

Received December 16, 2020, accepted December 28, 2020, date of publication December 31, 2020, date of current version January 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3048309

Recent Advances in Variational Autoencoders With Representation Learning for Biomedical Informatics: A Survey

RUOQI WEI^{ID} AND AUSIF MAHMOOD^{ID}, (Member, IEEE)

Department of Computer Science and Engineering, University of Bridgeport, Bridgeport, CT 06604, USA

Corresponding author: Ruoqi Wei (ruoqiwei@my.bridgeport.edu)

ABSTRACT Variational autoencoders (VAEs) are deep latent space generative models that have been immensely successful in multiple exciting applications in biomedical informatics such as molecular design, protein design, medical image classification and segmentation, integrated multi-omics data analyses, and large-scale biological sequence analyses, among others. The fundamental idea in VAEs is to learn the distribution of data in such a way that new meaningful data with more intra-class variations can be generated from the encoded distribution. The ability of VAEs to synthesize new data with more representation variance at state-of-art levels provides hope that the chronic scarcity of labeled data in the biomedical field can be resolved. Furthermore, VAEs have made nonlinear latent variable models tractable for modeling complex distributions. This has allowed for efficient extraction of relevant biomedical information from learned features for biological data sets, referred to as unsupervised feature representation learning. In this article, we review the various recent advancements in the development and application of VAEs for biomedical informatics. We discuss challenges and future opportunities for biomedical research with respect to VAEs.

INDEX TERMS Deep learning, variational autoencoders (VAEs), data representation, generative models, unsupervised learning, representation learning, latent space, biomedical informatics.

I. INTRODUCTION

Over the past decade, there has been a remarkable increase in the amount of available large-scale biomedical data such as molecule compound structures [1], DNA/protein sequencing [2], [3], computer tomography (CT)/magnetic resonance imaging (MRI) [4], [5], and electronic health record (EHR) [6], [7], among others. Gaining insights and knowledge from heterogeneous, high-dimensional, and complex, biomedical data remains a key challenge in transforming bioinformatics research.

Recent advances in several factors have led to an increased interest in the use of artificial intelligence (AI) approaches [8], [9], that have greatly improved the performance of biomedical data analyses. However, due to the scarcity of labeled training data, data generation is a fundamental problem in several areas of deep learning. This is especially useful in imbalanced dataset problems and few-shot learning where a few classes may have low

representation in the dataset [10]–[12]. On the other hand, due to the powerful development of deep learning techniques in recent years, such as Convolutional Neural Networks (CNN) [13]–[16], the ability to learn meaningful nonlinear feature embeddings with little or no supervision has become a key improvement toward applying AI to the enormous unlabeled data acquired in the world, where a system, fed with raw data, provides its own representations. However, the CNN is mainly designed to automatically and adaptively learn features of spatial hierarchies for object-classification tasks [8]. Recently, Bengio *et al.* [17] proposed a concept of meta-priors to learn mapping from high-dimensional space to a meaningful low-dimensional embedding. In this concept, the high-dimensional inputs can be reconstructed from the low-dimensional manifold representations [18].

Recently, deep generative models have gained a lot of attention due to numerous applications in data generation [17]. Among them, variational autoencoder (VAE) [19]–[21] is regarded as one of the most popular approaches to generative modeling as well as a low-dimensional manifold representation learning. The VAE can also be regarded as

The associate editor coordinating the review of this manuscript and approving it for publication was Kumaradevan Punithakumar^{ID}.

a mixture of an encoder and a decoder Bayesian network. The encoder maps an input data (e.g., an image) x to a latent vector z , and then, the decoder maps the latent vector z back to image or data space [20]. The VAEs are able to learn the smooth latent representations of the input data [17] produced by the encoder and thus generate new meaningful samples, balancing the dataset with more intra-class variants in an unsupervised manner via the decoder. In addition, a key benefit of VAEs is the ability to control the distribution of the latent representation vector z , which can combine VAEs with representation learning to further improve the downstream tasks [18], [22]. Moreover, the generated image quality and diversity are improved by the existing VAE-variants such as β -VAE [23] and InfoVAE [24], which combine VAEs with disentanglement learning, GMVAE [90] and VaDE [91], which give the VAE the ability for classification with unsupervised clustering, f-VAEGAN-D2 [92] and Zero-VAE-GAN [93], which combine VAEs with GANs and few-shot learning, S -VAE [94], which combines VAEs with spherical latent representation, VQ-VAE [95], which combines VAEs with discrete latent representation, VAE-GAN [96], which combines VAEs and GANs to generate a high-quality image, and S3VAE [97], which combines VAEs with disentangled representations of sequential data. Therefore, apart from the VAE being used as a powerful generative model, it is particularly important that its excellent nonlinear latent feature representation learning idea be used to produce a series of new research directions and various applications in biomedical informatics.

Due to the above characteristics of VAEs, current VAE research in biomedical informatics focuses primarily in two directions 1) data generation approach and 2) representation learning approach. In this article, we aim to provide a concise and insightful discussion of the latest advances in applying VAEs to biomedical informatics. We particularly highlight the most important techniques in successfully applying VAEs in this field. Table 1 shows VAE research in biomedical informatics.

The structure of our paper is organized as follow: Section II overviews some background work about VAEs. Section III provides overview the application of VAEs in biomedical informatics. Conclusion and future work are given in Section IV and references are delineated at the end.

II. OVERVIEW OF VAEs

In this part, we first present the theory behind VAE. Additionally, we introduce the data augmentation approach and representation learning approach of VAEs. It should be emphasized that data augmentation is also the one of the results of representation learning. In many studies, the two approaches can often be utilized at the same time in order to achieve the similar goal.

A. PRELIMINARIES OF VAEs

The VAE is an unsupervised generative model that provides a principled way for performing variational inference

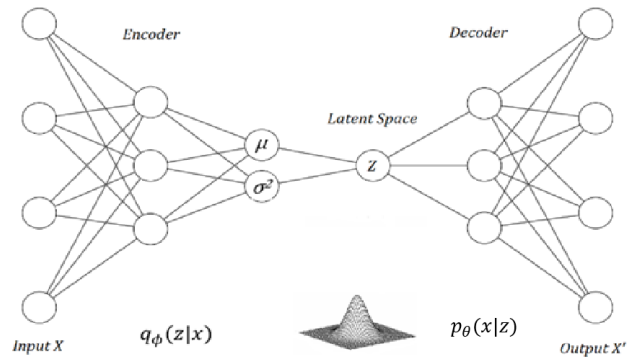


FIGURE 1. Architecture of Variational Autoencoder (VAE).

utilizing an Autoencoder (AE) architecture [98], [99]. As shown in Figure 1, the VAEs enhance a normal AE by adding a Bayesian component that learns the parameters representing the probability distribution of the data. The main difference between AE and VAE is the AE learns the compressed representation of the input, and its decompression to match the given input. In contrast, the VAE is a Bayesian model which learns the compressed representation of the AE, and constructs the parameters representing the probability distribution of the data. It can sample from this distribution and generate new input data samples. Therefore, VAE is a generative model, whereas an AE which just does reconstruction does not have an obvious generative interpretation.

The VAEs use distribution estimation and sampling to achieve generation of new data [100]. To explain this further, suppose in a continuous or discrete high-dimensional space, there is dataset $X = \{x_i\}_{i=1}^N$, and the regenerated dataset is $X' = \{x'_i\}_{i=1}^N$. Suppose encoding process produces a latent variable Z in a relatively low-dimensional space. Then, the generated model can be divided into two processes:

- 1) Latent variable Z approximates posterior distribution $q_\phi(z|x)$ - the inferred network - through the inference process:

$$p_\theta(z|x) = p_\theta(x|z)p_\theta(z)/p_\theta(x)$$

- 2) The generation process of the variable X' which the data-likelihood can be defined as:

$$p_\theta(x) = \int p_\theta(z)p_\theta(x|z)dz$$

The distribution of the latent random variable Z cannot be estimated directly, and the integral of the marginal likelihood P_θ is intractable. Therefore, the EM algorithm can't be utilized to compute the variational inference. To overcome this difficulty, VAE present an inferred model $q_\phi(z|x)$ instead of the true posterior distribution.

Specifically, the VAEs consist of the following parts: an encoder network which parameterizes a posterior distribution $q(z|x)$ of discrete latent random variables z given the input data x , a prior distribution $p(z)$, and a decoder with a distribution $p(x|z)$ over input data. Suppose we want to approximate a distribution $p(x|z)$ with some $q(z|x)$ distribution via the

TABLE 1. VAE research in biomedical informatics.

Categories	Subcategories	VAE Research References
molecular design	string representation design	[25], [26], [27], [28], [29], [30], [31]
	graph representation design	[32], [33], [34], [35], [36], [37], [38], [39], [40]
sequence datasets analyses	sequence engineering	[41], [42], [43], [44], [45], [46]
	dimensionality reduction	[47], [48], [49], [50], [51], [52], [53], [54]
	integrated multi-omics data analyses	[55], [56]
	predict effects of mutations	[57], [46]
	gene expression analyses	[55], [58], [53], [59], [60], [56], [61], [62], [54]
	DNA methylation analyses	[63], [64], [65], [66]
medical imaging and image analyses	image augmentation for downstream tasks	[67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86]
	representation learning for decision-making	[68], [70], [87], [78], [86], [88], [89]

Kullback-Leibler (KL) divergence [101], then by definition of KL,

$$\begin{aligned}
 D_{KL} [q(Z|X) \| p(Z|X)] &= \sum_Z q(Z|X) \log \left[\frac{q(Z|X)}{p(Z|X)} \right] \\
 &= E \left[\log \left[\frac{q(Z|X)}{p(Z|X)} \right] \right] = E [\log [q(Z|X)] - \log [p(Z|X)]]
 \end{aligned} \quad (1)$$

Since D_{KL} is always positive, we can conclude that:

$$\log p(X) \geq E[\log p(X|Z)] - D_{KL} [q(Z|X) \| p(Z)] \quad (3)$$

Equation (3) is an important result and is known as the Evidence Lower Bound (ELBO) [102]. In a deep neural network implementation of a VAE, equation (3) is used as the loss function during training of the network. The $E[\log p(X|Z)]$ term denotes the reconstruction i.e., the generation of output from the latent representation z . The $D_{KL} [q(Z|X) \| p(Z)]$ measures the similarity of the distribution of the latent space with the target distribution $p(z)$. Thus, the two components of equation (3) try to make the output similar to the input while keeping the distribution of the latent space as close to the target distribution $p(z)$ as possible.

The ELBO is tight if $q(z) = p(z|x)$, indicating that $q(z)$ is optimized to approximate the true posterior. For scalability to

larger datasets, we do not optimize $q(z)$ for every data point X . Instead an inference network $q(z|x)$ is introduced that is parameterized by a neural network that outputs a probability distribution for each data point X . Therefore, the final objective is to maximize:

$$\Lambda(\theta, \phi) = \mathbb{E}_{q_\phi(z|x)} [\log p_\theta(x|z)] - D_{KL}(q_\phi(z|x) \| p_\theta(z)) \quad (4)$$

According to the objective described in equation (4), after we introduced $q_\phi(z|x)$ to approximate $p_\theta(z|x)$, if we want to sample Z from $q_\phi(z|x)$, an easy choice is to assume that $q_\phi(z|x)$ obeys the Gaussian distribution and that the sampling of Z can be done in the following reparameterization way [103]:

$$z^i = \mu^i + \sigma^i * \epsilon^i \quad (5)$$

where ϵ is an auxiliary noise variable such that $\epsilon \sim N(0, 1)$ i.e., let $q(z|x)$ be a Gaussian with parameters $q(z|x)$ and $p(z)$. Then the KL divergence between $q(z|x)$ and $p(z)$ can be computed in closed form as follows:

$$\begin{aligned}
 D_{KL}[N(\mu(x), \Sigma(x)) \| N(0, 1)] &= \frac{1}{2} \sum_k ((\exp \Sigma(x)) + \mu^2(x) - 1 - \log \Sigma(x))
 \end{aligned} \quad (6)$$

B. DATA AUGMENTATION APPROACH

The explosion of big data and the development of GPUs provide sufficient training samples and advanced hardware facilities, which help to enhance the performance of deep learning. However, many applications of deep learning can only be realized under the premise of having a massive amounts of high-quality labeled data [100] and there still exist many domains lacking sufficient ideal training samples.

