pre-processing steps of scATAC-seq data analysis are often analogous to scRNA-seq pipelines. That is, the same tools are often used in both data modalities, although scRNA-seq tools are often not optimized for the particular properties of scATAC-seq data. scATAC-seq has low coverage, and the data analysis is highly sensitive to non-biological confounding factors. The data is pre-processed and assembled into a feature-per-cell matrix, where common choices for "feature" are fixed-size genomic bins and signal peaks at biological events. This matrix displays particular numerical properties which entail computational challenges: it is extremely high-dimensional, sparse, and near-binary in nature (presence/absence of signal). Several packages have been recently developed specifically for scATAC-seq data, with all

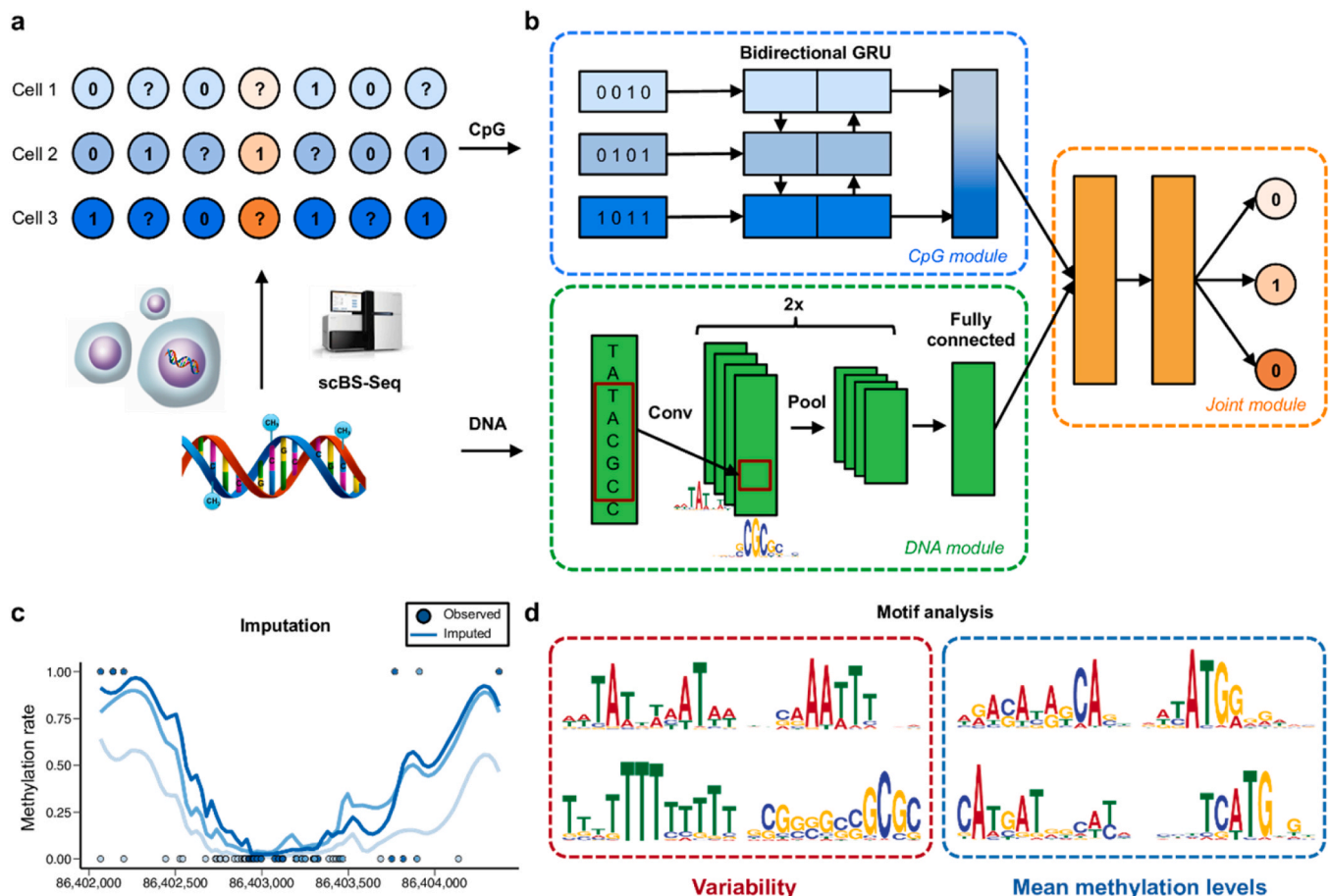


Fig. 5. An overview of DeepCpG (taken from Angermueller et al.) (a) scBS-seq (Smallwood et al.) scRRBS-seq (Farlik et al.) provide sparse single-cell CpG profiles, which are then pre-processed as binary values (methylated CpG sites are marked by ones, unmethylated CpG sites are denoted by zeros, and CpG sites with unknown methylation state [missing data] are represented by question marks) (b) Two convolutional and pooling layers are used to detect predictive motifs from the local sequence context, and one fully connected layer is used to model motif interactions in the DNA model. Applying a bidirectional gated recurrent network (GRU), the CpG model scans the CpG neighborhood of numerous cells (rows in b), providing compressed information in a vector of constant size. To forecast methylation states in all cells, the Joint model learns connections between higher-level properties generated from the DNA- and CpG models (c, d) The DeepCpG model can be utilized for a variety of downstream tasks after training, such as genome-wide imputation of missing CpG sites (c) and the finding of DNA sequence motifs linked to DNA methylation levels or cell-to-cell variability (d).

of them having some major limitations [174]. The tools for processing scATAC-seq are diverse and nascent, and thus, there is no consensus on the best practices scATAC-seq data analysis. Further development of scATAC-centric computational tools and benchmark studies are much needed. scATAC-seq analyses can help to elucidate cell types and differentially accessible regions across cells. Moreover, it can decipher regulatory networks of cis-acting elements like promoters and enhancers, and trans-acting elements like transcription factors (TFs), and infer gene activity [175]. scATAC-seq data could also be integrated with RNA-seq and other omics data. However, most current software only integrates the derived gene activity matrix with expression data, and important information from whole-genome chromatin accessibility is lost.

Currently, there are a variety of DL models for bulk ATAC-seq data, such as LanceOtron's CNN for peak calling [176] and CoRE-ATAC for functional classification of cis-regulatory elements [177]. Thibodeau et al. demonstrated CoRE-ATAC's transferable capability on cell clusters inferred from single nuclei ATAC-seq data with a small decrease in model prediction accuracy (mean micro-average precision of 0.80 from bulk vs. 0.69 in single-cell clusters).

One common way to reduce the dimensionality of scRNA-data is to identify the most variable genes (e.g. with PCA), since they carry the most biologically relevant information. However, scATAC-seq data is binary, and therefore prohibiting the identification of variable peaks for dimensionality reduction. Instead, dimensionality reduction of scATAC data is done through Latent Semantic Indexing (LSI), a technique used for natural language processing. Although this approach is scalable to large number of cells and features, it may fail to capture the complex reliance of peaks, since LSI is a linear method.

Single-Cell ATAC-seq analysis via Latent feature Extraction (SCALE) [170] combines a deep generative framework with a probabilistic Gaussian Mixture Model (GMM) as a prior over the latent variables, in order to learn a nonlinear latent space of scATAC-seq features. Given the nature of scATAC-seq, GMM is a suitable distribution for modeling the high-dimensional, sparse multimodal scATAC-seq data. SCALE can also be used for denoising and imputing missing data, recovering signals from missing peaks. In the benchmarking study by Xiong et al., they demonstrated that SCALE outperformed conventional non-DL scATAC-seq tools for dimensionality reduction, such as PCA and LSI. Furthermore, they show that SCALE is scalable to large datasets (on the order of 80,000 single cells). While SCALE succeeds at learning nonlinear cell representations with higher accuracy, it assumes that the read depth is constant across cells and ignores potential batch effects. These pitfalls motivated the development of scalable and accurate invariant representation learning scheme (SAILER) [178].

SAILER is a deep generative model inspired by VAEs which also learns a low-dimensional latent representation of each cell. For SAILER, the authors aimed to design an invariant representation learning scheme, where they discard the learned component associated with confounding factors from various technical sources. SAILER captures nonlinear dependencies among peaks, faithfully separating biologically relevant information from technical noise, in a manner that is easily scalable to millions of cells (when using GPUs). Similar to SAILER, SCALE also offers a unified strategy for scATAC-seq denoising, clustering, and imputation. However, in multi-sample scATAC-seq integration, SAILER can eliminate batch effects and properly recreate a chromatin accessibility landscape free of confusing variables, regardless of sequencing depths or batch effects.

4.2.3 Transcription Factor (TF)-gene relationship prediction. To unravel gene regulatory mechanisms and differentiate heterogeneous cells, understanding the genome-wide binding TF profile is crucial. Researchers have developed several methods that use expression data for infer gene-gene interactions such inferring coexpression, understanding functional assignments and reconstructing pathways [179]. However, each task in inferring gene-gene relationships is typically done using different techniques. DL-based methods, on the other hand, are

capable of learning multiple tasks jointly, which is very advantageous in inferring relationships between genes. Yuan et al. [179] introduced Convolutional Neural Network for Co-Expression (CNNC), a new encoding method for gene expression data based on CNNs. CNNC is a general computational technique for supervised gene relationship inference that builds on previous approaches in various tasks, including predicting TF targets and recognizing genes related to disease in order to infer cause and effect. The key idea behind CNNC is to turn data into co-occurrence histograms (as images), and then analyzing the histograms using CNNs. More specifically, Yuan et al. generated a histogram for each pair of genes and utilized CNNs to infer relationships among the various levels of expression encoded in the image. CNNC is adaptable and can easily be expanded to integrate other types of genomics data as well, resulting in additional performance gains. CNNC goes beyond previous approaches for predicting TF-gene and protein-protein interactions and predicting the pathway of a regulator-target gene pair. CNNC may also be used to draw causality inferences, functional assignments (such as biological processes and diseases), and as part of algorithms that recreate known pathways [179].

In another work, Fu et al. [180] presented Single Cell Factor Analysis Network (scFAN), a DL model for determining genome-wide TF binding profiles in individual cells. The scFAN pipeline consists of a "pre-trained model" trained on bulk data and then used to predict TF binding at the cellular level using DNA sequence, aggregated associated scATAC-seq data, and mappability data. ScFAN can help overcome the basic sparsity and noise constraints of scATAC-seq data. This model provides a valuable method for predicting TF profiles through individual cells and could be applied to analyze SC epigenomics and determine cell types. Fu et al. presented scFAN's ability to identify cell types by analyzing sequence motifs enriched within predicted binding peaks and studying the effectiveness of predicted TF peaks. They suggested a novel metric called "TF activity score" to classify each cell and demonstrated that the activity scores could accurately capture cell identity. Generally, scFAN is capable of connecting open chromatin states with transcript factor binding activity in individual cells, which is beneficial for a deeper understanding of regulatory and cellular dynamics [180].

4.2.3. Histone modification

Given the effects of protein-DNA interactions between histone marks and TF on the regulation of crucial cellular processes (including the organization of chromatin structures and gene expression), the identification of such interactions is highly significant in biomedical science. Chromatin immunoprecipitation followed by sequencing (ChIP-seq) is a widely used technique for mapping TFs, histone changes, and other protein-DNA interactions for genome-wide mapping [181]. ChIP-seq data is very sparse, therefore generally requiring imputation for more accurate analysis. Albretsch et al. introduced Single-cell ChIP-seq iMPutAtion (SIMPA) [182], an imputation algorithm that was tested on a single-cell ChIP-seq (scChIP-seq) dataset of the H3K4me3 and H3K27me3 histone marks in B-cells and T-cells. Unlike most SC imputation approaches, SIMPA integrates the sparse input of one single cell with a series of 2251 ENCODE ChIP-seq experiments to extract predictive information from bulk ChIP-seq data. SIMPA's goal is to identify statistical patterns that bind protein-DNA interacting sites through specific SC regions of target-specific ENCODE data for different cell types, as well as the presence or absence of potential sites for a single cell. Once the patterns are identified, SIMPA's DL models uses these patterns to make precise predictions. As a new approach in sc-seq, SIMPA's imputation strategy augmented sparse scChIP-seq data, leading to improved cell-type clustering and the detection of pathways that are specific to each cell type.

4.2.4. CNV and SNP prediction

A genome-wide association study (GWAS) is the standard method for identifying genetic biomarkers, such as single nucleotide polymorphisms (SNPs) and copy number variations (CNVs), that are linked

to numerous characteristics and illnesses. CNVs are DNA segments that are duplicated or deleted in relation to the reference genome and are greater than 1000 base pairs (bp) [183]. GWAS, on the other hand, has limitations due to its individual-CNV and SNP analysis technique, which is compounded by the high dimensionality of genomic datasets. When doing several individual association tests, strict thresholds must be used to allow for mistake rates that contribute to weak detection [184]. Several investigations have already shown that DL can be used in genomic datasets; nevertheless, to the best of our knowledge, there is just a limited literature review detailing its use to CNV-set construction. Several approaches for detecting CNV from single-cell DNA data have been developed, including *dudeML* [185], *DeepCNV* [186], *RDAClone* [187], and *rcCAE* [188].

dudeML is a DL-based method that detects CNVs pretty successfully, particularly in samples with low coverage or in pools, utilizing data easily provided from the sample [185]. Standard procedures for finding CNVs, on the other hand, rely on CNV tool predictions followed by tedious and heuristic filtering. When the sample size is big, however, manually distinguishing false positives from true positives becomes more expensive. Furthermore, manual filtering may be subjective and may induce bias, whereas in clinical settings workflow rigour and goals are desirable. *DeepCNV* is intended to address this void by automating the filtering process in a transparent and impartial manner. To improve distinguishability, *DeepCNV*'s hierarchical structure compresses the original data from intractable big dimensions into a viable and informative representation [186]. *RDAClone* is a nonlinear DL-based AE that has been shown to be a successful tool for analyzing genomics data. It may also be tweaked to assess cellular and gene networks for potential uses in biology and medicine [187].

The *rcCAE* technique is an AE-based approach for calculating SC copy numbers from scDNA-seq data. The approach employs a convolutional AE to learn latent representations of SCs while also improving read count quality, resulting in an integrated framework that enables efficient characterization of underlying cell subpopulations in a low-dimensional latent space, as well as accurate detection of SC copy numbers from improved read counts by effectively eliminating non-biological confounding factors [188].

SC genomics analysis through DL is a promising and emerging field with an incredible potential to advance our knowledge of fundamental biological matters. In this respect, DL can provide us with a better understanding of nature, and the intricacies of DNA structure and epigenomics effects on human diseases for both therapeutic and diagnostic purposes. Due to intrinsic challenges in SC genomics, such as sparsity, systematic noise, and higher-dimensionality of biological systems, developing new DL models are paramount to further advancing the SC genomics field.

4.3. Deep learning in spatial transcriptomics

Since being named the method of the year [189], spatial transcriptomics (ST) is becoming the natural extension of scRNA-seq, unbiasedly profiling transcriptome-wide gene expression. By not requiring tissue dissociation, spatial transcriptomics retain spatial information (adding a spatial component to conventional RNA-seq technologies). ST have the potential of revolutionizing the field by bridging the gap between the deep characterization of cellular states and the cellular diversity that constitutes tissue organization. Spatially resolved transcriptomics can provide the genetic profiles of cells while containing information about the positional distribution of the sequenced cells, enhancing our understanding of cell interactions, organ function and pathology. However, the high-throughput spatial characterization of complex tissues remains a challenge. Broadly, spatially-resolved transcriptomics techniques can be divided into two categories: (i) the cyclic RNA imaging techniques that achieve single-cell resolution and (ii) array-based spatially resolved RNA-seq techniques, such as *Visium Spatial Transcriptomics* [190], *Slide-sequencing* [191], or

high-definition spatial transcriptomics (HDST) [192]. Though both subgroups can provide spatial information on a single-cell level, the cyclic RNA methods are limited on the number of genes that they can multiplex. On the other hand, the array-based spatially resolved RNA-seq techniques achieve high-throughput data by capturing mRNA across thin tissue sections using a grid of microarrays or bead-arrays, relying on simple molecular biology and histology protocols. However, since array-based mRNA capture does not match cellular boundaries, spatial RNA-seq measurements are a combination of multiple cell-type gene expressions that can correspond either to multiple cells (*Visium*) or fractions of multiple cells (depending on the spatial resolution of each method).

To obtain a comprehensive characterization of underlying tissue, there is the need for computational methods that can produce coupled single-cell and spatially resolved transcriptomics strategy, mapping cellular profiles into a spatial context. Nonetheless, there are some techniques for retrieving relevant biological information from ST data. As remarked by Lähnemann et al. [193], detecting spatial gene expression patterns is one of the most pressing challenges in single-cell omics data science. Identifying such patterns can provide valuable insight on the spatial distribution of cell populations, pointing out gene marker candidates and potentially leading to identification of new rare cell subpopulations. Moreover, ST not only puts gene expression into a spatial context but also facilitates the integration of tissue-image information with gene expression information. Such data integration will enable researchers to utilize image processing techniques to investigate the morphological information, gaining more intuition in order to obtain more refined inferences, predictions, or cellular profiles. After addressing the mapping issues from gene expression to the spatial coordinates, further computational tools will be required to study cell-cell interactions within tissues and to model transcriptional relations between cell types [194].

The ST field is growing fast, with new datasets and analysis pipelines presented weekly. Although the space of DL models for ST analysis is now small, we anticipate the field to experience a paradigm shift toward DL models. In this way, Xu et al. presented *DeepST*, a DL framework for recognizing spatial domains that outperform existing approaches for benchmarking human dorsolateral prefrontal cortex datasets [195]. Further testing on the breast cancer, ST dataset demonstrated that *DeepST* can analyze the spatial domains of cancer tissues at a better scale. In addition, *DeepST* can not only efficiently batch integrate ST data generated from multiple batches or different technologies but is also scalable to handle other spatial omics data. Their results determined that *DeepST* has a special ability to identify spatial domains, making it a necessary tool for obtaining new visions from ST studies. Maseda et al. developed an algorithm for imputing spatial information onto a scRNA-seq dataset from a given spatial reference atlas and named it *DEEPsc* which requires a reference atlas [196]. They used *DEEPsc* for spatial imputation of four different biological systems (*Zebrafish*, *Drosophila*, *Cortex*, and *Follicle*), reaching precision to some existing models while maintaining higher accuracy and robustness. A weakness of *DEEPsc* is its training time, which depends non-linearly on the number of available locations. However, this issue could be diminished by considering a small subset of possible locations when training the model, or by using a more optimized design.

In 2021 Biancalani et al. introduced *Tangram*, a DL framework that addresses two challenges: Learn transcriptome-wide spatial gene expression maps at sc resolution and link them to histological and anatomical information from the same sample [197]. In another word, *Tangram* can align scRNA-seq or single-nucleus RNAseq profiles to spatial data. Biancalani et al. revealed that *Tangram* learns a precise mapping among the spatial data and scRNAseq gene expression after using fine- or coarse-grained spatial atlases. *Tangrams* can offer a variety of advantages based on spatial techniques, but they can generate reliable spatial mappings and overcome resolution or throughput limitations. This is helpful in many ST studies.

Tan et al. introduced SpaCell, a user-friendly DL tool that can combine thousands of gene expression measures from spatially barcoded areas in tissue with millions of pixel intensity values. To develop DL models that can forecast labels of tissue images with high resolution and accuracy or recognize cell types, their integration technique performs better than the use of gene-count data alone or imaging data alone [198].

The SEDR is an unsupervised spatially embedded deep representation of the spatial and transcriptional data that Fu et al. have expressed. With the help of deep AE and a variation graph AE, the SEDR pipeline creates a low-dimensional latent representation of gene expression that is co-embedded with the corresponding spatial data. They enhanced clustering precision using SEDR on data from the human dorsolateral prefrontal cortex, and they successfully followed the prenatal brain's growth in trajectory analysis [199]. Recently O'Connor et al. introduced DeLTA 2.0, a Python-based workflow that enables fast and precise analysis of images of individual cells on flat surfaces to measure gene expression and cell growth. The algorithm employed in DeLTA 2.0 utilizes DCNN to extract relevant information about single cells from time-lapse images, eliminating the need for human intervention once the system is trained [200].

Given the recency of ST, ML- and DL-based models for studying this type of data are rare and not fully developed. Given the complexity of the ST space, however, we predict that DL models will be the predominant method of choice for ST data integration and analysis, with perhaps many new models borrowing ideas and designs from the existing body of work in computer vision.

4.4. Integrating scRNA-seq and spatial transcriptomics (spot deconvolution)

Given that ST approaches normally detect mRNA expression from a mixture of cells, and that they do not distribute the sequenced samples to match cellular boundaries, it is imperative to integrate ST data with scRNA-seq data to obtain a comprehensive map. Using scRNA-seq as a reference, this integration aims to infer which cells belong to which gene expression counts detected at the various location across the tissue. Thus far, traditional ML and statistical models have shown promising results in this space. For example, SPOTlight [201] is a model for deconvoluting spatial transcriptomics spots with single-cell transcriptomes using seeded NMF regression. SPOTlight starts by obtaining cell-type-specific profiles that are representatives of the gene expressions associated with the different cells. This process is further refined by a proper initialization of the method that uses unique marker genes of specific cell types. Next, the method uses Non-Negative Least Squares to deconvolute the captured expression of each spot (location). Despite SPOTlight's accuracy and computational efficiency, SPOTlight lacks the flexibility of integrating datasets from different batches or sequencing technologies. In SPOTlight, several aspects of the underlying biological and technological variance in the data are not addressed nor accounted for, which limits its application. Furthermore, technical procedures to obtain RNA-seq data are notoriously delicate and subjected to notable sources of variation.