Data augmentation is a strategy that can significantly increase the data available for training models so that researchers do not need to actually collect new data and increase labor costs. In addition, insufficient or unbalanced data distribution can lead to over-fitting and over-parameterization problems, resulting in a significant drop in the effectiveness of learning results. To this end, previous research in data generation is augmented data by modifying images via simple transformations such as basic image processing [104]. However, naïve method has limitations such as lacks intra-class variations that cannot well represent the data variance. In order to solve these problems, attempts have been made to convert the original data variance into the feature variance.

In recent years, deep generative models have gained a lot of attention due to numerous applications in machine learning. A generative model aims to learn the features of the input and recover the original data or generate similar data from a latent space distribution, thereby increasing the variance of the dataset [105]. VAEs [19] and Generative Adversarial Networks (GANs) [106] are regarded as the two most popular approaches to generative modeling. However, VAEs do not suffer problems encountered in GANs, mainly: non-convergence causing mode collapse, and are hard to evaluate [105], [107]. What's more, VAEs have decent theoretical guarantee: first, by introducing the variational lower bound, the complicated calculation of the marginal likelihood probability is avoided. Second, by the reparameterization trick, the complicated Markov chain sampling process of latent variable is avoided.

Despite the above-mentioned advantages of VAEs, they do have some premise constraints such as compared to GANs, the samples it generates tend to be blurry and of lower quality [108]. In order to solve the problems, researchers have proposed many variations of the VAEs based on different task requirements such as representation learning, disentanglement and deep clustering with the goal of greatly improving the intra-class variations and quality of the generated data [109]. The ability of VAEs to synthesize images at state-of-art levels gives hope that the chronic scarcity of labeled data in the biomedical field can be resolved.

C. REPRESENTATION LEARNING APPROACH

The performance of models can be improved by selecting different representations to adjust the difficulty of machine learning [110]. Feature engineering [111] is one of the methods that can refine the representations from raw data.

Feature engineering refers to transforming raw data into advanced training data representations. However, in machine learning, manually selected features rely on human and professional knowledge, which is part of the most time-consuming and energy-intensive work, and its weakness is the inability to extract and organize discriminant information from the data. Moreover, although our world is inundated with data, a large part of the data is still unlabeled and unorganized. Therefore, the ability to learn meaningful nonlinear feature embeddings with little or no supervision has become a key improvement toward applying AI to the enormous unlabeled data acquired in the world. Recently, many representation learning models have been proposed based on the VAEs where the goal is to learn mapping from high-dimensional space to a meaningful low-dimensional embedding. Furthermore, they can learn useful disentangled representations automatically. The representation learning of VAEs is done by the meta-priors proposed by Bengio et al [17]. The goal of representation learning is to be useful for downstream tasks. The most important meta-prior is called “disentanglement” which is an unsupervised learning technique that breaks down, or disentangles, each feature into narrowly defined variables and encodes them as separate dimensions [17]. Assuming that the data is generated from independent factors of variation, and if the VAE is trained to reconstruct the sample well, then the latent space between the encoder and decoder keeps the important information of the original data. Intuitively, a factorial code disentangles the individual elements that were originally mixed in the sample, just as humans recognize complex things by disentangling independent elements. If the dimensions of the latent vector are independent of each other, it is factorial disentangled, i.e., a good representation. VAEs have made such nonlinear latent variable models tractable for modeling complex distributions, and efficient extraction of relevant biological information from learned features for biological data sets, referred to as unsupervised representation learning.

III. APPLICATION OF VAEs IN BIOMEDICAL INFORMATICS

Over the past decade, there has been a remarkable increase in the amount of available biomedical data available. Data types can include images [112], audio [113], textual information [1], high-dimensional omics data [2], heterogeneous data [114], and other information from wearable devices. Recent advances in several factors have led to increased interest in the use of VAE approaches within the biomedical informatics and pharmaceutical industry [9], [115]. It is particularly important that VAE not only be used as a powerful generative model but also its excellent nonlinear latent feature representation learning. However, using VAEs for validating and visualizing learning in biological datasets is particularly challenging and remains in its infancy. In the following sections, we aim to provide a concise and insightful discussion of the latest advances in applying VAEs to bioinformatics. We particularly highlight the most important techniques in

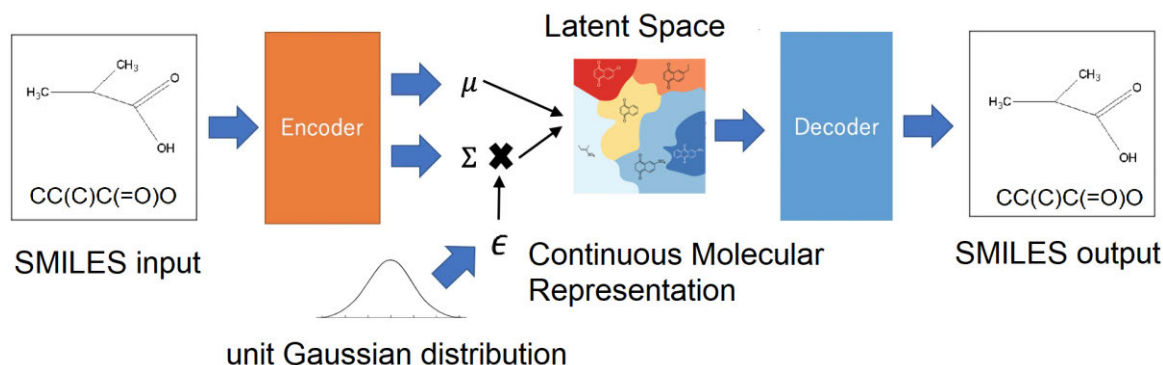


FIGURE 2. A diagram of the VAE used for molecular design.

successfully applying VAEs in this field: 1) molecular design; 2) sequence datasets analyses; and 3) medical imaging and image analyses.

A. MOLECULAR DESIGN

VAEs has numerous applications in drug discovery, including de novo molecular design. The general application of VAEs in compound dataset is to generate new chemical/molecule structures. Compound screening refers to the process of selecting compounds with high activity for a specific target through standardized experimental methods. In order to discover and optimize molecules, a chemical space of drug-like molecules estimated to be 10^{23} - 10^{60} must be searched. Screening out compounds that meet the activity index [116] is a time consuming process. Moreover, during the optimization process, adjusting one property by changing the molecular structure often has negative effects on another property [31]. VAEs can accelerate the development of this process, with $p(x|z)$ learned by the decoder aiding in appropriate chemical/molecule structure generation.

In VAEs, chemicals/molecules are represented as continuous and differentiable vectors residing on a probabilistic manifold. By grouping molecules according to properties of interest, new molecules with desired profiles could be generated by decoding latent vectors from the organized continuous latent space back to discrete molecules. For a given molecule, we can sample nearby latent spaces to decode similar molecules. As we increase this distance, increasingly dissimilar molecules can be decoded. A diagram of the VAE used for molecular design is shown in Figure 2.

The two main molecular representations used by previously reported generative models are 1) a string notation called Simplified Molecular Input Line Entry System (SMILES) [117], and 2) graphs [118]. Figure 2 shows that given a starting from a discrete molecular representation, such as a SMILES string, the encoder network converts each molecule into a vector in the latent space, which is effectively a continuous molecular representation. Given a point in the latent space, the decoder network produces a corresponding

```
Input:
Cc1ccc(cc1C)C(=O)Nc2ccc(cc2)C(=O)OC
Label:
Cc1ccc(cc1C)C(=O)Nc2ccc(cc2)C(=O)OC
Output:
Cc1cccccc1)CC(=O)Nc2cccccc2)CC000)C
```

FIGURE 3. A molecule represented as a string notation called Simplified Molecular Input Line Entry System (SMILES).



FIGURE 4. Molecule represented as a graph.

SMILES string. A molecule represented as a string notation and as graph are shown in Figure 3 and Figure 4, respectively. Note that the sampling results are generated by the code provided in the original papers [25], [32]. Table 2 shows literatures for the application of VAEs in molecular design.

Initial work in generative models for chemical/molecule compound structures focused on SMILES based methods. Gómez-Bombarelli *et al.* [25] implemented a VAE-based method that can convert discrete variables of a molecule to multidimensional continuous variables to generate new molecules. The encoder converts the discrete variables of the molecule into continuous variables, and the decoder converts these continuous variables back into discrete variables. Continuous variables of molecules can automatically generate new chemical structures by decoding vectors from the latent space. However, above methods often produce outputs that are not valid. Kusner *et al.* [26] used SMILES grammar, extracted from the training set, to improve the quality of the latent space. Although they saw an improvement, the models still struggled with invalid outputs and undesirable chemical structures such as large carbon rings and uncommon

TABLE 2. Literatures for the application of VAEs in molecular design.

Author & Year & Ref	Description	Data Type	Dataset
Gómez-Bombarelli et al.2018 [25]	Capable of generate new molecules representation	string	ZINC and QM9
Kusner et al.2017 [26]	Utilize grammar to improve the molecules design	string	ZINC
Schiff et al.2020 [27]	Novel encoding the latent space for 3D topological structure of molecules	string	ZINC
Pang et al.2020 [28]	Learns energy-based prior model with SMILES molecules in latent space	string	ZINC
Samanta et al.2020 [29]	A metric for calculating molecular similarity	string	ZINC
Yan et al.2020 [30]	Re-balancing VAE Loss to generate more valid SMILES molecules	string	ZINC
Tavakoli et al.2020 [119]	Continuous representation of molecules using graph VAE	graph	ZINC and QM9
Jin et al. 2018 [32]	Utilize tree structure to implement the graph representation	graph	ZINC
Liu et al. 2018 [33]	Capable of generating graph representations sequentially	graph	ZINC and QM9 and CEPDB
Simonovsky et al. 2018 [34]	Generating graphs from a continuous representation in the latent space	graph	QM9 and ZINC
Shervani-Tabar et al. 2020 [35]	Computing the statistics of molecular properties given small size training data set	graph	QM9
Nesterov et al. 2020 [36]	Generating 3-d molecular structures and explore molecular domains within a continuous low-dimensional representation of the molecules.	graph	QM9
Mohammadi et al. 2019 [31]	Incorporates a penalty term on the decoder of the VAE results in improvement in latent space quality	string	ZINC and PCBA
Court et al. 2020 [120]	A deep learning pipeline for generation of 3D crystal structures and prediction of their properties	electron-density maps	Materials Project
Ragoza et al. 2020 [37]	Modeling 3D molecular structures using VAE	graph	MolPort
Mahmood et al. 2020 [38]	A VAE-based masked graph model for molecular graphs	graph	QM9 and ChEMBL
Koge et al. 2020 [39]	Molecular embedding learning that combines metric learning and VAEs	graph	QM9
Kwon et al. 2020 [40]	Generates compressed graph representation for scalable molecular graphs	graph	ChEMBL

functional groups. Since complex non-linear methods were used, molecules possessing desired properties could be located in multiple locations in the latent space, thus molecule generation with an entire profile of properties is difficult. Mohammadi *et al.* [31] proposed the Penalized Variational Autoencoder (PVAE), which incorporates a penalty term on the decoder of the VAE and operates directly on SMILES strings. It demonstrates that the PVAE results in a significant improvement in latent space quality and transferability to new chemistry over all previous VAE approaches. Schiff *et al.* [27] examined the latent space of a VAE trained on molecular SMILES representations, and demonstrate how well the VAE's latent space encodes 3D topological structure of molecules. Pang *et al.* [28] proposed a VAE to learn energy-based prior model with SMILES molecules in latent space. Samanta *et al.* [29] trained a VAE that provides a rapid and novel metric for calculating molecular similarity. Yan *et al.* [30] propose a re-balancing VAE Loss to generate more valid SMILES molecules.