Methods like Seurat 3 [202] attempt to account for the intrinsic technical variability of RNA-seq procedures thorough developing an "anchor"-based method for integrating datasets. However, it is important to remember that there are biological variabilities in the number of cells across positions or the amount of mRNA expressed by each cell or cell type. Therefore, there is an unmet need for techniques that can integrate datasets while properly accounting for the various sources of variability. Several studies have proposed statistical frameworks as viable candidates to account for the variability when integrating different datasets. For example, Stereoscope [203] build their model on a statistical framework. They model gene expressions counts as occurrences under a negative binomial distribution. Stereoscope follows previous approaches of obtaining a gene expression profile for each cell

type. That is, Andersson et al. follow a two-step approach: First, they estimate the parameter of the negative binomial distribution for all genes within each cell type. Similar parameters for a distribution of the RNA-seq and spatial expression mixture are then formed by a linear combination of the single-cell parameters. The next step is to search a set of weights that can best fit the spatial data [203]. These computed weights reflect the contribution of each cell type to the gene expression counts found in each location, thus explaining the abundance of each cell type across the spots.

Following the previous methodology, cell2location [204] builds on a Bayesian framework and models gene expression counts as a negative binomial distribution. This approach enables controlling the sources of variability, which is crucial when working with data from different technologies. Cell2location integrates scRNA-seq information into the spatial model using a similar statistical framework as Stereoscope: In addition to modeling the gene-specific unobserved rate (mean) as a weighted sum of the cell signature gene expression, cell2location also adds various parameters to provide the model with prior information regarding technology sensitivity.

That is, to further improve integration between different technologies, Kleshchevnikov et al. allow for four hyperparameters that will scale the weighted sum of cell contributions. Kleshchevnikov et al. show that cell2location integration results while being more computationally efficient than competing models (such as Stereoscope).

In another study by Lopez et al. a Deconvolution of Spatial Transcriptomics profiles using Variational Inference (DestVI), was introduced [205]. Their algorithm was applied for multi-resolution analysis used for spatial transcriptomics that certainly represents continuous variation inside cell types. They showed that DestVI provides greater resolution compared to current approaches and could estimate gene expression in each cell type at all spots. Additionally, when applying DestVI to a mouse tumor model, it can precisely deconvolve spatial organization, as well as provide a high-resolution and precise spatial characterization of tissue cellular organization.

4.5. Deep learning in the integration of single-cell multimodal omics data

Single-cell sequencing (sc-seq) was chosen as Method of the Year in 2013 due to its ability to sequence DNA and RNA in individual cells [206]. Sc-seq allows for gene expression measurements at an unprecedented single-cell resolution, which can provide a comprehensive view of the genome, transcriptome, or epigenome. In addition, recent technological advances now allow for multimodal omics measurements from the same experiment. In order to gain an accurate and comprehensive view of the cellular composition of the control (normal) and disease groups, it is essential to integrate all omics data for one sample simultaneously [207]. The various omics technologies can assess various modalities in one experiment (i.e. conduct multimodal studies) or integrate diverse omics datasets from multiple experiments.

Given the tremendous potentials of these approaches, single-cell multimodal omics was named the 2019 Method of the Year [206]. Omics integration holds the promise of linking even small datasets across orthogonal biochemical domains, amplifying biologically significant signals in the process [208]. By analyzing multi-omics data, researchers can produce novel hypotheses or design mathematical algorithms for prediction tasks, such as drug sensitivity and efficacy, gene dependence prediction, and patient stratification. Hao et al. [209] proposed a non-DL based method, called the weighted-nearest neighbor method which has shown promising results. WNN is an unsupervised framework for defining cellular identity by leveraging multiple data types for creating a multimodal reference atlas. Hao et al. apply their technique to a dataset of human PBMCs that includes paired transcriptomes and measurements of 228 surface proteins, forming a multimodal immune system atlas. They evaluate multimodal datasets from single cells, including paired measurements of RNA and chromatin state, and extend beyond the transcriptome to define cellular identity in

a coherent and multimodal manner [209].

Another non-DL-based method was presented by Clark et al., a sc approach for parallel chromatin accessibility, DNA methylation, and transcriptome profiling, and introduced its single-cell nucleosome, methylation, and transcription sequencing (scNMT-seq). GpC methyltransferase is used to tag open chromatin in scNMT-seq, which is then followed by bisulfite and RNA sequencing [210].

Although traditional ML models such as WNN have shown promising results, multi-omics data pose unique challenges for holistic integration, in addition to the other traditional difficulties such as batch effects from multiple sources. Multi-omics data reflect molecular phenotypes at different molecular systems, and thus each omics dataset could follow a different and specific distribution. To overcome these hurdles, sophisticated statistical and computational strategies are required. Among the various proposed algorithms thus far, only DL-based algorithms provide the computational versatility necessary to effectively model and incorporate virtually any form of omics data in an unsupervised or supervised manner [208].

Most DL-based algorithms in this area aim to simultaneously calculate several modalities in a single experiment. For example, Zuo et al. [211] introduced a single-cell multimodal VAE model (scMVAE) for profiling both transcriptomic and chromatin accessibility information in the same individual cells. Given the scRNA-seq and scATAC-seq of the same individual cells, scMVAE's uses three joint-learning strategies to learn a non-linear joint embedding that can be used for various downstream tasks (e.g. clustering). This joint learning distinguishes scMVAE from other VAE-based models (such as scVI) which process individual omics data separately. Zuo et al. note that scMVAE's feature embeddings are more distinct than the scVI's for each omics data, indicating that the joint learning representation of multi-omics data will produce a more robust and more valuable representation [211].

Currently, there are only a few studies that have used DL for data integration, but their success thus far calls for additional investigation of DL models in this domain [212]. Despite the significant advances made with single-cell multimodal omics technologies, several obstacles remain: First, these techniques are prohibitively expensive when used on a large scale to analyze complex heterogeneous samples and distinguish rare cell types within a tissue. On the other hand, data sparsity is a significant limitation of high-throughput single-cell multimodal omics assays. Furthermore, existing methods cover only a small portion of the epigenome and transcriptome of individual cells, making it challenging to separate technical noise from cell-to-cell variability. While future modification of these approaches will eventually close the gap, fundamentally new algorithms or strategies may be needed to resolve this constraint completely [213]. Amodio et al. [214] propose Manifold-Aligning GAN (MAGAN), which is a GAN-based model which aligns two manifolds coming from different domains with the assumption that different measurements for the same underlying system contain complementary information. They show MAGAN's potential in the problem of single-cell data integration (CyTOF and scRNA-seq data) and demonstrate the generalization potential of this method in the integration of other data types. However, the performance of MAGAN decreases in the absence of correspondence complementary information among samples [215]. Cao et al. introduced UnionCom for unsupervised topological alignment of single-cell omics integration without a need for correspondence information among cells or among features, which can be very useful in data integration. However, UnionCom is not scalable to large datasets that are on the order of millions of cells [216]. Other models such as SMILE (Single-cell Mutual Information Learning) also allow for unmatched feature types. SMILE is a deep clustering algorithm for different tissues and modalities, even when the feature types are unmatched. SMILE can remove batch effects and learn a discriminative representation for data integration using a cell-pairing maximization algorithm [217].

Single-Cell Data Integration via Matching (SCIM) [218] is another deep generative approach that constructs a technology-invariant latent

space to recover cell correspondences among datasets, even with unpaired feature sets. The architecture is a modified auto-encoder with an integrated discriminator network, similar to the one in GANs, allowing the network to be trained in an adversarial manner. Multi-modal datasets are integrated by pairing cells across technologies using a bipartite matching scheme that operates on the low-dimensional latent representations [218]. Another data integration model is integrative analysis of multi-omics at single-cell resolution (GLUER) [219], which employs three computational approaches: a nonnegative matrix factorization (NMF), mutual nearest neighbor, and a DNN to integrate multi-omics data. The NMF stage helps to identify shared signals across data sets of different modalities, resulting in "factor loading matrices" (FLM) for each modality. The FLM from one data modality is defined as the reference matrix while the other FLM is used as query matrices, which are used to calculate putative cell pairs. These cell pairs are then used in a DNN which tries to learn a mapping between the query FLMs and the reference FLMs, resulting in co-embedded datasets [219].

Donno et al. proposed single-cell population-level integration (scPoli) in 2022 as a semi-supervised conditional deep generative algorithm for data integration, label transfer, and query-to-reference mapping. Unlike other algorithms, scPoli can integrate and annotate newly created query datasets while developing an uncertainty mechanism for recognizing unknown populations. It also learns both sample and cell representations [220]. For combining sparse and high-dimensional scATAC-seq data, Kopp et al. created BAVARIA. Their method tackles the lack of a particular scATAC-seq tool that permits simultaneous data integration from several complex datasets and dimensionality reduction [221]. They announced that their algorithm is the first one to enable data integration and precise batch correction across various scATAC protocols.

SciPENN, is a multi-use DL approach presented by Lakkis et al. Their method supports CITE-seq (matched RNA and protein profilings) and scRNA-seq data integration, protein expression prediction for scRNA-seq and protein expression imputation for CITE-seq, as well as the transfer of cell type label from CITE-seq to scRNA-seq [222].

BABEL was presented by Wu et al. an algorithm that integrates the chromatin profiles and transcriptome of a single cell using a DNN model. After being trained on relevant data, BABEL can accurately predict sc expression from a cell's scATAC-seq and vice versa. Protein epitope profiling is one of the extra sc data modalities that BABEL is also capable of incorporating. This enables translation across chromatin, RNA, and protein [86]. Ma et al. introduced DeepMAPS, a DL-based Multi-omics Analysis Platform. scRNA-seq, CITE-seq, matched scRNA, and ATAC-seq data sets can all be analyzed collectively using this method. The key component of the method is a transformer with a graph attention mechanism and a heterogeneous graph representation of cell-gene connections.

VAEs and attention transfer were combined to create the deep cross-omics cycle attention (DCCA) model, which Zuo et al. suggested for the integration analysis of sc multi-omics data. The DCCA model learned a coordinated but independent representation for each omics data given the raw scRNA-seq or scEpigenomics data as input.

Total Variational Inference (totalVI), described by Gayoso et al., is a pairwise measurement of RNA and surface protein abundance in SC analysis, using CITE-seq that links transcriptional variation with cell phenotypes and functions.

Some studies have formulated the integration problem as a transfer-learning task. For example, Lin et al. pose the integration problem as a transfer learning question, where the model is co-trained on labeled RNA and unlabeled ATAC data. Their model, scJoint, integrates atlas-scale collections of scRNA-seq and scATAC-seq data, using a neural network framework. Their approach uses scATAC-seq to gain a complementary layer of information at the single-cell resolution which is then added to the gene expression data from scRNA-seq. However, scJoint requires that both input matrices share the same dimensions, and scATAC-seq is first converted to gene activity scores, where a single

encoder can share the weights for both RNA and ATAC data [223].

In 2022, Lotfollahi et al. introduced scArches, a DNN-based approach for integrating query datasets onto a reference dataset without sharing raw data. It employs transfer learning and parameter optimization for efficient and decentralized reference building. scArches preserves biological state information and eliminates batch effects across various datasets, including mouse brain, pancreas, immune, whole-organism atlases, and COVID-19 disease variations [88].

In a more recent development in 2023, Lotfollahi et al. introduced ExpiMap, a DL-based algorithm that enhances single-cell atlas integration using (conditional) VAE. ExpiMap maps single-cell references while capturing biologically significant components representing gene programs. It outperforms existing methods and provides interpretability in integrative single-cell analysis. Validation experiments include analyzing single-cell perturbation responses across tissues and species and COVID-19 patient response analysis across different cell types. [89].

5. Conclusions and future perspective

In conclusion, the field of SC sequencing data analysis using DL approaches is rapidly advancing, but one of the most pressing needs is for benchmarking studies. Most papers published using DL algorithms have compared performance to standard methods, but have not gone into great depth when comparing across different types of DL models. SC experiments can be vastly different, with tissue samples containing known cell types, or much more complex tissues such as in diseases such as cancer or COVID-19. However, most methods have only claimed superior performance based on a set of example datasets from specific SC experiments. To overcome these issues, one potential way would be to better understand when these DL models fail or what their limitations are. Understanding the types of DL approaches and model structures that can be beneficial in some cases as compared to others would be very important for developing new approaches and guiding the field in terms of what methods perform better under specific conditions. Additionally, the development of a human cell atlas, i.e., the aggregation of many different human SC expression data across many institutions to cover all major organ systems within the body, would allow for large amounts of annotated SC sequencing data and more comprehensive benchmarking studies [90–93].

DL has played a significant role in the analysis of scRNA-seq data, with an increasing number of studies utilizing these techniques. However, the use of DL for scRNA-seq data analysis is still in its infancy, as the datasets available for analysis are relatively small, which may make it challenging to identify rare cell populations and understand how they change under different disease states. The datasets used in computer vision tasks, where DL has achieved significant improvements, are orders of magnitude larger than those used in scRNA-seq analysis. However, with the increasing availability of scRNA-seq data, the use of these smaller datasets for computational analyses may be changing. A recent work by Sikkema et al., using a combination of 46 different datasets with 2.2 million cells of lung tissue across healthy and diseased patients, has shown that DL methods can outperform standard pipeline approaches in more complex settings. The use of transfer learning, similar to approaches such as scArches, can also be used to save the information gained from large-scale training sets, which can be utilized by researchers who do not have access to such large and diverse datasets. The field of scRNA-seq is also embracing the concept of open-source data sharing, and new toolkits such as scverse aim to provide a unified framework for large-scale scRNA-seq analyses. This information, in combination with other large-scale data collection efforts such as TCGA, can be used to better understand how cellular changes correlate with disease.

The advancements in SC isolation and barcoding techniques have enabled the production of diverse omics data at the SC level. Integrative analyses of multi-omics data at the bulk level have demonstrated the potential to provide a comprehensive understanding of molecular

mechanisms in biology and medicine. DL has proven to be superior to traditional ML methods in the analysis of bulk multi-omics data, due to its ability to capture informative latent features from the high-dimensional heterogeneous multi-omics feature space and its flexible architecture that can model each modality separately and combine them later to aggregate information. Drawing on the success in bulk multi-omics data, integrating scRNA-seq data with other SC omics data as well as multi-omics data at the bulk level using DL may provide a better and deeper understanding of biological mechanisms. However, there are specific and distinct challenges in integrating multi-omics data, particularly in joint-modality SC-seq and the integration of single modality SC-seq data. These challenges include increased noise and sparsity in the data, difficulties in balancing both modalities during the embedding process, and inherent biases between different institutions. Additionally, joint sequencing models are less frequent than single modality sequencing methods, making the development of methods to integrate two different modalities with unique cell populations a significant computational challenge. Despite these challenges, there is still a great potential for DL to advance our understanding of SC omics and contribute to the field of biology and medicine.

ST is a new approach to SC analysis that preserves the spatial relationship of RNA-sequencing within a tissue. While ST has the advantage of spatial resolution, current technology such as the 10x Visium platform generates 50-micron spots that are pooled for analysis, losing the ability to identify the transcriptome of a single cell. Other approaches, such as Slide-Seqv2, aim to improve resolution but have limitations such as limited ability to detect low expressing genes compared with scRNA-seq methods. It is important to realize that scRNA-seq and ST can act in a complimentary manner. ST will require unique computational and DL algorithms separate from scRNA-seq. For example, the method PASTE shows that scRNA-seq related methods are insufficient to properly analyze ST data. Additionally, cell-cell communication networks can be elucidated using newly developed algorithms. However, scRNA-seq can currently provide unique gene information that has been leveraged during ST analysis. For example, DestVI uses a reference scRNA-seq dataset to deconvolve or attempt to identify unique cell types within a given ST spot. Furthermore, work has been done to jointly embed seqFISH data and an scRNA-seq atlas, to annotate specific cell types in the seqFISH dataset. Given current ST spatial resolution constraints and detection limitations, SRT and scRNA-seq can act synergistically. Additionally, the AE structures used in the context of scRNA-seq and in this review can also be components used within ST analysis. By leveraging the strengths of both technologies, we can gain a more comprehensive understanding of cellular biology and the underlying mechanisms of diseases.

We anticipate an increase in the use of DL techniques for addressing the current challenges mentioned in the SC field. Furthermore, we believe that the biological generalizability and interpretability of deep-learned models in understanding complex pathological phenotypes such as cancer, drug resistance, and neurobiology would be of great interest and importance to the omics field.

Declarations

Ethics approval and consent to participate

Not applicable.

Funding Statement

None.

CRedit authorship contribution statement

Nafiseh Erfanian, A. Ali Heydari, Amirhossein Sahebkar, Pablo Iañez: Conceptualization. **Nafiseh Erfanian, Afshin Derakhshani, A.**

Ali Heydari, Mohammad Ghasemigol, Mohsen Farahpour, Saeed Nasseri: Writing – original draft. **Pablo Iañez, Adib Miraki Feriz, Seyyed Mohammad Razavi, Hossein Safarpour, Amirhossein Sahebkar:** Writing – review & editing. All Authors: Approval.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Data availability

No data was used for the research described in the article.

Acknowledgments

Not applicable.

Consent for publication

Not applicable.