As a result of challenges encountered with SMILES-based methods, attention has shifted to graph-based methods with the proposal that it could be a superior molecular representation for VAE approaches [121]. Jin *et al.* [32] developed a graph based VAE that operates on a vocabulary of subgraphs extracted from the training set. Their method greatly improved the quality of the latent space over previous

approaches. Liu *et al.* [33] proposed a variational autoencoder model in which both encoder and decoder are graph-structured, i.e., a sequential generative model for graphs built from a VAE with Gated Graph Neural Networks (GGNN) for the application of molecule generation. This approach achieved state-of-the-art generation and optimization results. Simonovsky and Komodakis [34] addressed the problem of generating graphs from a continuous embedding in the context of VAEs, which combines standard graph matching algorithm to align the output to the ground truth. Tavakoli and Baldi [119] proposed a generative model in the form of a VAE which operates on the 2D-graph structure of molecules in order to continuously represent molecules. Shervani-Tabar and Zabaras [35] presented a VAE-based framework for computing the statistics of molecular properties given small size training data set. Nesterov *et al.* [36] proposed an extended version of the VAE, which allows to efficiently generate 3-d molecular structures and explore molecular domains within a continuous low-dimensional representation of the molecules. Ragoza *et al.* [37] proposed a VAE based model and a fitting algorithm to generate 3D molecular structures by converting continuous grids to discrete molecular structures. Mahmood *et al.* [38] proposed a mask map model based on VAEs, which can generate novel molecular graphs by iteratively generating a subset of the graph components. Kwon *et al.* [40] constructed a VAE-based model to

compress graph representation for scalable molecular graph generation. While the literature related to the creation of novel molecules using VAE models has proliferated, similar works that strive to do the same for inorganic crystal structures are less common but are on the rise. Court *et al.* [120] present an VAE-based deep-representation learning pipeline for geometrically optimized 3-D crystal structures that simultaneously predicts the values of eight target properties. Graph-based molecule sampled from VAE Gaussian prior distribution is shown in Figure 5. Note that the sampling results are generated by the code provided in the original paper [32].

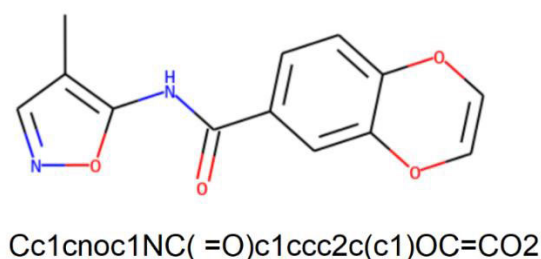


FIGURE 5. Graph-based molecule sampled from VAE Gaussian prior distribution.

In addition to generating new molecules, there are also studies on representation learning for decision-making based on VAEs. Koge *et al.* [39] proposed a method of molecular embedding learning using a combination of VAE and metric learning. This method can simultaneously maintain the consistency of the relationship between molecular structural features and physical properties, resulting in better predictions.

B. SEQUENCE DATASET ANALYSES

Another application of VAEs in bioinformatics is the application to sequence data such as genetic and amino acid datasets. Using deep learning tools for DNA/amino acid analyses usually requires converting sequences to numbers because many deep learning algorithms cannot work with categorical data directly. We can do this by one-hot encoding our representations dataset. A one-hot encoding is a representation of categorical variables as binary vectors. Figure 6 shows an example of a wild-type sequence represented as one-hot $m \times n$ matrix. Here m is the number of categories and n is the length of the sequence. Red dots correspond to a certain category at

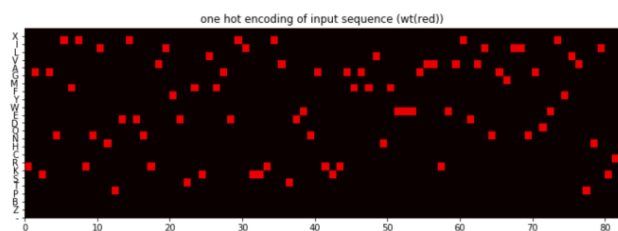


FIGURE 6. A amino acid sequence represented as one-hot matrix.

a certain position being utilized in an amino acid. Note that the sampling results are generated by the code provided in the original paper [46].

As a generative model, The VAEs can be used for data augmentation with more meaningful intra-class variations. Therefore, the basic applications of VAEs in sequence data are functional sequence engineering. Moreover, the VAEs can be used as feature dimensionality reduction technique. The main benefit of feature dimensionality reduction is to eliminate any redundant features and noise, which can improve the accuracy of prediction or generalization ability and support the interpretability of research results. Feature dimensionality reduction methods can be divided into two categories: supervised and unsupervised techniques [122]. Supervised techniques such as filter techniques [123], multivariate wrapper techniques [124] and embedded feature reduction techniques [125] select relevant features based on their ability to compute differences between groups. Unsupervised techniques such as Principal Component Analysis (PCA) [126], Independent Component Analysis (ICA) [127], AE, VAE and Coordinate-Based Meta-Analysis (CBMA) [128] are capable of selecting relevant features based on the results of interest. VAE utilizes the AE structure as a data pre-processing approach to generating representations that represent the structure of input data, and reduce its complexity, but does not reduce the quality or performance of the data.

In addition, since the structure of a VAE is a pre-expandable structure, it can easily handle the integration of multiple heterogeneous data. Therefore, the VAEs can be successfully utilized in a more representation learning setting to learn the representation in complex integrative analyses of data, and eventually lead to more stable and accurate results [55]. Specifically, for many sequence databases there are different information sets besides sequence information can be obtained. Such as The Cancer Genome Atlas (TCGA) including gene expression levels information, so they can be used for more in-depth biomedical analyses such as predict effects of mutations, estimate gene expression levels, and DNA methylation analyses. A pipeline for building a VAE-based system in specific biomedical sequences analyses is shown in Figure 7.

The biggest strength of VAE in representation learning is that it makes nonlinear latent variable models easy to handle and can model complex distributions. This has allowed for efficient extraction of relevant biomedical information from learned features for biological data sets, referred to as unsupervised feature representation learning. Since sequence space is exponentially large, and experiments are costly and time consuming, accurate computational methods are essential for sequence annotation and design. In sequence analyses, every possible higher order interaction between sequences needs explicit incorporation of a unique free parameter that must be estimated. However, traditional methods such as pairwise model for sequence analyses are still unable to model higher-order effects. Rather than describing them by explicit inclusion of parameters for each type interaction, it is possible

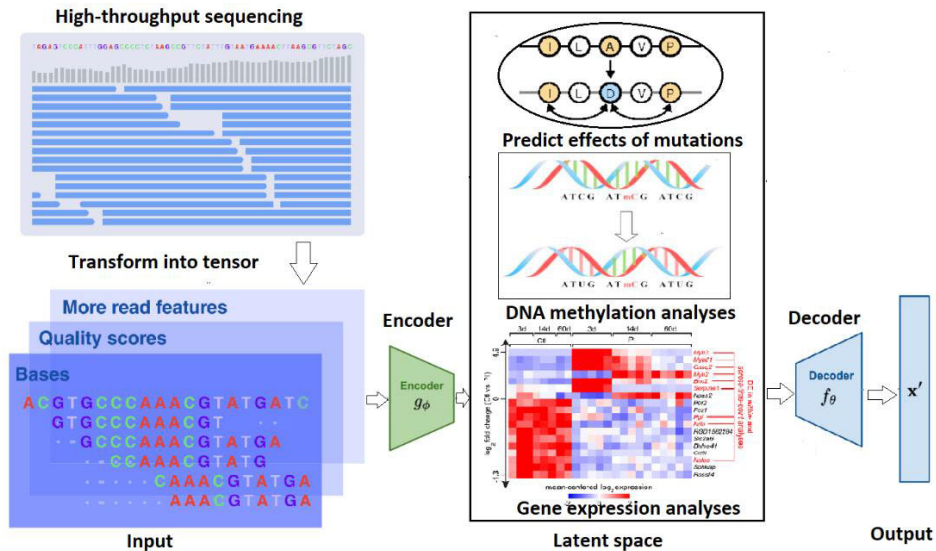


FIGURE 7. VAEs in sequence analyses Source: Adapted from [8].

to instead implicitly capture higher-order correlations by means of latent variables. Latent variables models posit hidden factors of variation that explain observed data and involve joint estimation of both hidden variables for each data point as well as global parameters describing how these hidden variables affect the observed. Towards this goal, numerous machine learning models have been developed to feature learning from evolutionary sequence data. Two widely used models for the analyses of genetic data, PCA and admixture analyses [129] can be cast as latent variable models with linear dependencies. Although these linear latent variable models are restricted in the types of correlations that they can model, replacing their linear dependencies with flexible nonlinear transformations can in principle allow the models to capture arbitrary order correlations between observed variables [57]. Learning robust predictive models based on sequence data is a key step in the development of precision medicine. Extracting meaningful low-dimensional feature representations from molecular data is the key to successfully solving high-dimensional molecular problems. Recent advances in VAEs have made such nonlinear latent variable models tractable for modeling complex distributions. This has allowed for efficient extraction of relevant biomedical information from learned features for biological data sets, referred to as unsupervised feature representation learning. Figure 8 shows the contrast between traditional models that factorize dependency in sequence families with a nonlinear latent variable z in a VAE model that can jointly influence many positions at the same time.

The main research areas based on the application of VAEs on sequence datasets include:

- 1) functional sequence engineering [41]–[45], [46];
- 2) sequence structures dimensionality reduction [47], [51]–[53];
- 3) integrated multi-omics data analyses [55];

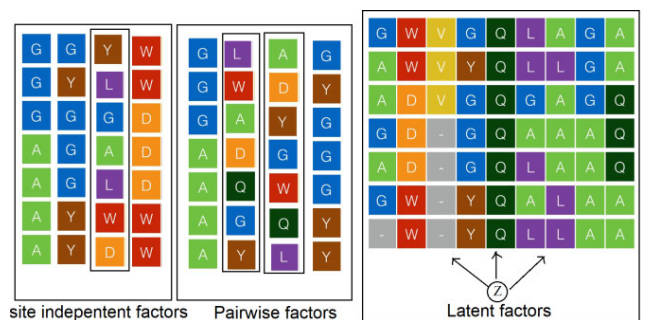


FIGURE 8. A nonlinear latent variable model VAE captures higher-order dependencies in proteins.

- 4) predict effects of mutations [46], [57];
- 5) gene expression analyses [53], [58]–[60] and
- 6) DNA methylation analyses [55], [63]–[65].

Table 3 shows literatures for the application of VAEs in sequence analyses. Note that if the dataset is public, we show the name of the data set. If the dataset can only be accessed by the owner or needs appropriate permission, we show it as private.