References

- [1] L. Haghverdi, A.T. Lun, M.D. Morgan, J.C. Marioni, Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors, *Nat. Biotechnol.* 36 (5) (2018) 421–427.
- [2] F.S. Fritsch, C. Dusny, O. Frick, A. Schmid, Single-cell analysis in biotechnology, systems biology, and biocatalysis, *Annu. Rev. Chem. Biomol. Eng.* 3 (2012) 129–155.
- [3] B. Anchang, T.D. Hart, S.C. Bendall, P. Qiu, Z. Bjornson, M. Linderman, et al., Visualization and cellular hierarchy inference of single-cell data using SPADE, *Nat. Protoc.* 11 (7) (2016) 1264–1279.
- [4] A.L. Haber, M. Biton, N. Rogel, R.H. Herbst, K. Shekhar, C. Smillie, et al., A single-cell survey of the small intestinal epithelium, *Nature* 551 (7680) (2017) 333–339.
- [5] X. Han, Z. Zhou, L. Fei, H. Sun, R. Wang, Y. Chen, et al., Construction of a human cell landscape at single-cell level, *Nature* 581 (7808) (2020) 303–309.
- [6] C. Tabula Muris, C. Overall, C. Logistical, C. Organ, processing, P. Library, et al., Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris, *Nature* 562 (7727) (2018) 367–372.
- [7] A. Roth, J. Khattra, D. Yap, A. Wan, E. Laks, J. Biele, et al., PyClone: statistical inference of clonal population structure in cancer, *Nat. Methods* 11 (4) (2014) 396–398.
- [8] H. Zafar, N. Navin, K. Chen, L. Nakhleh, SiCloneFit: Bayesian inference of population structure, genotype, and phylogeny of tumor clones from single-cell genome sequencing data, *Genome Res.* 29 (11) (2019) 1847–1859.
- [9] A. Regev, S.A. Teichmann, E.S. Lander, I. Amit, C. Benoist, E. Birney, et al., Science forum: the human cell atlas, *Elife* 6 (2017), e27041.
- [10] M. Hosokawa, Y. Nishikawa, M. Kogawa, H. Takeyama, Massively parallel whole genome amplification for single-cell sequencing using droplet microfluidics, *Sci. Rep.* 7 (1) (2017) 1–11.
- [11] R. Zilionis, J. Nainys, A. Veres, V. Savova, D. Zemmour, A.M. Klein, et al., Single-cell barcoding and sequencing using droplet microfluidics, *Nat. Protoc.* 12 (1) (2017) 44–73.
- [12] J. Cao, M. Spielmann, X. Qiu, X. Huang, D.M. Ibrahim, A.J. Hill, et al., The single-cell transcriptional landscape of mammalian organogenesis, *Nature* 566 (7745) (2019) 496–502.
- [13] J. Zheng, K. Wang, Emerging deep learning methods for single-cell RNA-seq data analysis, *Quant. Biol.* 7 (4) (2019) 247–254.
- [14] W.S. McCulloch, Pitts WJTBomb, *A Log. Calc. ideas immanent Nerv. Act.* 5 (1943) 115–133.
- [15] Rosenblatt Fjpr. The perceptron: a probabilistic model for information storage and organization in the brain. 1958;65(6):386.
- [16] Werbos PJPT, Committee on Applied Mathematics, Harvard University, Cambridge, MA. Beyond regression: New tools for prediction and analysis in the behavioral sciences, 1974.
- [17] G.E. Hinton, S. Osindero, Teh Y-WJNC, A fast Learn. Algorithm Deep Belief nets 18 (7) (2006) 1527–1554.
- [18] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, *nature* 323 (6088) (1986) 533–536.
- [19] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (8) (1997) 1735–1780.
- [20] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, et al., Attention is all you need, *Adv. Neural Inf. Process. Syst.* (2017) 30.
- [21] D.H. Hubel, T.N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol.* 160 (1) (1962) 106–154.
- [22] J.C. Kimmel, A.S. Brack, W.F. Marshall, Deep convolutional and recurrent neural networks for cell motility discrimination and prediction, *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 18 (2) (2019) 562–574.
- [23] Pang M., Tegnér J. Multitask learning for Transformers with application to large-scale single-cell transcriptomes. bioRxiv, 2020.
- [24] Y. LeCun, Y. Bengio, Convolutional networks for images, speech, and time series, *Handb. brain Theory Neural Netw.* 3361 (10) (1995) 1995.
- [25] Y. Xu, Z. Zhang, L. You, J. Liu, Z. Fan, X. Zhou, scIGANs: single-cell RNA-seq imputation using generative adversarial networks, *Nucleic Acids Res.* 48 (15) (2020) e85-e.
- [26] G. Eraslan, L.M. Simon, M. Mircea, N.S. Mueller, F.J. Theis, Single-cell RNA-seq denoising using a deep count autoencoder, *Nat. Commun.* 10 (1) (2019) 1–14.
- [27] M.B. Badsha, R. Li, B. Liu, Y.I. Li, M. Xian, N.E. Banovich, et al., Imputation of single-cell gene expression with an autoencoder neural network Running title: Autoencoder for imputation of single-cell gene expression, *Quant. Biol.* 8 (1) (2020) 78–94.
- [28] Lotfollahi M., Litinetskaya A., Theis F.J. Multigrade: single-cell multi-omic data integration. bioRxiv, 2022.
- [29] Kingma D., Welling M., editors. Efficient gradient-based inference through transformations between bayes nets and neural nets. International Conference on Machine Learning; 2014: PMLR.
- [30] Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization. CoRR, 2015; abs/1412.6980.
- [31] Huang H., Li Z., He R., Sun Z., Tan T. Introvae: Introspective variational autoencoders for photographic image synthesis. arXiv preprint 2018:arXiv:1807.06358.
- [32] Heydari A.A., Mehmood A., editors. SRVAE: super resolution using variational autoencoders. Pattern Recognition and Tracking XXXI; 2020: International Society for Optics and Photonics.
- [33] M.A. Moreno Ayala, Z. Li, M. DuPage, Treg programming and therapeutic reprogramming in cancer, *Immunology* 157 (3) (2019) 198–209.
- [34] J. Lucas, G. Tucker, R.B. Grosse, M. Norouzi, Don't blame the Elbo! a linear Vae perspective on posterior collapse, *Adv. Neural Inf. Process. Syst.* 32 (2019) 9408–9418.
- [35] Yang Z., Hu Z., Salakhutdinov R., Berg-Kirkpatrick T., editors. Improved variational autoencoders for text modeling using dilated convolutions. International conference on machine learning; 2017: PMLR.
- [36] Severyn A., Barth E., Semeniuta S. A hybrid convolutional variational autoencoder for text generation. arXiv preprint. 2017:arXiv:1702.02390.
- [37] Tolstikhin I., Bousquet O., Gelly S., Schoelkopf B., editors. Wasserstein autoencoders. Proceedings of the 6th International Conference on Learning Representations; 2018: ICLR.
- [38] Zhao S., Song J., Ermon S. Infovae: Information maximizing variational autoencoders. arXiv preprint 2017:arXiv:1706.02262.
- [39] A.A. Heydari, C.A. Thompson, A. Mehmood, Softadapt: Techniques for adaptive loss weighting of neural networks with multi-part loss functions, arXiv Prepr. 1406 (2019) 2661.
- [40] R. Lopez, J. Regier, M.B. Cole, M.I. Jordan, N. Yosef, Deep generative modeling for single-cell transcriptomics, *Nat. Methods* 15 (12) (2018) 1053–1058.
- [41] V. Svensson, A. Gayoso, N. Yosef, L. Pachter, Interpretable factor models of single-cell RNA-seq via variational autoencoders, *Bioinformatics* 36 (11) (2020) 3418–3421.
- [42] A.A. Heydari, O.A. Davalos, L. Zhao, K.K. Hoyer, S.S. Sindi, ACTIVA: realistic single-cell RNA-seq generation with automatic cell-type identification using introspective variational autoencoders, *Bioinformatics* 38 (8) (2022) 2194–2201.
- [43] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., Generative adversarial networks, arXiv Prepr. 1406 (2014) 2661.
- [44] G.K. Dziugaite, D.M. Roy, Z. Ghahramani, Training generative neural networks via maximum mean discrepancy optimization, arXiv Prepr. 1505 (2015) 03906.
- [45] Vondrick C., Pirsaviash H., Torralba A. Generating videos with scene dynamics. arXiv preprint, 2016: arXiv:1609.02612.
- [46] Zhu J.-Y., Krähenbühl P., Shechtman E., Efros A.A., editors. Generative visual manipulation on the natural image manifold. European conference on computer vision; 2016: Springer.
- [47] Fedus W., Goodfellow I., Dai A.M. Maskgan: better text generation via filling in the. arXiv preprint 2018:arXiv:1801.07736.
- [48] Yang X., Chen Y.-N., Hakkani-Tür D., Crook P., Li X., Gao J., et al., editors. End-to-end joint learning of natural language understanding and dialogue manager. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2017: IEEE.
- [49] Engel J., Agrawal K.K., Chen S., Gulrajani I., Donahue C., Roberts A. Gansynth: Adversarial neural audio synthesis. arXiv preprint. 2019:arXiv:1902.08710.
- [50] Esteban C., Hyland S.L., Rätsch G. Real-valued (medical) time series generation with recurrent conditional gans. arXiv preprint 2017:arXiv:1706.02633.
- [51] M. Marouf, P. Machart, V. Bansal, C. Kilian, D.S. Magruder, C.F. Krebs, et al., Realistic in silico generation and augmentation of single-cell RNA-seq data using generative adversarial networks, *Nat. Commun.* 11 (1) (2020) 1–12.
- [52] Z. Wang, Q. She, T.E. Ward, Generative adversarial networks in computer vision: A survey and taxonomy, *ACM Comput. Surv. (CSUR)* 54 (2) (2021) 1–38.
- [53] Arjovsky M., Chintala S., Bottou L., editors. Wasserstein generative adversarial networks. International conference on machine learning; 2017: PMLR.
- [54] Larsen A.B.L., Sønderby S.K., Larochelle H., Winther O., editors. Autoencoding beyond pixels using a learned similarity metric. International conference on machine learning; 2016: PMLR.
- [55] Metz L., Poole B., Pfau D., Sohl-Dickstein J. Unrolled generative adversarial networks. arXiv preprint, 2016:arXiv:1611.02163.
- [56] J. Park, H. Kim, J. Kim, M. Cheon, A practical application of generative adversarial networks for RNA-seq analysis to predict the molecular progress of Alzheimer's disease, *PLoS Comput. Biol.* 16 (7) (2020), e1008099.