1) FUNCTIONAL SEQUENCE ENGINEERING

Protein engineering is of increasing importance in modern therapeutics. VAEs provide an alternative and potentially complementary approach capable of exploiting the information available in sequence and structure databases. Such conditional generation is of particular interest for protein design where it is frequently desirable to maintain a particular function while modifying a property such as stability or solubility. Guo *et al.* [41] proposed an approach to generating functionally-relevant three-dimensional structures of a protein and show the promise of generative models in directly revealing the latent space for sampling novel tertiary structures. Killoran *et al.* [42] propose VAE-based methods

TABLE 3. Literatures for the application of VAEs in sequence analyses.

Author & Year & Ref	Description	Application	Dataset
Guo et al.2020 [41]	Capable of generate three-dimensional structures of a protein	Protein design	Private
Killoran et al. 2017 [42]	Capable of generate DNA sequences and tune them to have desired properties	DNA design	chr1 and exons-50-400-500 and protein-binding
Davidsen et al.2019 [43]	Capable of generate unseen sequences and learn the rules of VDJ recombination	DNA design and predict cohort frequency	ImmuneACCESS
Greener et al. 2018 [44]	Capable of generate protein based on required structural properties	Protein design	UniRef90
Hawkins-Hooker et al. 2020 [45]	Capable of generate new protein structures	Protein design	InterPro
Ding et al. 2018 [47]	Dimensionality reduction and visualization of structures in single-cell gene expression data	Sequence structures dimensionality reduction	Gene Expression Omnibus
Dony et al.2020 [48]	Capable of learning biologically informative embeddings by VAE with a VAMP prior	Sequence structures dimensionality reduction	Private
Oh et al.2020 [49]	Convert high-dimensional microbiome data into robust low-dimensional representations	Sequence structures dimensionality reduction	Colorectal
Way et al.2020 [50]	Compress data dimensionalities and learns complementary biological representations	Sequence structures dimensionality reduction	TCGA, GTEx, and TARGET
Wang et al. 2018 [51]	Capture and visualize the low-dimensional structures in scRNA-seq data	Sequence structures dimensionality reduction	PBMC3k
Simidjievsk et al. 2019 [55]	Reduce the genetic integrated data-dimensionality and leading to good representation	Integrated multi-omics data analyses and Gene expression analyses	Private multi-omics and clinical data
Ding et al.2019 [52]	Embed scRNA-seq data into low-dimensional hyperspherical spaces	Sequence structures dimensionality reduction	Single Cell Portal
Riesselman et al.2018 [57]	Predict the effects of mutations and non-linear feature learning in protein sequences data	Predict the effects of mutations	Private
Sinai et al.2017 [46]	Capable of generate new protein structures and predict the effects of mutations	Protein design and predict the effects of mutations	PABP
Grønbech et al. 2018 [58]	Capable of estimate expected gene expression levels and model representations	Gene expression analyses	PBMC and MB-IM
Hu et al.2019 [53]	Capable of learn structure of bulk gene expression data by mapping it into a low dimensional space	Sequence structures dimensionality reduction and Gene expression analyses	Private
Way et al.2017 [59]	Capable of model cancer gene expression, and capture biologically-relevant features	Gene expression analyses	Private
Dincer et al.2018 [60]	Capable of extract a low-dimensional feature representation from gene expression data	Gene expression analyses	GEO
He et al.2020 [56]	Representation learning of unlabeled heterogeneous omics data; Predict anti-cancer drug sensitivity from somatic mutations via the assistance of gene expressions	Integrated multi-omics data analyses and Gene expression analyses	CCLL
Kim et al.2020 [61]	Capable of extract genes significant features that can be used for patient survival prediction	Gene expression analyses	TCGA cancer
Bica et al.2020 [62]	Data representation and separation of biological factors for gene expression data	Gene expression analyses	CEL-seq2
Titus et al.2018 [63]	Accurately classifies disease subtypes on the learned latent methylome	DNA methylation analyses	TCGA BRCA
Titus et al. 2018 [64]	Representation learning on methylation patterns among clinical subtypes of tumors	DNA methylation analyses	Private
Wang et al. 2018 [65]	Learns a biological relevant methylome	DNA methylation analyses	Genomic Data Commons
Qiu et al. 2020 [66]	Representation learning for transcriptome and methylome data	DNA methylation analyses and Sequence structures dimensionality reduction	TCGA and DNA methylation data
Rashid et al.2019 [54]	Dimensionality reduction in single cell genomic data	Sequence structures dimensionality reduction and Gene expression analyses	Oligodendroglioma

to generate DNA sequences and adjust them to have the required characteristics. Davidsen *et al.* [43] proposed a VAE model parameterized by deep neural networks for T cell receptor (TCR) repertoires and show that simple VAE models can perform accurate cohort frequency estimation, learn the rules of VDJ recombination, and generalize well to unseen sequences. Another approach has been to use trained models to directly move towards sampling sequences that have measurable and desired attributes. Greener *et al.* [44] indicated that CVAEs are able to carry out protein design tasks by conditioning output sequences on desired structural properties. Hawkins-Hooker *et al.* [45] proposed a VAE-based model that can be used to generate aligned sequence input and raw sequence input, and showed that both can reproduce the amino acid usage pattern of the family. Guo *et al.* [41] proposed a VAE-based interpretative approach to generate functionally related 3D structures of proteins.

2) SEQUENCE STRUCTURES DIMENSIONALITY REDUCTION

At present, the utilize of dimensionality reduction to explain the structure of single-cell sequencing data is still a challenge. Sequence datasets usually have high dimensionality, performing dimensionality reduction, followed by visualization or downstream analyses has become a key strategy for exploratory data analyses in sequence datasets, e.g., Single-cell RNA sequencing (scRNA-seq) [130], [131]. Ding *et al.* [47] provided a VAE based computational framework to compute low-dimensional embeddings of scRNA-seq data while preserving global structure of the high-dimensional measurements. Wang and Gu [51] proposed the deep variational autoencoder for the unsupervised dimensionality reduction and visualization of scRNA-seq data and find the nonlinear hierarchical feature representations of the original data as well as provide better representations for very rare cell populations in the 2D visualization. Hu and Greene [53] evaluated the performance of a simple VAE model developed for bulk gene expression data under various parameter settings and found substantial performance differences with hyperparameter tuning. Ding and Regev [52] introduced a hyperspherical VAE based model to embed cells into low-dimensional hyperspherical spaces, as a more accurate representation of the data which overcomes cell crowding and facilitates interactive visualization of large datasets. Rashid *et al.* [54] introduced a VAEs based method converts the single-cell genome data into a feature space with a smaller dimension, so that the tumor subpopulations can be divided more effectively. The analysis of the encoding feature space reveals the evolutionary relationship between cell subpopulations. In addition, this method tolerates dropout of gene expression in single-cell RNA sequencing datasets. Dony *et al.* [48] found that VAE with a VAMP prior is capable of learning biologically informative embeddings without compromising on generative properties. Oh and Zhang [49] utilized various autoencoders to convert high-dimensional microbiome data into robust low-dimensional representations, and apply machine learning

classification algorithms to the learned representations. Way *et al.* [50] compared different methods include VAE to compress data dimensionalities and learn complementary biological representations.

3) INTEGRATED MULTI-OMICS DATA ANALYSES

Multi-omics data covers a wide range of data generated from the transcriptome, proteome, genome, epigenome and metabolome. Due to a comprehensive understanding of diseases and human health, it is necessary to explain the molecular intricacy and multi-level variations of the multi-omics data. By providing an integrated system-level approach, the availability of multi-omics data has revolutionized the fields of biology and medicine. Simultaneously analyzing and integrating different data types can better help researchers study the mechanism and internal structure of biomedical processes at the molecular level. Simidjievski *et al.* [55] proposed VAE-based architecture is utilized to integrate heterogeneous cancer data types which including multi-omics and clinical data, and perform extensive analysis. They designed a detailed computational framework for VAEs as a system, which can correctly model the nonlinear representations in the integration data, while still being able to reduce the data dimension and acquire good representation learning. He and Xie [56] adopt a VAE-based approach on the unlabeled heterogeneous omics data in predicting anti-cancer drug sensitivity from somatic mutations via the assistance of gene expressions

4) PREDICT EFFECTS OF MUTATIONS

Accurate prediction of the effects of sequence variation is a major goal in biological research. The traditional method is through site independent or pairwise models [132]–[135]. However, they are unable to model higher-order correlations. The most recent improvement in model performance was made possible by the VAE. Riesselman *et al.* [57] used VAEs to infer biological sequences from large multi-sequence alignments and predict the effects of mutations and organize sequence information, all while being grounded with biologically motivated architecture learned in unsupervised fashion. The accuracy achieved with the VAE approaches exceeds the site-independent or pairwise interaction models. Sinai *et al.* [46] generated protein sequences using VAEs with the goal of predicting how mutations affect protein function. A nonlinear latent variable model VAE captures higher-order dependencies in proteins is shown in Figure 8.

5) GENE EXPRESSION ANALYSES

Since the VAEs model can perform meaningful representation learning in the latent space, it can generate and explore hypothetical gene expression under the perturbation of different molecules and genetic sequences. As an example, this could be used to predict a tumor's response to specific therapies, or to characterize complex gene expression activations existing in differential proportions in different tumors. The low dimensional latent space generated by the

VAEs have been used to reveal complex patterns and novel biological signals from large-scale gene expression data and carry out drug response predictions. Grønbech *et al.* [58] proposed a novel variational auto-encoder-based method for analyses of scRNA-seq data. It is able to estimate the expected gene expression levels and model a latent representation for each cell, with support for several count likelihood functions. They use a variant of the variational auto-encoder which has a priori clustering in the latent space. Way *et al.* [59] proposed the research to which a VAE can be trained to model cancer gene expression, and capture biologically-relevant features. Dincer *et al.* [60] utilized a VAEs-based model to extract low-dimensional features from public unlabeled gene expression data, and the effectiveness of these low-dimensional representations is demonstrated. And they further demonstrate that the learned features are related to drug response predictions. Kim *et al.* [61] introduced a VAE-based survival prediction model to extract genes' significant features that can be used for patient survival prediction. Bica *et al.* [62] proposed a method based on VAEs, which can model cell differentiation by constructing low-dimensional meaningful representations from complex and high-dimensional gene expression data.

6) DNA METHYLATION ANALYSES

DNA methylation is a well-defined epigenetic biomarker for monitoring cancer development and its treatment efficacy because of role it plays in pathways and regulation of gene expression. DNA Methylation using VAEs have been shown to learn latent representations of the DNA methylation landscape from three independent breast tumor datasets and demonstrate the feasibility of VAEs to track representative differential methylation patterns among clinical subtypes of tumors. Wang and Wang [65] showed that DNA methylation data can be used with VAEs that can learn meaningful signals from merged datasets that credibly represent the subtype of the samples. Qiu *et al.* [66] showed a representation learning framework for transcriptome and methylome data based on VAE.

C. MEDICAL IMAGING AND IMAGE ANALYSES

Some of the popular medical imaging techniques for the early detection and diagnosis of diseases include magnetic resonance imaging (MRI), computed tomography (CT), ultrasound, X-rays, mammography and positron emission tomography (PET). Imaging enables scientists to see the phenotype and behavior of host, organ, tissue, cell, and subcellular components. Digital image analyses reveal hidden biology and pathology as well as drug action mechanisms. VAEs have also succeeded in biological image analyses, and many studies show superior performance. The main research area based on the VAE use in medical imaging datasets includes:

1) Medical image data augmentation for downstream tasks include image classification [67], [70], [78], [79], image segmentation [71]–[75], [86], image restoration [84], [85], and image reconstruction [71], [80]–[82].

2) Improve the interpretability of representation learning [78], [86], [88], [89].

Table 4 shows the literatures for the application of VAEs in medical imaging and image analyses. Note that if the dataset is public, we show the name of the data set. If the dataset can only be accessed by the owner or needs appropriate permission, we show it as private.