- [57] S. Aldridge, S.A. Teichmann, Single cell transcriptomics comes of age, *Nat. Commun.* 11 (1) (2020) 1–4.
- [58] S. Linnarsson, S.A. Teichmann, Single-cell genomics: coming of age, *Genome Biol.* 17 (1) (2016) 1–3.
- [59] O. Schwartzman, A. Tanay, Single-cell epigenomics: techniques and emerging applications, *Nat. Rev. Genet.* 16 (12) (2015) 716–726.
- [60] D.T. Montoro, A.L. Haber, M. Biton, V. Vinarsky, B. Lin, S.E. Birket, et al., A revised airway epithelial hierarchy includes CFTR-expressing ionocytes, *Nature* 560 (7718) (2018) 319–324.
- [61] M. Plass, J. Solana, F.A. Wolf, S. Ayoub, A. Misios, P. Glazár, et al., Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics, *Science* 360 (2018) 6391.
- [62] G. Eraslan, L.M. Simon, M. Mircea, N.S. Mueller, F.J. Theis, Single-cell RNA-seq denoising using a deep count autoencoder, *Nat. Commun.* 10 (1) (2019) 390.
- [63] M. Amodio, D. van Dijk, K. Srinivasan, W.S. Chen, H. Mohsen, K.R. Moon, et al., Exploring single-cell data with deep multitasking neural networks, *Nat. Methods* 16 (11) (2019) 1139–1145.
- [64] D. Talwar, A. Mongia, D. Sengupta, A. Majumdar, AutoImpute: Autoencoder based imputation of single-cell RNA-seq data, *Sci. Rep.* 8 (1) (2018) 16329.
- [65] C. Arisdakessian, O. Poirion, B. Yunits, X. Zhu, L.X. Garmire, DeepImpute: an accurate, fast, and scalable deep neural network method to impute single-cell RNA-seq data, *Genome Biol.* 20 (1) (2019) 1–14.
- [66] A. Mongia, D. Sengupta, A. Majumdar, deepMc: deep matrix completion for imputation of single-cell RNA-seq data, *J. Comput. Biol.* 27 (7) (2020) 1011–1019.
- [67] S. Sun, J. Zhu, Y. Ma, X. Zhou, Accuracy, robustness and scalability of dimensionality reduction methods for single-cell RNA-seq analysis, *Genome Biol.* 20 (1) (2019) 269.
- [68] R. Yan, C. Fan, Z. Yin, T. Wang, X. Chen, Potential applications of deep learning in single-cell RNA sequencing analysis for cell therapy and regenerative medicine, *Stem Cells* 39 (5) (2021) 511–521.
- [69] H.T.N. Tran, K.S. Ang, M. Chevrier, X. Zhang, N.Y.S. Lee, M. Goh, et al., A benchmark of batch-effect correction methods for single-cell RNA sequencing data, *Genome Biol.* 21 (1) (2020) 12.
- [70] M.D. Luecken, M. Büttner, K. Chaichoompu, A. Danese, M. Interlandi, M. F. Mueller, et al., Benchmarking atlas-level data integration in single-cell genomics, *Nat. Methods* 19 (1) (2022) 41–50.
- [71] X. Li, K. Wang, Y. Lyu, H. Pan, J. Zhang, D. Stambolian, et al., Deep learning enables accurate clustering with batch effect removal in single-cell RNA-seq analysis, *Nature, Communications* 11 (1) (2020) 2338.
- [72] W. Hou, Z. Ji, H. Ji, S.C. Hicks, A systematic evaluation of single-cell RNA-seq imputation methods, *Genome Biol.* 21 (1) (2020) 218.
- [73] Molho D., Ding J., Li Z., Wen H., Tang W., Wang Y., et al. Deep Learning in Single-Cell Analysis. arXiv preprint arXiv:221012385, 2022.
- [74] J. Xu, J. Xu, Y. Meng, C. Lu, L. Cai, X. Zeng, et al., Graph embedding and Gaussian mixture variational autoencoder network for end-to-end analysis of single-cell RNA sequencing data, *Cell Reports, Methods* (2023), 100382.
- [75] E. Becht, L. McInnes, J. Healy, C.-A. Dutertre, I.W. Kwok, L.G. Ng, et al., Dimensionality reduction for visualizing single-cell data using UMAP, *Nat. Biotechnol.* 37 (1) (2019) 38–44.
- [76] Mårtens K., Yau C., editors. BasisVAE: Translation-invariant feature-level clustering with Variational Autoencoders. International Conference on Artificial Intelligence and Statistics; 2020: PMLR.
- [77] J. Peng, X. Wang, X. Shang, Combining gene ontology with deep neural networks to enhance the clustering of single cell RNA-Seq data, *BMC Bioinforma.* 20 (8) (2019) 284.
- [78] L. Chen, Y. Zhai, Q. He, W. Wang, M. Deng, Integrating Deep Supervised, Self-Supervised and Unsupervised Learning for Single-Cell RNA-seq Clustering and Annotation, *Genes* 11 (7) (2020) 792.
- [79] Y. Zeng, Z. Wei, F. Zhong, Z. Pan, Y. Lu, Y. Yang, A parameter-free deep embedded clustering method for single-cell RNA-seq data, *Brief. Bioinforma.* 23 (2022) 5.
- [80] S. He, T. Yu, J. Fan, G3DC: a Gene-Graph-Guided selective Deep Clustering method for single cell RNA-seq data. bioRxiv, 2023, 2023.01.15.524109.
- [81] Z. Liu, D. Sun, C. Wang, Evaluation of cell-cell interaction methods by integrating single-cell RNA sequencing data with spatial information, *Genome Biol.* 23 (1) (2022) 218.
- [82] Z. Chen, W.C. King, A. Hwang, M. Gerstein, J. Zhang, DeepVelo: Single-cell transcriptomic deep velocity field learning with neural ordinary differential equations, *Sci. Adv.* 8 (48) (2022) eabq3745.
- [83] C. Zuo, L. Chen, Deep-joint-learning analysis model of single cell transcriptome and open chromatin accessibility data, *Brief. Bioinforma.* 22 (4) (2021) bbaa287.
- [84] C. Zuo, H. Dai, L. Chen, Deep cross-omics cycle attention model for joint analysis of single-cell multi-omics data, *Bioinformatics* 37 (22) (2021) 4091–4099.
- [85] A. Gayoso, Z. Steier, R. Lopez, J. Regier, K.L. Nazor, A. Streets, et al., Joint probabilistic modeling of single-cell multi-omic data with totalVI, *Nat. Methods* 18 (3) (2021) 272–282.
- [86] K.E. Wu, K.E. Yost, H.Y. Chang, J. Zou, BABEL enables cross-modality translation between multiomic profiles at single-cell resolution, *Proc. Natl. Acad. Sci.* 118 (15) (2021), e2023070118.
- [87] Ma A., Wang X., Wang C., Li J., Xiao T., Wang J., et al. Deepmaps: Single-cell biological network inference using heterogeneous graph transformer. bioRxiv. 2021.
- [88] M. Lotfollahi, M. Naghipourfar, M.D. Luecken, M. Khajavi, M. Büttner, M. Wagenstetter, et al., Mapping single-cell data to reference atlases by transfer learning, *Nat. Biotechnol.* 40 (1) (2022) 121–130.
- [89] M. Lotfollahi, S. Rybakov, K. Hrovatin, S. Hediyyeh-Zadeh, C. Talavera-López, A. V. Misharin, et al., Biologically informed deep learning to query gene programs in single-cell atlases, *Nat. Cell Biol.* 25 (2) (2023) 337–350.
- [90] S. Jin, C.F. Guerrero-Juarez, L. Zhang, I. Chang, R. Ramos, C.-H. Kuan, et al., Inference and analysis of cell-cell communication using CellChat, *Nat. Commun.* 12 (1) (2021) 1088.
- [91] M. Efremova, M. Vento-Tormo, S.A. Teichmann, R. Vento-Tormo, CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes, *Nat. Protoc.* 15 (4) (2020) 1484–1506.
- [92] M. Solovey, A. Scialdone, COMUNET: a tool to explore and visualize intercellular communication, *Bioinformatics* 36 (15) (2020) 4296–4300.
- [93] Y. Wang, R. Wang, S. Zhang, S. Song, C. Jiang, G. Han, et al., iTALK: an R Package to Characterize and Illustrate Intercellular Communication, bioRxiv (2019), 507871.
- [94] G.X. Zheng, J.M. Terry, P. Belgrader, P. Ryvkin, Z.W. Bent, R. Wilson, et al., Massively parallel digital transcriptional profiling of single cells, *Nat. Commun.* 8 (1) (2017) 1–12.
- [95] A.M. Klein, L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, et al., Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells, *Cell* 161 (5) (2015) 1187–1201.
- [96] E. Azizi, A.J. Carr, G. Plitas, A.E. Cornish, C. Konopacki, S. Prabhakaran, et al., Single-cell map of diverse immune phenotypes in the breast tumor microenvironment, *Cell* 174 (5) (2018) 1293–1308, e36.
- [97] S. Parekh, C. Ziegenhain, B. Vieth, W. Enard, I. Hellmann, zUMIs-a fast and flexible pipeline to process RNA sequencing data with UMIs, *Gigascience* (2018).
- [98] A. Lafzi, C. Moutinho, S. Picelli, H. Heyn, Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies, *Nat. Protoc.* 13 (12) (2018) 2742–2757.
- [99] P. Jiang, J.A. Thomson, R. Stewart, Quality control of single-cell RNA-seq by SinQC, *Bioinformatics* 32 (16) (2016) 2514–2516.
- [100] J.A. Griffiths, A. Scialdone, J.C. Marioni, Using single-cell genomics to understand developmental processes and cell fate decisions, *Mol. Syst. Biol.* 14 (4) (2018), e8046.
- [101] R. Bacher, L.-F. Chu, N. Leng, A.P. Gasch, J.A. Thomson, R.M. Stewart, et al., SCnorm: robust normalization of single-cell RNA-seq data, *Nat. Methods* 14 (6) (2017) 584.
- [102] S.A. Hogan, A. Courtier, P.F. Cheng, N.F. Jaberg-Bentele, S.M. Goldinger, M. Manuel, et al., Peripheral blood TCR repertoire profiling may facilitate patient stratification for immunotherapy against melanoma, *Cancer Immunol. Res.* 7 (1) (2019) 77–85.
- [103] A.T. Lun, K. Bach, J.C. Marioni, Pooling across cells to normalize single-cell RNA sequencing data with many zero counts, *Genome Biol.* 17 (1) (2016) 1–14.
- [104] F. Büttner, K.N. Natarajan, F.P. Casale, V. Proserpio, A. Scialdone, F.J. Theis, et al., Computational analysis of cell-to-cell heterogeneity in single-cell RNA-seq data reveals hidden subpopulations of cells, *Nat. Biotechnol.* 33 (2) (2015) 155–160.
- [105] B. Ding, L. Zheng, Y. Zhu, N. Li, H. Jia, R. Ai, et al., Normalization and noise reduction for single cell RNA-seq experiments, *Bioinformatics* 31 (13) (2015) 2225–2227.
- [106] C.A. Vallejos, J.C. Marioni, S. Richardson, BASiCS: Bayesian analysis of single-cell sequencing data, *PLoS Comput. Biol.* 11 (2015) 6.
- [107] M.D. Luecken, F.J. Theis, Current best practices in single-cell RNA-seq analysis: a tutorial, *Mol. Syst. Biol.* 15 (6) (2019), e8746.
- [108] W. Gong, I.-Y. Kwak, P. Pota, N. Koyano-Nakagawa, D.J. Garry, DRImpute: imputing dropout events in single cell RNA sequencing data, *BMC Bioinforma.* 19 (1) (2018) 220.
- [109] P.V. Kharchenko, L. Silberstein, D.T. Scadden, Bayesian approach to single-cell differential expression analysis, *Nat. Methods* 11 (7) (2014) 740–742.
- [110] P. Qiu, Embracing the dropouts in single-cell RNA-seq analysis, *Nat. Commun.* 11 (2020) 1–9.
- [111] D. Ran, S. Zhang, N. Lytal, L. An, scDoc: correcting drop-out events in single-cell RNA-seq data, *Bioinformatics* 36 (15) (2020) 4233–4239.
- [112] L. Zappia, B. Phipson, A. Oshlack, Splatter: simulation of single-cell RNA sequencing data, *Genome Biol.* 18 (1) (2017) 174.
- [113] P. Lin, M. Troup, J.W. Ho, CIDR: Ultrafast and accurate clustering through imputation for single-cell RNA-seq data, *Genome Biol.* 18 (1) (2017) 59.
- [114] D. Sengupta, N.A. Rayan, M. Lim, B. Lim, Prabhakar S. Fast, scalable and accurate differential expression analysis for single cells, *BioRxiv* (2016), 049734.
- [115] D. van Dijk, J. Nainys, R. Sharma, P. Kaithail, A.J. Carr, K.R. Moon, et al., MAGIC: A diffusion-based imputation method reveals gene-gene interactions in single-cell RNA-sequencing data, *BioRxiv* (2017), 111591.
- [116] M. Huang, J. Wang, E. Torre, H. Dueck, S. Shaffer, R. Bonasio, et al., SAVER: gene expression recovery for single-cell RNA sequencing, *Nat. Methods* 15 (7) (2018) 539–542.
- [117] W.V. Li, J.J. Li, An accurate and robust imputation method scImpute for single-cell RNA-seq data, *Nat. Commun.* 9 (1) (2018) 1–9.
- [118] M. Moussa, Mandoiu I. Locality sensitive imputation for single cell RNA-seq data, *J. Comput. Biol.* 26 (8) (2019) 822–835.
- [119] D. Talwar, A. Mongia, D. Sengupta, A. Majumdar, AutoImpute: Autoencoder based imputation of single-cell RNA-seq data, *Sci. Rep.* 8 (1) (2018) 1–11.
- [120] M. Amodio, D. Van Dijk, K. Srinivasan, W.S. Chen, H. Mohsen, K.R. Moon, et al., Exploring single-cell data with deep multitasking neural networks, *Nat. Methods* 16 (11) (2019) 1139–1145.
- [121] Y. Deng, F. Bao, Q. Dai, L.F. Wu, S.J. Altschuler, Scalable analysis of cell-type composition from single-cell transcriptomics using deep recurrent learning, *Nat. Methods* 16 (4) (2019) 311–314.