1) MEDICAL IMAGE AUGMENTATION FOR DOWNSTREAM TASKS

Deep learning models have shown significant success in analyzing clinical images and utilizing them for downstream tasks such as medical image classification. However, many applications of deep learning can only be realized on the premise of having a large amount of labeled data, and it is impractical to obtain large amounts of data in clinical medicine. Inspired by humans' ability to learn quickly from a small number of samples, few-sample learning has become a hot topic in artificial intelligence. The few-sample learning solves the problem of the limited training data set through the process of imitating human brain, thus getting closer to the application scenario. Data augmentation is a strategy that allows researchers to have a significantly increased amount of data available for training models without actually collecting new data. The easiest way for data generation is through simple image transformation [104] such as rotation, color jittering, noise addition, and image translation. However, these methods only generate duplicate versions of the original data, so the entire data set still lacks intra-class variation. Since VAEs learn variations in the data, they can be used for data augmentation to effectively improve performance of the downstream tasks [136], [137]. This is especially useful in imbalanced dataset problems and few-shot learning where a few classes may have low representation in the dataset [138].

To the medical image classification tasks, Pesteie *et al.* [67] proposed conditional generative model based on VAEs that learns the latent space independent of the labels to further improve the image classification task. Biffi *et al.* [70] proposed a 3D convolution generation model based on VAEs, which can be utilized to automatically classify images of cardiac disease patients related to structural remodeling and also improves the interpretability of the VAEs to further enhancing the clinical value. Uzunova *et al.* [78] proposed learns the meaningful perturbations of pathological regions by defining plausible pathology perturbation based on replacing the pixel values with healthy looking tissue learned by VAEs, which also improve the interpretability of the black box classifiers. Uzunova *et al.* [79] learns the entire variability of healthy data and detect pathologies by a conditional VAE. Díaz Berenguer *et al.* [68] proposed explainable semi-supervised representation learning based on VAE for COVID-19 diagnosis from CT Imaging. Thiagarajan *et al.* [69] proposed a VAE-based model to predict the type of lesion. This model is used to distinguish different types of lesions by extracting interpretable features.

TABLE 4. Literatures for the application of VAEs in medical imaging and image analyses.

Author & Year & Ref	Description	Application	Dataset
Pesteie et al.2019 [67]	Learns the embedding space independent of the labels to further improve the image classification task	Classification	A-Transverse US Images of the Spine and B-MR Volumes of the Brain
Díaz Berenguer et al.2020 [68]	COVID-19 classification from CT imaging based on VAE	Classification and Representation learning	Covid-ct-dataset
Thiagarajan et al.2020 [69]	Distinguish different types of lesions by extracting interpretable features	Classification	ISIC 2018 lesion diagnosis challenge
Biffi et al.2018 [70]	Automatically classify images of cardiac disease patients	Classification and Representation learning	Private
Biffi et al. 2019 [71]	Reconstructs a 3D segmentation from 2D standard cardiac views representation	3D segmentation and reconstruction	Private
Sedai et al. 2017 [72]	Capable of segment the optic cup	Segmentation	Private
Chen et al. 2019 [73]	Brain tumor and white matter hyperintensities segmentation combined with semi-supervised learning	Segmentation	BraTS and White Matter Hyperintensities Challenge
Myronenko et al. 2018 [74]	Automated segmentation of brain tumors from 3D MRIs	Segmentation	BraTS Challenge
Ouyang et al.2019 [75]	Few-shot learning with segmentation	Segmentation	MICCAI Multi-modality Whole Heart Segmentation Challenge
Qian et al.2020 [76]	Estimates object shape uncertainty in medical images	Segmentation	Private
Volokitin et al.2020 [77]	3D MR brain images reconstruction and brain anatomy segmentation	Reconstruction and Segmentation	Human Connectome Project
Lafarge et al.2020 [87]	Unsupervised representation learning for computational pathology	Representation learning	TUPAC-ROI
Uzunova et al. 2019 [78]	Learns the meaningful perturbations of pathological regions	Classification and Representation learning	Retinal OCT and LPBA40
Uzunova et al. 2018 [79]	Learns the entire variability of healthy data and detect pathologies by a conditional VAE	Classification	Private
Edupuganti et al. 2019 [80]	Analyze the uncertainty in compressed MR image recovery	Reconstruction	Private
Gomez et al. 2019 [81]	Combine overlapping image patches into a fused reconstruction of the object of interest	Reconstruction	Private
Mostapha et al. 2019 [82]	Automatic quality control of structural MR images	Reconstruction	BCP and EBDS
Tudosiu et al.2020 [83]	Data compression with image fidelity	Reconstruction	ADNI
Zilvan et al. 2019 [84]	Utilize denoising convolutional VAE as feature extractor and also as denoiser for disease detection tasks	Restoration	Plantvillage
Prakash et al.2020 [85]	Utilize fully-convolutional VAE to generate diverse and plausible denoising solutions	Restoration	herlev
Chartsias et al.2019 [86]	Learns decomposed meaningful spatial disentangled representation of cardiac imaging data, and improved semi supervised segmentation results.	Segmentation and Representation learning	ACDC
Zhao et al.2019 [88]	Unsupervised representation learning with supervised regression	Representation learning	Private
Puyol-Antón et al. 2019 [89]	Unsupervised representation learning with supervised regression	Representation learning	CMR

To the medical image segmentation tasks, Myronenko [74] proposed a semantic segmentation network for automatically segmenting brain tumors based on the VAE architecture and 3D MRI data sets. This network not only improves the performance of automated segmentation but also consistently acquires excellent training accuracy for random initialization. To the few shot learning problem in medical image segmentation, Ouyang *et al.* [75] proposed a method that combines

VAEs and domain adaptation of transfer learning techniques to learn the feature latent space shared by both the source domain and the target domain, which can be utilized in the segmentation process with only a few target set. Another approach for few shot learning is the used of semi-supervised learning in segmentation with VAEs: Sedai *et al.* [72] proposed a semi-supervised VAEs-based method to segment optical cup, and utilized a small number of labeled data to

accurately localize the anatomical structure. Chen *et al.* [73] proposed a VAEs-based semi-supervised image segmentation method for Brain tumor and white matter hyperintensities segmentation. Qian *et al.* [76] build a novel VAE for estimating object shape uncertainty in medical images.

To the medical image reconstruction tasks, Biffi *et al.* [71] proposed a CVAE based model, which can reconstruct a high-resolution 3D segmentation image of LV myocardium from three segmentations of a 2D cardiac image. Edupuganti *et al.* [80] introduced a VAEs-based method to analyze the uncertainty in compressed MR image recovery. Gomez *et al.* [81] proposed an image reconstruction architecture based on β -VAE, which can combine numerous overlapping image patches into a fusion reconstruction of the real fetal ultrasound images. Mostapha *et al.* [82] proposed VAEGAN-based framework to the automatic quality control of structural MR images. Tudosiu *et al.* [83] proposed a model based on VQ-VAE, which can effectively encode full-resolution 3D brain volume, compressing data to 0.825% of the original size, while maintaining image fidelity. Volokitin *et al.* [77] proposed a method to model the 3D MR brain volumes distribution by combining the 2D samples obtained by VAEs and the Gaussian model.

Image restoration is the task of removing unwanted noise and distortions, giving us clean images that are closer to the true but unknown signal. Prakash *et al.* [85] introduced a fully-convolutional VAE to generate diverse and plausible denoising solutions, sampled from a learned posterior. It cannot only produce diverse results, but can also be leveraged for downstream processing. Zilvan *et al.* [84] proposed denoising convolutional VAE as feature extractor and also as a denoiser for disease detection tasks.

2) REPRESENTATION LEARNING FOR DECISION-MAKING

The VAEs can apply to large dataset with high dimensional features, due to high computational complexity. And it can learn meaningful fixed-size low-dimensional feature representation in unsupervised manner. Lafarge *et al.* [87] proposed the VAE-based framework to learn orientation-wise disentangled generative factors of histopathology images. The result shows the aggregated representation of subpopulations of cells produces higher performances in subsequent tasks.

However, due to the scarcity of labelled training data of medical datasets, one drawback of such unsupervised deep neural networks is the lack of interpretability. Since many neurons in neural networks turns into a many-to-many entangled mess, the learned latent representations are usually not directly interpretable. In medical research area, model interpretability is not only important but also necessary, since clinicians are increasingly relying on data-driven solutions for patient monitoring and decision-making.

Established methods like guided backpropagation [139] and gradCAMs [140] try to gain an insight into how neural networks learn and create intuitive visualizations based on the learned network weights. However, they are mostly

heuristic and depend on the architecture of the neural network. Another possibility is by using perturbations [141]. e.g., Uzunova *et al.* [78] tackle the interpretability problem of generating plausible explanations by meaningful perturbations using VAEs. Recent scientific advances have combined the interpretability of supervised settings with the power of VAEs. Zhao *et al.* [88] proposed a VAEs based unified probabilistic model for learning the latent space of imaging data and performing supervised regression. Their results allow for intuitive interpretation of the structural developmental patterns of the human brain. Puyol-Antón *et al.* [89] proposed to use a VAE featuring a regression loss in the latent space to simultaneously learn efficient representations of cardiac function and map their change with regard to differences in systolic blood pressure (a measure of hypertension). Chartsias *et al.* [86] proposed VAE based model to learn decomposed meaningful spatial disentangled representation of cardiac imaging data, and leveraging these for improved semi-supervised segmentation results.

IV. CONCLUSION AND FUTURE OPPORTUNITIES

In this article, we comprehensively summarized the essential concepts of VAEs and its applications in the areas of molecular design, sequence dataset analyses, and medical imaging analyses. It is particularly important that VAE not only be used as a powerful generative model, but also its excellent nonlinear latent feature representation learning capability be used to produce a series of new research directions, and various applications in biomedical informatics. We particularly highlighted the most important techniques in successfully applying VAEs in the following fields: 1) molecular design, 2) sequence dataset analyses, and 3) medical imaging and image analyses.

In molecular design, VAEs can produce molecular generation in both the SMILES string, and graphs of molecular representations. The encoder converts the discrete variables of the molecule into continuous variables, and the decoder converts these continuous variables back into discrete variables. However, since molecules possessing desired properties could be located in multiple locations in the latent space, molecular generation with an entire profile of properties is difficult. Therefore, future work may include an extension to tune VAEs to have desired properties and further experimental validation. In addition, compared with sequence dataset analyses, medical imaging and image analyses, there are few studies on representation learning for decision-making approach using VAEs on the molecular dataset. We hope that more research will appear in this field in the future.

In sequence dataset analyses, VAE can assert not only its advantages as a generative model and feature dimensionality reduction technique, but also its benefits in nonlinear representation learning. In particular, VAEs provide an alternative and potentially complementary approach capable of exploiting the information available in sequences and tune them to have desired properties. Moreover, VAEs can be used for more in-depth biomedical analyses, such as

integrated multi-omics data analyses, prediction of the effects of mutations, gene expression analyses, and DNA methylation analyses, so that more accurate and stable results can be achieved. In addition, the VAEs can be used as a feature dimensionality reduction technique. The main benefit of feature dimensionality reduction is to eliminate any redundant features and noise, which can improve the accuracy of prediction or generalization ability and support the interpretability of research results. Although the VAE model in sequence data set analysis has challenges, such as lack of interpretability and increased potential for overfitting, VAE may be increasingly important in the high-throughput design and biological sequence annotation. The applications of VAEs in sequence dataset analyses are currently in a relatively early stage of development. Therefore, its application in this area as well as the improvement of its generality such as increase in quality of feature reduction approaches [142], and interpretability will be more extensive in the future.