- [122] L. Patrino, D. Maspero, F. Craighero, F. Angaroni, M. Antoniotti, A. Graudenzi, A review of computational strategies for denoising and imputation of single-cell transcriptomic data, *Brief. Bioinforma.* 22 (2020) 4.
- [123] T. Fei, T. Yu, scBatch: batch-effect correction of RNA-seq data through sample distance matrix adjustment, *Bioinformatics* (2020).
- [124] S. Lin, Y. Lin, J.R. Nery, M.A. Urich, A. Breschi, C.A. Davis, et al., Comparison of the transcriptional landscapes between human and mouse tissues, *Proc. Natl. Acad. Sci.* 111 (48) (2014) 17224–17229.
- [125] W.E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods, *Biostatistics* 8 (1) (2007) 118–127.
- [126] X. Luo, Y. Wei, Batch effects correction with unknown subtypes, *J. Am. Stat. Assoc.* 114 (526) (2019) 581–594.
- [127] M. Büttner, Z. Miao, F.A. Wolf, S.A. Teichmann, F.J. Theis, A test metric for assessing single-cell RNA-seq batch correction, *Nat. Methods* 16 (1) (2019) 43–49.
- [128] F. Zhang, Y. Wu, W. Tian, A novel approach to remove the batch effect of single-cell data, *Cell Discov.* 5 (1) (2019) 1–4.
- [129] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, et al., limma powers differential expression analyses for RNA-sequencing and microarray studies, *Nucleic Acids Res.* 43 (7) (2015) e47–e.
- [130] G. Finak, A. McDavid, M. Yajima, J. Deng, V. Gersuk, A.K. Shalek, et al., MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data, *Genome Biol.* 16 (1) (2015) 1–13.
- [131] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15 (12) (2014) 1–21.
- [132] B. Hie, B. Bryson, B. Berger, Efficient integration of heterogeneous single-cell transcriptomes using Scanorama, *Nat. Biotechnol.* 37 (6) (2019) 685–691.
- [133] K. Polański, M.D. Young, Z. Miao, K.B. Meyer, S.A. Teichmann, J.-E. Park, BBKNN: fast batch alignment of single cell transcriptomes, *Bioinformatics* 36 (3) (2020) 964–965.
- [134] He K., Zhang X., Ren S., Sun J., editors. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
- [135] U. Shaham, K.P. Stanton, J. Zhao, H. Li, K. Raddassi, R. Montgomery, et al., Removal of batch effects using distribution-matching residual networks, *Bioinformatics* 33 (16) (2017) 2539–2546.
- [136] X. Li, K. Wang, Y. Lyu, H. Pan, J. Zhang, D. Stambolian, et al., Deep learning enables accurate clustering with batch effect removal in single-cell RNA-seq analysis, *Nat. Commun.* 11 (1) (2020) 1–14.
- [137] U. Shaham, Batch Effect Removal via Batch-Free Encoding, *bioRxiv* (2018), 380816.
- [138] T. Wang, T.S. Johnson, W. Shao, Z. Lu, B.R. Helm, J. Zhang, et al., BERMUDA: a novel deep transfer learning method for single-cell RNA sequencing batch correction reveals hidden high-resolution cellular subtypes, *Genome Biol.* 20 (1) (2019) 1–15.
- [139] D. Gan, J. Li, SCIBER: a simple method for removing batch effects from single-cell RNA-sequencing data, *Bioinformatics* 39 (1) (2023) btac819.
- [140] D. Wang, J. Gu, VASC: dimension reduction and visualization of single-cell RNA-seq data by deep variational autoencoder, *Genom., Proteom. Bioinforma.* 16 (5) (2018) 320–331.
- [141] K. Pearson, L.I.I. On lines and planes of closest fit to systems of points in space, *Lond., Edinb., Dublin Philos. Mag. J. Sci.* 2 (11) (1901) 559–572.
- [142] L. Van der Maaten, G. Hinton, Visualizing data using t-SNE, *J. Mach. Learn. Res.* 9 (2008) 11.
- [143] L. Haghverdi, F. Büttner, F.J. Theis, Diffusion maps for high-dimensional single-cell analysis of differentiation data, *Bioinformatics* 31 (18) (2015) 2989–2998.
- [144] Titsias M., Lawrence N.D., editors. Bayesian Gaussian process latent variable model. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics; 2010: JMLR Workshop and Conference Proceedings.
- [145] F. Büttner, F.J. Theis, A novel approach for resolving differences in single-cell gene expression patterns from zygote to blastocyst, *Bioinformatics* 28 (18) (2012) i626–i632.
- [146] J. Ding, A. Condon, S.P. Shah, Interpretable dimensionality reduction of single cell transcriptome data with deep generative models, *Nat. Commun.* 9 (1) (2018) 1–13.
- [147] E. Pierson, C. Yau, ZIFA: Dimensionality reduction for zero-inflated single-cell gene expression analysis, *Genome Biol.* 16 (1) (2015) 1–10.
- [148] ARMACKI A. Application of Autoencoders on Single-cell Data: University OF Novi Sad; 2018.
- [149] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, *J. Big Data* 6 (1) (2019) 1–48.
- [150] Z. Zhang, F. Cui, C. Lin, L. Zhao, C. Wang, Q. Zou, Critical downstream analysis steps for single-cell RNA sequencing data, *Brief. Bioinforma.* (2021).
- [151] L. Yang, J. Liu, Q. Lu, A.D. Riggs, X. Wu, SAIC: an iterative clustering approach for analysis of single cell RNA-seq data, *BMC Genom.* 18 (6) (2017) 689.
- [152] X. Han, R. Wang, Y. Zhou, L. Fei, H. Sun, S. Lai, et al., Mapping the mouse cell atlas by microwell-seq, *Cell* 172 (5) (2018) 1091–1107, e17.
- [153] K. Davie, J. Janssens, D. Koldere, M. De Waegeneer, U. Pech, L. Kreft, et al., A single-cell transcriptome atlas of the aging *Drosophila* brain, *Cell* 174 (4) (2018) 982–998, e20.
- [154] O. Rozenblatt-Rosen, M.J. Stubbington, A. Regev, S.A. Teichmann, The Human Cell Atlas: from vision to reality, *Nat. N.* 550 (7677) (2017) 451.
- [155] R. Zheng, M. Li, Z. Liang, F.-X. Wu, Y. Pan, J. Wang, SinNLR: a robust subspace clustering method for cell type detection by non-negative and low-rank representation, *Bioinformatics* 35 (19) (2019) 3642–3650.
- [156] Guo X., Zhu E., Liu X., Yin J., editors. Deep embedded clustering with data augmentation. Asian conference on machine learning; 2018.
- [157] H.T.N. Tran, K.S. Ang, M. Chevrier, X. Zhang, N.Y.S. Lee, M. Goh, et al., A benchmark of batch-effect correction methods for single-cell RNA sequencing data, *Genome Biol.* 21 (1) (2020) 1–32.
- [158] T. Tian, J. Wan, Q. Song, Z. Wei, Clustering single-cell RNA-seq data with a model-based deep learning approach, *Nat. Mach. Intell.* 1 (4) (2019) 191–198.
- [159] C.H. Grønbech, M.F. Vording, P.N. Timshel, C.K. Sønderby, T.H. Pers, O. Winther, scVAE: Variational auto-encoders for single-cell gene expression data, *Bioinformatics* 36 (16) (2020) 4415–4422.
- [160] J. Wang, J. Xia, H. Wang, Y. Su, C.-H. Zheng, scDCCA: deep contrastive clustering for single-cell RNA-seq data based on auto-encoder network, *Brief. Bioinforma.* 24 (1) (2023) bbac625.
- [161] D. Hu, K. Liang, S. Zhou, W. Tu, M. Liu, X. Liu, scDFC: A deep fusion clustering method for single-cell RNA-seq data, *Brief. Bioinforma.* (2023) bbad216.
- [162] R. Qi, A. Ma, Q. Ma, Q. Zou, Clustering and classification methods for single-cell RNA-sequencing data, *Brief. Bioinform* 21 (4) (2020) 1196–1208.
- [163] V.Y. Kiselev, T.S. Andrews, M. Hemberg, Challenges in unsupervised clustering of single-cell RNA-seq data, *Nat. Rev. Genet.* 20 (5) (2019) 273–282.
- [164] Jiang Q., Yang W., Xu Z., Luo M., Cai Y., Xu C., et al. DeepCCI: a deep learning framework for identifying cell-cell interactions from single-cell RNA sequencing data. *bioRxiv*. 2022:2022.11.11516061.
- [165] Y. Yuan, Z. Bar-Joseph, GCNG: graph convolutional networks for inferring gene interaction from spatial transcriptomics data, *Genome Biol.* 21 (1) (2020) 1–16.
- [166] L. Atta, A. Sahoo, J. Fan, VeloViz: RNA velocity-informed embeddings for visualizing cellular trajectories, *Bioinformatics* 38 (2) (2022) 391–396.
- [167] V. Bergen, M. Lange, S. Peidli, F.A. Wolf, F.J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling, *Nat. Biotechnol.* 38 (12) (2020) 1408–1414.
- [168] Cui H., Maan H., Wang B. DeepVelo: Deep Learning extends RNA velocity to multi-lineage systems with cell-specific kinetics. *bioRxiv*. 2022.
- [169] V. Bergen, R.A. Soldatov, P.V. Kharchenko, F.J. Theis, RNA velocity—current challenges and future perspectives, *Mol. Syst. Biol.* 17 (8) (2021) e10282.
- [170] L. Xiong, K. Xu, K. Tian, Y. Shao, L. Tang, G. Gao, et al., SCALE method for single-cell ATAC-seq analysis via latent feature extraction, *Nat. Commun.* 10 (1) (2019) 1–10.
- [171] C. Angermueller, H.J. Lee, W. Reik, O. Stegle, DeepCpG: accurate prediction of single-cell DNA methylation states using deep learning, *Genome Biol.* 18 (1) (2017) 1–13.
- [172] S.M. Linker, L. Urban, S.J. Clark, M. Chhatrivala, S. Amatya, D.J. McCarthy, et al., Combined single-cell profiling of expression and DNA methylation reveals splicing regulation and heterogeneity, *Genome Biol.* 20 (1) (2019) 1–14.
- [173] F. Yan, D.R. Powell, D.J. Curtis, N.C. Wong, From reads to insight: a hitchhiker's guide to ATAC-seq data analysis, *Genome Biol.* 21 (1) (2020) 22.
- [174] Cao Y., Fu L., Wu J., Peng Q., Nie Q., Zhang J., et al. SAILER: Scalable and Accurate Invariant Representation Learning for Single-Cell ATAC-Seq Processing and Integration. *bioRxiv*, 2021.
- [175] S. Baek, I. Lee, Single-cell ATAC sequencing analysis: From data preprocessing to hypothesis generation, *Comput. Struct. Biotechnol. J.* (2020).
- [176] Hentges L.D., Sergeant M.D., Downes D.J., Hughes J.R., Taylor S. LanceOtron: a deep learning peak caller for ATAC-seq, ChIP-seq, and DNase-seq. *bioRxiv*. 2021.
- [177] A. Thibodeau, S. Khetan, A. Eroglu, R. Tewhey, M.L. Stitzel, D. Ucar, CoRE-ATAC: A deep learning model for the functional classification of regulatory elements from single cell and bulk ATAC-seq data, *PLoS Comput. Biol.* 17 (12) (2021), e1009670.
- [178] Y. Cao, L. Fu, J. Wu, Q. Peng, Q. Nie, J. Zhang, et al., SAILER: scalable and accurate invariant representation learning for single-cell ATAC-seq processing and integration, *Bioinformatics* 37 (Supplement 1) (2021) i317–i326.
- [179] Y. Yuan, Z. Bar-Joseph, Deep learning for inferring gene relationships from single-cell expression data, *Proc. Natl. Acad. Sci.* 116 (52) (2019) 27151–27158.
- [180] L. Fu, L. Zhang, E. Dollinger, Q. Peng, Q. Nie, X. Xie, Predicting transcription factor binding in single cells through deep learning, *Sci. Adv.* 6 (51) (2020) eab9031.
- [181] T.S. Furey, ChIP-seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions, *Nat. Rev. Genet.* 13 (12) (2012) 840–852.
- [182] Albrecht S., Andreani T., Andrade-Navarro M.A., Fontaine J.-F. Interpretable machine learning models for single-cell ChIP-seq imputation. *bioRxiv*, 2021: 2019.12.20.883983.
- [183] Eghbal-zadeh H., Fischer L., Popitsch N., Kromp F., Taschner-Mandl S., Koutini K., et al. Deep snp: An end-to-end deep neural network with attention-based localization for break-point detection in snp array genomic data. *arXiv preprint arXiv:180608840*, 2018.
- [184] C. Gawad, W. Koh, S.R. Quake, Single-cell genome sequencing: current state of the science, *Nat. Rev. Genet.* 17 (3) (2016) 175–188.
- [185] T. Hill, R.L. Unckless, A Simple Deep Learning Approach for Detecting Duplications and Deletions in Next-Generation Sequencing Data, *bioRxiv* (2019), 657361.
- [186] J.T. Gleissner, X. Hou, C. Zhong, J. Zhang, M. Khan, F. Brand, et al., DeepCNV: a deep learning approach for authenticating copy number variations, *Brief. Bioinforma.* 22 (5) (2021) bbaa381.
- [187] J. Xia, L. Wang, G. Zhang, C. Zuo, L. Chen, RDAClone: Deciphering Tumor Heterozygosity through Single-Cell Genomics Data Analysis with Robust Deep Autoencoder, *Genes* 12 (12) (2021) 1847.