In medical imaging and image analyses, since VAEs learn variations in the data, therefore, they can be used for image augmentation to effectively improve the performance of downstream tasks such as medical image classification, segmentation, image restoration, and reconstruction. This is especially useful in imbalanced dataset problems, and few-shot learning where a few classes may have low representation in the dataset. However, high sensitivity to input parameters and high running time are some disadvantages of the deep learning based VAEs [143]. In the future, we expect that there will be more research achieving high classification accuracy with low computational complexity by the use of VAEs. On the other hand, even though the VAE can learn meaningful fixed-size low-dimensional feature representation in an unsupervised manner, one drawback of such unsupervised deep neural networks is the lack of interpretability.

Currently, the application of deep learning in EHR data for clinical informatics research has increased. Compared to traditional methods in clinical informatics, deep learning methods offer better performance and require less time, data preprocessing, and representation learning costs. However, there is no research on the application of VAEs in EHR data yet [7]. EHR data is very heterogeneous and can include multiple data types. This is in sharp contrast to the homogeneity of the original input data type with only image pixels or only Neuro-Linguistic Programming (NLP) characters. However, the key requirement of EHR dataset analyses is to acquire recognized and meaningful findings from such high-dimensional, sparse, and complex clinical data. The feature representation learning ability in a higher-level abstraction of VAEs mentioned in this article may be utilized to acquire meaningful and powerful research results from EHRs. In the future, we expect that there will be more research focusing on the application of VAEs in EHR data analyses.

One of the most important goals of representation learning for biomedical informatics research is to make decisions based on VAEs. There are many biomedical applications that require decisions. In this survey, we have described

researchers have accordingly attempted to use VAEs for decision-making applications, such as mutation-effect prediction for genomic sequences [57]. To make decisions based on VAEs, researchers implicitly appeal to Bayesian decision theory, such as taking the action that minimizes expected loss under the posterior distribution [144]. Lopez *et al.* [145] proposed a three-step procedure for using VAE for decision-making.

One drawback of such unsupervised deep neural networks is the lack of interpretability. Since many neurons in neural networks turn into a many-to-many entangled complexity, the learned latent representations are usually not directly interpretable. In medical research, model interpretability is not only important but also necessary, as clinicians are increasingly relying on data-driven solutions for patient monitoring and decision making. Recent scientific advances have combined the interpretability of supervised settings with the power of VAEs. However, more research is needed to ensure the worst-case performance of VAE-based models in diagnostic problems before they are used in high-stakes decision-making scenarios.

We provided a comprehensive review of the VAEs and its current representation learning research directions in biomedical informatics. We also provided an interactive visual code on GitHub for some papers to make this survey beneficial. We hope that this will be useful in improving the state-of-the-art leading to research breakthroughs in related fields.

REFERENCES

- [1] J. J. Irwin and B. K. Shoichet, "ZINC—A free database of commercially available compounds for virtual screening," *J. Chem. Inf. Model.*, vol. 45, no. 1, pp. 177–182, 2005.
- [2] K. Tomczak, P. Czerwińska, and M. Wizerowicz, "The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge," *Contemp. Oncol.*, vol. 19, no. 1A, pp. A68–A77, 2015.
- [3] S. Hunter, R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman, D. Binns, P. Bork, U. Das, L. Daugherty, and L. Duquenne, "InterPro: The integrative protein signature database," *Nucleic acids Res.*, vol. 37, no. 1, pp. D211–D215, Jan. 2009.
- [4] K. Yan, X. Wang, L. Lu, and R. M. Summers, "Deeplesion: Automated mining of large-scale lesion annotations and universal lesion detection with deep learning," *J. Med. Imag.*, vol. 5, no. 3, 2018, Art. no. 036501.
- [5] R. A. Poldrack and K. J. Gorgolewski, "OpenfMRI: Open sharing of task fMRI data," *NeuroImage*, vol. 144, pp. 259–261, Jan. 2017.
- [6] P.-Y. Wu, C.-W. Cheng, C. D. Kaddi, J. Venugopalan, R. Hoffman, and M. D. Wang, "Omic and electronic health record big data analytics for precision medicine," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 2, pp. 263–273, Feb. 2016.
- [7] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, "Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 5, pp. 1589–1604, Sep. 2018.
- [8] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean, "A guide to deep learning in healthcare," *Nature Med.*, vol. 25, no. 1, pp. 24–29, Jan. 2019.
- [9] B. Norgeot, B. S. Glicksberg, and A. J. Butte, "A call for deep-learning healthcare," *Nature Med.*, vol. 25, no. 1, pp. 14–15, Jan. 2019.
- [10] A. Mishra, S. K. Reddy, A. Mittal, and H. A. Murthy, "A generative model for zero shot learning using conditional variational autoencoders," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 2188–2196.

- [11] E. Schonfeld, S. Ebrahimi, S. Sinha, T. Darrell, and Z. Akata, "Generalized zero- and few-shot learning via aligned variational autoencoders," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8247–8255.
- [12] W. Wang, Y. Pu, V. K. Verma, K. Fan, and Y. Zhang, "Zero-shot learning via class-conditioned deep generative models," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 1–8.
- [13] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [14] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [15] A. Jin, S. Yeung, J. Jopling, J. Krause, D. Azagury, A. Milstein, and L. Fei-Fei, "Tool detection and operative skill assessment in surgical videos using region-based convolutional neural networks," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2018, pp. 691–699.
- [16] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent.* New York, NY, USA: Springer, 2015, pp. 234–241.
- [17] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.
- [18] M. Tschannen, O. Bachem, and M. Lucic, "Recent advances in autoencoder-based representation learning," 2018, *arXiv:1812.05069*. [Online]. Available: <http://arxiv.org/abs/1812.05069>
- [19] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," 2013, *arXiv:1312.6114*. [Online]. Available: <http://arxiv.org/abs/1312.6114>
- [20] C. Doersch, "Tutorial on variational autoencoders," 2016, *arXiv:1606.05908*. [Online]. Available: <http://arxiv.org/abs/1606.05908>
- [21] D. P. Kingma and M. Welling, "An introduction to variational autoencoders," 2019, *arXiv:1906.02691*. [Online]. Available: <http://arxiv.org/abs/1906.02691>
- [22] F. Locatello, S. Bauer, M. Lucic, G. Rätsch, S. Gelly, B. Schölkopf, and O. Bachem, "Challenging common assumptions in the unsupervised learning of disentangled representations," 2018, *arXiv:1811.12359*. [Online]. Available: <http://arxiv.org/abs/1811.12359>
- [23] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed, and A. Lerchner, "Beta-VAE: Learning basic visual concepts with a constrained variational framework," in *Proc. ICLR*, 2017, vol. 2, no. 5, p. 6.
- [24] S. Zhao, J. Song, and S. Ermon, "InfoVAE: Information maximizing variational autoencoders," 2017, *arXiv:1706.02262*. [Online]. Available: <http://arxiv.org/abs/1706.02262>
- [25] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik, "Automatic chemical design using a data-driven continuous representation of molecules," *ACS Central Sci.*, vol. 4, no. 2, pp. 268–276, Feb. 2018.
- [26] M. J. Kusner, B. Paige, and J. M. Hernández-Lobato, "Grammar variational autoencoder," in *Proc. 34th Int. Conf. Mach. Learn.*, vol. 70, 2017, pp. 1945–1954.
- [27] Y. Schiff, V. Chenthamarakshan, K. Natesan Ramamurthy, and P. Das, "Characterizing the latent space of molecular deep generative models with persistent homology metrics," 2020, *arXiv:2010.08548*. [Online]. Available: <http://arxiv.org/abs/2010.08548>
- [28] B. Pang, T. Han, and Y. Nian Wu, "Learning latent space energy-based prior model for molecule generation," 2020, *arXiv:2010.09351*. [Online]. Available: <http://arxiv.org/abs/2010.09351>
- [29] S. Samanta, S. O'Hagan, N. Swainston, T. J. Roberts, and D. B. Kell, "VAE-sim: A novel molecular similarity measure based on a variational autoencoder," *Molecules*, vol. 25, no. 15, p. 3446, Jul. 2020.
- [30] C. Yan, S. Wang, J. Yang, T. Xu, and J. Huang, "Re-balancing variational autoencoder loss for molecule sequence generation," in *Proc. 11th ACM Int. Conf. Bioinf., Comput. Biol. Health Informat.*, Sep. 2020, pp. 1–7.
- [31] S. Mohammadi, B. O'Dowd, C. Paulitz-Erdmann, and L. Goerlitz, "Penalized variational autoencoder for molecular design," *ChemRxiv*, 2019, doi: [10.26434/chemrxiv.7977131.v2](https://doi.org/10.26434/chemrxiv.7977131.v2).
- [32] W. Jin, R. Barzilay, and T. Jaakkola, "Junction tree variational autoencoder for molecular graph generation," 2018, *arXiv:1802.04364*. [Online]. Available: <http://arxiv.org/abs/1802.04364>
- [33] Q. Liu, M. Allamanis, M. Brockschmidt, and A. Gaunt, "Constrained graph variational autoencoders for molecule design," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 7795–7804.
- [34] M. Simonovsky and N. Komodakis, "Graphvae: Towards generation of small graphs using variational autoencoders," in *Proc. Int. Conf. Artif. Neural Netw.* New York, NY, USA: Springer, 2018, pp. 412–422.
- [35] N. Shervani-Tabar and N. Zabarab, "Physics-constrained predictive molecular latent space discovery with graph scattering variational autoencoder," 2020, *arXiv:2009.13878*. [Online]. Available: <http://arxiv.org/abs/2009.13878>
- [36] V. Nesterov, M. Wieser, and V. Roth, "3DMolNet: A generative network for molecular structures," 2020, *arXiv:2010.06477*. [Online]. Available: <http://arxiv.org/abs/2010.06477>
- [37] M. Ragoza, T. Masuda, and D. Ryan Koes, "Learning a continuous representation of 3D molecular structures with deep generative models," 2020, *arXiv:2010.08687*. [Online]. Available: <http://arxiv.org/abs/2010.08687>
- [38] O. Mahmood, E. Mansimov, R. Bonneau, and K. Cho, "Masked graph modeling for molecule generation," *ChemRxiv*, 2020, doi: [10.26434/chemrxiv.13143167.v1](https://doi.org/10.26434/chemrxiv.13143167.v1).
- [39] D. Koge, N. Ono, M. Huang, M. Altaf-Ul-Amin, and S. Kanaya, "Embedding of molecular structure using molecular hypergraph variational autoencoder with metric learning," *Mol. Informat.*, 2020, doi: [10.1002/minf.202000203](https://doi.org/10.1002/minf.202000203).
- [40] Y. Kwon, D. Lee, Y.-S. Choi, K. Shin, and S. Kang, "Compressed graph representation for scalable molecular graph generation," *J. Cheminformatics*, vol. 12, no. 1, pp. 1–8, Dec. 2020.
- [41] X. Guo, S. Tadepalli, L. Zhao, and A. Shehu, "Generating tertiary protein structures via an interpretative variational autoencoder," 2020, *arXiv:2004.07119*. [Online]. Available: <http://arxiv.org/abs/2004.07119>
- [42] N. Killoran, L. J. Lee, A. Delong, D. Duvenaud, and B. J. Frey, "Generating and designing DNA with deep generative models," 2017, *arXiv:1712.06148*. [Online]. Available: <http://arxiv.org/abs/1712.06148>
- [43] K. Davidsen, B. J. Olson, W. S. DeWitt, J. Feng, E. Harkins, P. Bradley, and F. A. Matsen, "Deep generative models for t cell receptor protein sequences," *eLife*, vol. 8, Sep. 2019, Art. no. e46935.
- [44] J. G. Greener, L. Moffat, and D. T. Jones, "Design of metalloproteins and novel protein folds using variational autoencoders," *Sci. Rep.*, vol. 8, no. 1, Dec. 2018, Art. no. 16189.
- [45] A. Hawkins-Hooker, F. Depardieu, S. Baur, G. Couairon, A. Chen, and D. Bikard, "Generating functional protein variants with variational autoencoders," *BioRxiv*, 2020, doi: [10.1101/2020.04.07.029264](https://doi.org/10.1101/2020.04.07.029264).
- [46] S. Sinai, E. Kelsic, G. M. Church, and M. A. Nowak, "Variational auto-encoding of protein sequences," 2017, *arXiv:1712.03346*. [Online]. Available: <http://arxiv.org/abs/1712.03346>
- [47] J. Ding, A. Condon, and S. P. Shah, "Interpretable dimensionality reduction of single cell transcriptome data with deep generative models," *Nature Commun.*, vol. 9, no. 1, pp. 1–13, Dec. 2018.
- [48] L. Dony, M. König, D. S. Fischer, and F. J. Theis, "Variational autoencoders with flexible priors enable robust distribution learning on single-cell RNA sequencing data," in *Proc. ICML*, 2020, pp. 1–8.
- [49] M. Oh and L. Zhang, "DeepMicro: Deep representation learning for disease prediction based on microbiome data," *Sci. Rep.*, vol. 10, no. 1, Dec. 2020, Art. no. 6026.
- [50] G. P. Way, M. Zietz, V. Rubinetti, D. S. Himmelstein, and C. S. Greene, "Compressing gene expression data using multiple latent space dimensionalities learns complementary biological representations," *Genome Biol.*, vol. 21, no. 1, pp. 1–27, Dec. 2020, doi: [10.1186/s13059-020-02021-3](https://doi.org/10.1186/s13059-020-02021-3).
- [51] D. Wang and J. Gu, "VASC: Dimension reduction and visualization of single-cell RNA-seq data by deep variational autoencoder," *Genomics, Proteomics Bioinf.*, vol. 16, no. 5, pp. 320–331, Oct. 2018.
- [52] J. Ding and A. Regev, "Deep generative model embedding of single-cell RNA-Seq profiles on hyperspheres and hyperbolic spaces," *BioRxiv*, vol. 1, Jan. 2019, Art. no. 853457.
- [53] Q. Hu and C. S. Greene, "Parameter tuning is a key part of dimensionality reduction via deep variational autoencoders for single cell RNA transcriptomics," in *Proc. Biocomputing*, Nov. 2018, pp. 362–373.
- [54] S. Rashid, S. Shah, Z. Bar-Joseph, and R. Pandya, "Dhaka: Variational autoencoder for unmasking tumor heterogeneity from single cell genomic data," *Bioinformatics*, p. btz095, 2019, doi: [10.1093/bioinformatics/btz095](https://doi.org/10.1093/bioinformatics/btz095).
- [55] N. Simidjievski, C. Bodnar, I. Tariq, P. Scherer, H. A. Terre, Z. Shams, M. Jammik, and P. Liü, "Variational autoencoders for cancer data integration: Design principles and computational practice," *Frontiers Genet.*, vol. 10, p. 1025, Dec. 2019.
- [56] D. He and L. Xie, "A cross-level information transmission network for predicting phenotype from new genotype: Application to cancer precision medicine," 2020, *arXiv:2010.04824*. [Online]. Available: <http://arxiv.org/abs/2010.04824>