- [188] Yu Z., Liu F., Shi F., Du F. rcCAE: a convolutional autoencoder based method for detecting tumor clones and copy number alterations from single-cell DNA sequencing data. *bioRxiv*. 2022:2022.12.04.519013.
- [189] V. Marx, Method of the Year: spatially resolved transcriptomics, *Nat. Methods* 18 (1) (2021) 9–14.
- [190] P.L. Ståhl, F. Salmén, S. Vickovic, A. Lundmark, J.F. Navarro, J. Magnusson, et al., Visualization and analysis of gene expression in tissue sections by spatial transcriptomics, *Science* 353 (6294) (2016) 78–82.
- [191] S.G. Rodrigues, R.R. Stickels, A. Goeva, C.A. Martin, E. Murray, C.R. Vanderburg, et al., Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution, *Science* 363 (6434) (2019) 1463–1467.
- [192] S. Vickovic, G. Eraslan, F. Salmén, J. Klughammer, L. Stenbeck, D. Schapiro, et al., High-definition spatial transcriptomics for in situ tissue profiling, *Nat. Methods* 16 (10) (2019) 987–990.
- [193] D. Lähnemann, J. Köster, E. Szczurek, D.J. McCarthy, S.C. Hicks, M.D. Robinson, et al., Eleven grand challenges in single-cell data science, *Genome Biol.* 21 (1) (2020) 1–35.
- [194] Pham D.T., Tan X., Xu J., Grice L.F., Lam P.Y., Raghuram A., et al. stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues. *bioRxiv*, 2020.
- [195] C. Xu, X. Jin, S. Wei, P. Wang, M. Luo, Z. Xu, et al., DeepST: identifying spatial domains in spatial transcriptomics by deep learning, *Nucleic Acids Res.* 50 (22) (2022) e131-e.
- [196] F. Maseda, Z. Cang, Q. Nie, DEEPsc: a deep learning-based map connecting single-cell transcriptomics and spatial imaging data, *Front. Genet.* 12 (2021), 636743.
- [197] T. Biancalani, G. Scalia, L. Buffoni, R. Avasthi, Z. Lu, A. Sanger, et al., Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram, *Nat. Methods* 18 (11) (2021) 1352–1362.
- [198] X. Tan, A. Su, M. Tran, Q. Nguyen, SpaCell: integrating tissue morphology and spatial gene expression to predict disease cells, *Bioinformatics* 36 (7) (2019) 2293–2294.
- [199] H. Fu, H. Xu, K. Chong, M. Li, K.S. Ang, H.K. Lee, et al., Unsupervised spatially embedded deep representation of spatial transcriptomics, *Biorxiv* (2021).
- [200] O.M. O'Connor, R.N. Alnahhas, J.-B. Lugagne, M.J. Dunlop, DeLTA 2.0: A deep learning pipeline for quantifying single-cell spatial and temporal dynamics, *PLOS Comput. Biol.* 18 (1) (2022), e1009797.
- [201] M. Elosua-Bayes, P. Nieto, E. Mereu, I. Gut, H. Heyn, SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes, *Nucleic Acids Res.* 49 (9) (2021) e50-e.
- [202] T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W.M. Mauck III, et al., Comprehensive integration of single-cell data, *Cell* 177 (7) (2019) 1888–1902, e21.
- [203] A. Andersson, J. Bergenstråhle, M. Asp, L. Bergenstråhle, A. Jurek, J.F. Navarro, et al., Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography, *Commun. Biol.* 3 (1) (2020) 565.
- [204] Kleshchevnikov V., Shmatko A., Dann E., Aivazidis A., King H.W., Li T., et al. Comprehensive mapping of tissue cell architecture via integrated single cell and spatial transcriptomics. *bioRxiv*, 2020:2020.11.15.378125.
- [205] R. Lopez, B. Li, H. Keren-Shaul, P. Boyeau, M. Kedmi, D. Pilzer, et al., DestVI identifies continuums of cell types in spatial transcriptomics data, *Nat. Biotechnol.* (2022) 1–10.
- [206] S. Teichmann, M. Efrimova, Method of the Year 2019: single-cell multimodal omics, *Nat. Methods* 17 (1) (2020).
- [207] N. Wani, K. Raza, Integrative approaches to reconstruct regulatory networks from multi-omics data: a review of state-of-the-art methods, *Comput. Biol. Chem.* 83 (2019), 107120.
- [208] D. Grapov, J. Fahrman, K. Wanichthanarak, S. Khoomrung, Rise of deep learning for genomic, proteomic, and metabolomic data integration in precision medicine, *Omics: a J. Integr. Biol.* 22 (10) (2018) 630–636.
- [209] Y. Hao, S. Hao, E. Andersen-Nissen, W.M. Mauck, S. Zheng, A. Butler, et al., Integrated analysis of multimodal single-cell data, *Cell* 184 (13) (2021) 3573–3587, e29.
- [210] S.J. Clark, R. Argelaguet, C.-A. Kapourani, T.M. Stubbs, H.J. Lee, C. Alda-Catalinas, et al., scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells, *Nat. Commun.* 9 (1) (2018) 1–9.
- [211] C. Zuo, L. Chen, Deep-joint-learning analysis model of single cell transcriptome and open chromatin accessibility data, *Brief. Bioinforma.* 22 (2020) 4.
- [212] Heydari A.A., Sindi S.S. Deep Learning in Spatial Transcriptomics: Learning From the Next Next-Generation Sequencing. *bioRxiv*. 2022.
- [213] C. Zhu, S. Preissl, B. Ren, Single-cell multimodal omics: the power of many, *Nat. Methods* 17 (1) (2020) 11–14.
- [214] Amodio M., Krishnaswamy S., editors. MAGAN: Aligning biological manifolds. International Conference on Machine Learning; 2018: PMLR.
- [215] Liu J., Huang Y., Singh R., Vert J.-P., Noble W.S., editors. Jointly embedding multiple single-cell omics measurements. Algorithms in bioinformatics: International Workshop, WABI, proceedings WABI (Workshop); 2019: NIH Public Access.
- [216] K. Cao, X. Bai, Y. Hong, L. Wan, Unsupervised topological alignment for single-cell multi-omics integration, *Bioinformatics* 36 (Supplement_1) (2020) i48–i56.
- [217] Y. Xu, P. Das, R.P. McCord, SMILE: mutual information learning for integration of single-cell omics data, *Bioinformatics* 38 (2) (2022) 476–486.
- [218] S.G. Stark, J. Ficek, F. Locatello, X. Bonilla, S. Chevrier, F. Singer, et al., SCIM: universal single-cell matching with unpaired feature sets, *Bioinformatics* 36 (Supplement_2) (2020) i919–i927.
- [219] Peng T., Chen G., Tan K. GLUER: integrative analysis of single-cell omics and imaging data by deep neural network. *bioRxiv*. 2021.
- [220] De Donno C., Hediye-Zadeh S., Wagenstetter M., Moïnfar A.A., Zappia L., Lotfollahi M., et al. Population-level integration of single-cell datasets enables multi-scale analysis across samples. *bioRxiv*, 2022.
- [221] W. Kopp, A. Akalin, U. Ohler, Simultaneous dimensionality reduction and integration for single-cell ATAC-seq data using deep learning, *Nat. Mach. Intell.* 4 (2) (2022) 162–168.
- [222] J. Lakkis, A. Schroeder, K. Su, M.Y. Lee, A.C. Bashore, M.P. Reilly, et al., A multi-use deep learning method for CITE-seq and single-cell RNA-seq data integration with cell surface protein prediction and imputation, *Nat. Mach. Intell.* 4 (11) (2022) 940–952.
- [223] Y. Lin, T.-Y. Wu, S. Wan, J.Y. Yang, W.H. Wong, Y. Wang, scJoint integrates atlas-scale single-cell RNA-seq and ATAC-seq data with transfer learning, *Nat. Biotechnol.* (2022) 1–8.