- [57] A. J. Riesselman, J. B. Ingraham, and D. S. Marks, "Deep generative models of genetic variation capture the effects of mutations," *Nature Methods*, vol. 15, no. 10, pp. 816–822, Oct. 2018.
- [58] C. H. Grønbech, M. F. Vording, P. N. Timshel, C. K. Sønderby, T. H. Pers, and O. Winther, "scVAE: Variational auto-encoders for single-cell gene expression data," *BioRxiv*, vol. 4, May 2018, Art. no. 318295.
- [59] G. P. Way and C. S. Greene, "Extracting a biologically relevant latent space from cancer transcriptomes with variational autoencoders," *BioRxiv*, vol. 5, Jan. 2017, Art. no. 174474.
- [60] A. B. Dincer, S. Celik, N. Hiranuma, and S.-I. Lee, "DeepProfile: Deep learning of cancer molecular profiles for precision medicine," *BioRxiv*, vol. 4, Apr. 2018, Art. no. 278739.
- [61] S. Kim, K. Kim, J. Choe, I. Lee, and J. Kang, "Improved survival analysis by learning shared genomic information from pan-cancer data," *Bioinformatics*, vol. 36, no. 1, pp. i389–i398, Jul. 2020.
- [62] I. Bica, H. Andrés-Terré, A. Cvejić, and P. Liū, "Unsupervised generative and graph representation learning for modelling cell differentiation," *Sci. Rep.*, vol. 10, no. 1, Dec. 2020, Art. no. 9790.
- [63] A. J. Titus, C. A. Bobak, and B. C. Christensen, "A new dimension of breast cancer epigenetics," in *Proc. Joint Conf. Biomed. Eng. Syst. Technol.*, 2018, pp. 1–6.
- [64] A. J. Titus, O. M. Wilkins, C. A. Bobak, and B. C. Christensen, "Unsupervised deep learning with variational autoencoders applied to breast tumor genome-wide DNA methylation data with biologic feature extraction," *BioRxiv*, vol. 1, Jan. 2018, Art. no. 433763.
- [65] Z. Wang and Y. Wang, "Exploring DNA methylation data of lung cancer samples with variational autoencoders," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2018, pp. 1286–1289.
- [66] Y. L. Qiu, H. Zheng, and O. Gevaert, "Genomic data imputation with variational auto-encoders," *GigaScience*, vol. 9, no. 8, Aug. 2020, Art. no. g1aa082.
- [67] M. Pesteie, P. Abolmaesumi, and R. N. Rohling, "Adaptive augmentation of medical data using independently conditional variational autoencoders," *IEEE Trans. Med. Imag.*, vol. 38, no. 12, pp. 2807–2820, May 2019.
- [68] A. D. Berenguer *et al.*, "Explainable-by-design semi-supervised representation learning for COVID-19 diagnosis from CT imaging," 2020, *arXiv:2011.11719*. [Online]. Available: <http://arxiv.org/abs/2011.11719>
- [69] J. J. Thiagarajan, B. Venkatesh, D. Rajan, and P. Sattigeri, "Improving reliability of clinical models using prediction calibration," in *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis*. New York, NY, USA: Springer, 2020, pp. 71–80.
- [70] C. Biffi, O. Oktay, G. Tarroni, W. Bai, A. De Marvao, G. Doumou, M. Rajchl, R. Bedair, S. Prasad, S. Cook, D. O'Regan, and D. Rueckert, "Learning interpretable anatomical features through deep generative models: Application to cardiac remodeling," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2018, pp. 464–471.
- [71] C. Biffi, J. J. Cerrolaza, G. Tarroni, A. de Marvao, S. A. Cook, D. P. O'Regan, and D. Rueckert, "3D high-resolution cardiac segmentation reconstruction from 2D views using conditional variational autoencoders," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1643–1646.
- [72] S. Sedai, D. Mahapatra, S. Hewavitharanage, S. Maetschke, and R. Garnavi, "Semi-supervised segmentation of optic cup in retinal fundus images using variational autoencoder," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2017, pp. 75–82.
- [73] S. Chen, G. Bortsova, A. G.-U. Juárez, G. van Tulder, and M. de Bruijne, "Multi-task attention-based semi-supervised learning for medical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2019, pp. 457–465.
- [74] A. Myronenko, "3D MRI brain tumor segmentation using autoencoder regularization," in *Proc. Int. MICCAI Brainlesion Workshop*. New York, NY, USA: Springer, 2018, pp. 311–320.
- [75] C. Ouyang, K. Kamnitsas, C. Biffi, J. Duan, and D. Rueckert, "Data efficient unsupervised domain adaptation for cross-modality image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2019, pp. 669–677.
- [76] L. Qian, J. Chen, T. Urakov, W. Gu, and L. Liang, "CQ-VAE: Coordinate quantized VAE for uncertainty estimation with application to disk shape analysis from lumbar spine MRI images," 2020, *arXiv:2010.08713*. [Online]. Available: <http://arxiv.org/abs/2010.08713>
- [77] A. Volokitin, E. Erdil, N. Karani, K. Can Tezcan, X. Chen, L. Van Gool, and E. Konukoglu, "Modelling the distribution of 3D brain MRI using a 2D slice VAE," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2020, pp. 657–666.
- [78] H. Uzunova, J. Ehrhardt, T. Kepp, and H. Handels, "Interpretable explanations of black box classifiers applied on medical images by meaningful perturbations using variational autoencoders," *Proc. SPIE Med. Imag. Image Process.*, vol. 10949, Mar. 2019, Art. no. 1094911.
- [79] H. Uzunova, H. Handels, and J. Ehrhardt, "Unsupervised pathology detection in medical images using learning-based methods," in *Proc. Bildverarbeitung Med.*, 2018, pp. 61–66.
- [80] V. Edupuganti, M. Mardani, S. Vasanawala, and J. Pauly, "Uncertainty quantification in deep MRI reconstruction," 2019, *arXiv:1901.11228*. [Online]. Available: <http://arxiv.org/abs/1901.11228>
- [81] A. Gomez, V. Zimmer, N. Toussaint, R. Wright, J. R. Clough, B. Khanal, M. P. M. van Poppel, E. Skelton, J. Matthews, and J. A. Schnabel, "Image reconstruction in a manifold of image patches: Application to whole-fetus ultrasound imaging," in *Proc. Int. Workshop Mach. Learn. Med. Image Reconstruct.* New York, NY, USA: Springer, 2019, pp. 226–235.
- [82] M. Mostapha, J. Prieto, V. Murphy, J. Girault, M. Foster, A. Rumble, J. Blocher, W. Lin, J. Elison, J. Gilmore, and S. Pizer, "Semi-supervised VAE-GAN for out-of-sample detection applied to MRI quality control," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2019, pp. 127–136.
- [83] P.-D. Tudosiu, T. Varsavsky, R. Shaw, M. Graham, P. Nachev, S. Ourselin, C. H. Sudre, and M. Jorge Cardoso, "Neuromorphologically-preserving volumetric data encoding using VQ-VAE," 2020, *arXiv:2002.05692*. [Online]. Available: <http://arxiv.org/abs/2002.05692>
- [84] V. Zilvan, A. Ramdan, E. Suryawati, R. B. S. Kusumo, D. Krisnandi, and H. F. Pardede, "Denoising convolutional variational autoencoders-based feature learning for automatic detection of plant diseases," in *Proc. 3rd Int. Conf. Informat. Comput. Sci. (ICICoS)*, Oct. 2019, pp. 1–6.
- [85] M. Prakash, A. Krull, and F. Jug, "DivNoising: Diversity denoising with fully convolutional variational autoencoders," 2020, *arXiv:2006.06072*. [Online]. Available: <http://arxiv.org/abs/2006.06072>
- [86] A. Chartsias, T. Joyce, G. Papanastasiou, S. Semple, M. Williams, D. E. Newby, R. Dharmakumar, and S. A. Tsiftaris, "Disentangled representation learning in cardiac image analysis," *Med. Image Anal.*, vol. 58, Dec. 2019, Art. no. 101535.
- [87] M. W. Lafarge, J. P. W. Pluim, and M. Veta, "Orientation-disentangled unsupervised representation learning for computational pathology," 2020, *arXiv:2008.11673*. [Online]. Available: <http://arxiv.org/abs/2008.11673>
- [88] Q. Zhao, E. Adeli, N. Honnorat, T. Leng, and K. M. Pohl, "Variational autoencoder for regression: Application to brain aging analysis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* New York, NY, USA: Springer, 2019, pp. 823–831.
- [89] E. Puyol-Antón, B. Ruijsink, J. R. Clough, I. Oksuz, D. Rueckert, R. Razavi, and A. P. King, "Assessing the impact of blood pressure on cardiac function using interpretable biomarkers and variational autoencoders," 2019, *arXiv:1908.04538*. [Online]. Available: <http://arxiv.org/abs/1908.04538>
- [90] N. Dilokthanakul, P. A. M. Mediano, M. Garnelo, M. C. H. Lee, H. Salimbeni, K. Arulkumaran, and M. Shanahan, "Deep unsupervised clustering with Gaussian mixture variational autoencoders," 2016, *arXiv:1611.02648*. [Online]. Available: <https://arxiv.org/abs/1611.02648>
- [91] Z. Jiang, Y. Zheng, H. Tan, B. Tang, and H. Zhou, "Variational deep embedding: An unsupervised and generative approach to clustering," 2016, *arXiv:1611.05148*. [Online]. Available: <https://arxiv.org/abs/1611.05148>
- [92] Y. Xian, S. Sharma, B. Schiele, and Z. Akata, "F-VAEGAN-D2: A feature generating framework for any-shot learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, May 2019, pp. 10267–10276.
- [93] R. Gao, X. Hou, J. Qin, J. Chen, L. Liu, F. Zhu, Z. Zhang, and L. Shao, "Zero-VAE-GAN: Generating unseen features for generalized and transductive zero-shot learning," *IEEE Trans. Image Process.*, vol. 29, pp. 3665–3680, Jan. 2020.
- [94] T. R. Davidson, L. Falorsi, N. De Cao, T. Kipf, and J. M. Tomczak, "Hyperspherical variational auto-encoders," 2018, *arXiv:1804.00891*. [Online]. Available: <https://arxiv.org/abs/1804.00891>
- [95] A. V. D. Oord, O. Vinyals, and K. Kavukcuoglu, "Neural discrete representation learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 6306–6315.

- [96] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, "Autoencoding beyond pixels using a learned similarity metric," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 1558–1566.
- [97] Y. Zhu, M. R. Min, A. Kadav, and H. P. Graf, "S3 VAE: Self-supervised sequential VAE for representation disentanglement and data generation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 6538–6547.
- [98] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [99] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, and L. Bottou, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, no. 12, pp. 1–38, 2010.
- [100] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*. [Online]. Available: <http://arxiv.org/abs/1511.06434>
- [101] C. Varin and P. Vidoni, "A note on composite likelihood inference and model selection," *Biometrika*, vol. 92, no. 3, pp. 519–528, Sep. 2005.
- [102] X. Yang, "Understanding the variational lower bound," Tech. Rep., 2017, pp. 1–4.
- [103] D. Jimenez Rezende and S. Mohamed, "Variational inference with normalizing flows," 2015, *arXiv:1505.05770*. [Online]. Available: <http://arxiv.org/abs/1505.05770>
- [104] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, p. 60, 2019.
- [105] L. Mi, M. Shen, and J. Zhang, "A probe towards understanding GAN and VAE models," 2018, *arXiv:1812.05676*. [Online]. Available: <http://arxiv.org/abs/1812.05676>
- [106] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [107] I. Goodfellow, "NIPS 2016 tutorial: Generative adversarial networks," 2017, *arXiv:1701.00160*. [Online]. Available: <http://arxiv.org/abs/1701.00160>
- [108] J. Tomczak and M. Welling, "VAE with a VampPrior," in *Proc. Int. Conf. Artif. Intell. Statist.*, Mar. 2018, pp. 1214–1223.
- [109] R. Wei, C. Garcia, A. El-Sayed, V. Peterson, and A. Mahmood, "Variations in variational autoencoders—A comparative evaluation," *IEEE Access*, vol. 8, pp. 153651–153670, 2020.
- [110] P. Langley, "Selection of relevant features in machine learning," in *Proc. AAAI Fall Symp. Relevance*, vol. 184, 1994, pp. 245–271.
- [111] C. R. Turner, A. Fuggetta, L. Lavazza, and A. L. Wolf, "A conceptual basis for feature engineering," *J. Syst. Softw.*, vol. 49, no. 1, pp. 3–15, Dec. 1999.
- [112] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burre, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.
- [113] B. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, R. P. Paiva, I. Chouvarda, P. Carvalho, and N. Maglaveras, "A respiratory sound database for the development of automated classification," in *Proc. Int. Conf. Biomed. Health Informat.* New York, NY, USA: Springer, 2017, pp. 33–37.
- [114] A. Jha, C. DesRoches, E. Campbell, K. Donelan, S. Rao, T. Ferris, A. Shields, S. Rosenbaum, and D. Blumenthal, "Use of electronic health records in US hospitals," *New England J. Med.*, vol. 360, no. 16, pp. 1628–1638, 2009.
- [115] J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer, and S. Zhao, "Applications of machine learning in drug discovery and development," *Nature Rev. Drug Discovery*, vol. 18, no. 6, pp. 463–477, 2019.
- [116] H. Zheng, H. Yang, D. Gong, L. Mai, X. Qiu, L. Chen, X. Su, R. Wei, and Z. Zeng, "Progress in the mechanism and clinical application of cilostazol," *Current Topics Med. Chem.*, vol. 19, no. 31, pp. 2919–2936, Jan. 2020.
- [117] D. Weininger, A. Weininger, and J. L. Weininger, "SMILES. 2. Algorithm for generation of unique SMILES notation," *J. Chem. Inf. Model.*, vol. 29, no. 2, pp. 97–101, May 1989.
- [118] B. Sanchez-Lengeling and A. Aspuru-Guzik, "Inverse molecular design using machine learning: Generative models for matter engineering," *Science*, vol. 361, no. 6400, pp. 360–365, Jul. 2018.
- [119] M. Tavakoli and P. Baldi, "Continuous representation of molecules using graph variational autoencoder," 2020, *arXiv:2004.08152*. [Online]. Available: <http://arxiv.org/abs/2004.08152>
- [120] C. J. Court, B. Yildirim, A. Jain, and J. M. Cole, "3-D inorganic crystal structure generation and property prediction via representation learning," *J. Chem. Inf. Model.*, vol. 60, no. 10, pp. 4518–4535, Oct. 2020.
- [121] S. R. Bowman, L. Vilnis, O. Vinyals, A. M. Dai, R. Jozefowicz, and S. Bengio, "Generating sentences from a continuous space," 2015, *arXiv:1511.06349*. [Online]. Available: <http://arxiv.org/abs/1511.06349>
- [122] B. Mwangi, T. S. Tian, and J. C. Soares, "A review of feature reduction techniques in neuroimaging," *Neuroinformatics*, vol. 12, no. 2, pp. 229–244, Apr. 2014.
- [123] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Jan. 2003.
- [124] R. Kohavi and H. George, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, pp. 1–2, Jan. 1997.
- [125] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc. B, Methodol.*, vol. 58, no. 1, pp. 267–288, Jan. 1996.
- [126] G. H. Dunteman, *Principal Components Analysis*, vol. 69. New York, NY, USA: Sage, 1989.
- [127] A. Hyvärinen, "Survey on independent component analysis," *Neural Comput. Surveys*, vol. 2, pp. 94–128, 1999.
- [128] C. Rottschy, R. Langner, I. Dogan, K. Reetz, A. R. Laird, J. B. Schulz, P. T. Fox, and S. B. Eickhoff, "Modelling neural correlates of working memory: A coordinate-based meta-analysis," *NeuroImage*, vol. 60, no. 1, pp. 830–846, Mar. 2012.
- [129] N. Patterson, A. L. Price, and D. Reich, "Population structure and eigenanalysis," *PLoS Genet.*, vol. 2, no. 12, p. e190, 2006.
- [130] A. Butler, P. Hoffman, P. Smibert, E. Papalex, and R. Satija, "Integrating single-cell transcriptomic data across different conditions, technologies, and species," *Nature Biotechnol.*, vol. 36, no. 5, pp. 411–420, May 2018.
- [131] M. D. Luecken and F. J. Theis, "Current best practices in single-cell RNA-seq analysis: A tutorial," *Mol. Syst. Biol.*, vol. 15, no. 6, 2019, Art. no. e8746.
- [132] R. Wei, A. Wang, and L. Huang, "Correlation of kidney deficiency syndrome with methylenetetrahydrofolate reductase gene polymorphism in postmenopausal women," *J. New Chin. Med.*, vol. 2, p. 33, May 2012.
- [133] K.-M. Yang, J. Jia, L.-N. Mao, C. Men, K.-T. Tang, Y.-Y. Li, H.-X. Ding, and Y.-Y. Zhan, "Methylenetetrahydrofolate reductase C677T gene polymorphism and essential hypertension: A meta-analysis of 10,415 subjects," *Biomed. Rep.*, vol. 2, no. 5, pp. 699–708, Sep. 2014.
- [134] M. Hecht, Y. Bromberg, and B. Rost, "Better prediction of functional effects for sequence variants," *BMC Genomics*, vol. 16, no. S8, p. 1, Dec. 2015.
- [135] T. A. Hopf, J. B. Ingraham, F. J. Poelwijk, C. P. I. Schärfe, M. Springer, C. Sander, and D. S. Marks, "Mutation effects predicted from sequence co-variation," *Nature Biotechnol.*, vol. 35, no. 2, p. 128, 2017.
- [136] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding data augmentation for classification: When to warp?" in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov. 2016, pp. 1–6.
- [137] T. DeVries and G. W. Taylor, "Dataset augmentation in feature space," 2017, *arXiv:1702.05538*. [Online]. Available: <http://arxiv.org/abs/1702.05538>
- [138] Z. Wan, Y. Zhang, and H. He, "Variational autoencoder based synthetic data generation for imbalanced learning," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Nov. 2017, pp. 1–7.
- [139] J. Tobias Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for simplicity: The all convolutional net," 2014, *arXiv:1412.6806*. [Online]. Available: <http://arxiv.org/abs/1412.6806>
- [140] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.
- [141] R. C. Fong and A. Vedaldi, "Interpretable explanations of black boxes by meaningful perturbation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 3429–3437.
- [142] S. F. Ershad and S. Hashemi, "To increase quality of feature reduction approaches based on processing input datasets," in *Proc. IEEE 3rd Int. Conf. Commun. Softw. Netw.*, May 2011, pp. 367–371.
- [143] S. Fekri-Ershad, "Pap smear classification using combination of global significant value, texture statistical features and time series features," *Multimedia Tools Appl.*, vol. 78, no. 22, pp. 31121–31136, 2019.
- [144] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. New York, NY, USA: Springer, 2013.
- [145] R. Lopez, P. Boyeau, N. Yosef, M. I. Jordan, and J. Regier, "Decision-making with auto-encoding variational bayes," 2020, *arXiv:2002.07217*. [Online]. Available: <http://arxiv.org/abs/2002.07217>



RUOQI WEI received the M.Sc. degree in computer science from the Department of Computer Science and Engineering, University of Bridgeport, Bridgeport, CT, USA, in 2016, where she is currently pursuing the Ph.D. degree. Her research interests include machine learning, deep learning, transfer learning, few-shot learning, computer vision, and biomedical informatics.



AUSIF MAHMOOD (Member, IEEE) is currently a Professor with the Department of Computer Science and Engineering. He is also the Director of the School of Engineering, University of Bridgeport. His research interests include computer vision, machine and deep learning, computer architecture, and parallel processing.

...