A Tour of Unsupervised Deep Learning for Medical Image Analysis

Nripendra Kumar Singh and Khalid Raza*
Department of Computer Science, Jamia Millia Islamia, New Delhi
kraza@jmi.ac.in

November 10, 2018

Abstract

Interpretation of medical images for diagnosis and treatment of disease from complex, high-dimensional and heterogeneous data remain a key challenge in transforming healthcare. In the last few years, both supervised and unsupervised deep learning achieved promising results in the field of medical imaging and image analysis. Unlike supervised learning which isbiased towards how it is being supervised and manual efforts to create class label for the algorithm, unsupervised learning derive insights directly from the data itself, group the data and help to make data driven decisions without any external bias. This chapter covers various unsupervised models applied to medical image analysis, including autoencoders and its several variants, Restricted Boltzmann machines, Deep belief networks, Deep Boltzmann machine and Generative network. Future research opportunities and challenges of unsupervised techniques for medical image analysis has also been discussed.

Keywords: Unsupervise learning; medical image analysis; autoencoders; restricted Boltzmann machine; Deep belief network

1. Introduction

Medical imaging techniques, such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), mammography, ultrasound, and X-ray, are used diagnostic system for the early detection, diagnosis, and treatment of various complex diseases (Wani & Raza, 2018). In the clinic, the images are mostly interpreted by human experts such as radiologists and physicians. Because of major variations in pathology and the potential fatigue of human experts, researchers and doctors have begun to benefit from computer-assisted interventions. The advancement in machine learning techniques, including deep learning, and availability of computing infrastructure through cloud computing, have given fuel to the field of computer-assisted medical image analysis and computer-assisted diagnosis (CAD). Deep learning is about learning representations, i.e, learning intermediate concept or features which are important to capture dependencies from input variables to output variables in supervised learning, or between subsets of variables in unsupervised learning. Both supervised and unsupervised machine learning approaches are widely applied in medical image analysis; each of them has their own pros and cons. Some of widely used supervised (deep) learning algorithms are Feedforward Neural Network (FFNN), Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Support Vector Machine (SVM) and so on. There are many scenarios where human supervisions are unavailable, inadequate or biased and therefore, supervised learning algorithm cannot be

directly used. Unsupervised learning algorithms, including its deep architecture, give a big hope with lots of advantages and have been widely applied in several areas of medical and engineering problems including medical image analysis.

This chapter presents unsupervised deep learning models, its applications to medical image analysis, and discusses opportunities and future challenges in the area.

2. Why Unsupervised Learning?

In the majority of machine learning projects, the work flow is designed in a supervised way, where algorithm is guided by us what to do and what not to! In such supervised architecture the potential of the algorithms are limited in three ways, *(i)* A huge manual effort to create labels and *(ii)* Biases related to how it is being supervised, which probabilities the algorithm to think for other corner cases during problem solving, and *(iii)* Reduce the scalabilities of target function at hand.

To intelligently solve these issues, unsupervised machine learning algorithm can be used. Unsupervised machine learning algorithms not only derives insights directly from data and group the data but also uses these insights for data-driven decisions making. Also, unsupervised models are more robust in the sense that they act as base for several different complex tasks where these can be utilised as the holy grail of learning and classification. In fact, classification is not the only task that we do; rather other tasks such as compression, dimensionality reduction, denoising, super resolution and some degree of decision making are also performed. Therefore, it is rather more useful to construct a model without knowing what tasks will be at hand and we will use representation (or model) for. In a nutshell, we can think of unsupervised learning as preparation (preprocessing) step for supervised learning task, where unsupervised learning of representation may allow better generalization of a classifier (Jabeen, et al., 2018).

3. Taxonomy of Unsupervised Learning Tasks

In unsupervised learning, we group the unlabeled data set on the basis of underlying hidden features. By grouping data through unsupervised learning, at least we learn something about raw data.

3.1 Density estimation

Density estimation is one of the popular categories of unsupervised learning which discovers the intrinsic feature and structure of large and complex unlabeled data set via another non-probabilistic approach. Density estimation is a non-parametric method which doesn't possess much restriction and distributional assumption unlike parametric estimation.

Estimation of univariate or multivariate density function without any prior functional assumptions get almost limitless function from data. There are some widely used non-parametric methods of estimation.

3.1.1 Kernel density estimation

Kernel density estimation (KDE) uses statistical model to produce a probabilistic distribution that resemble an observed variable as a random variable. Basically, KDE is used for data smoothing, exploratory data analysis and visualization. A large number of kernels have been proposed, namely *normal Gaussian mixture model* and *multivariate Gaussian mixture model.* Some of the advantages of Kernel density estimation are:

- o No need for model specification (data speaks itself).
- o Estimation converges to any density shape with sufficient sample.
- o Easily generalizes to higher dimension.
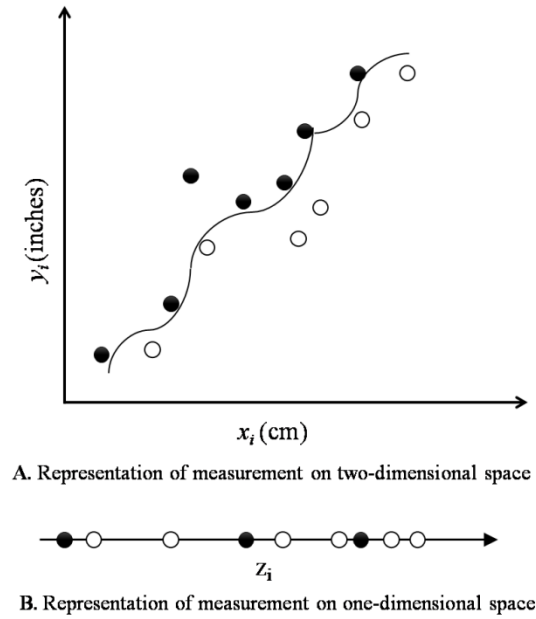- o Densities are multivariate and multimodal with irregular cluster shape.

3.1.2 Histogram density estimation

Histogram based technique mainly adds smoothness of density curve of reconstruction which can be optimized by kernel parameters and closely related to KNN density estimation algorithm (Bishop et al., 2006).

3.2 Dimensionality reduction

Why dimensionality reduction? There has been tremendous increase in deployment of sensors and various imaging technique's which are being used in industry and medical diagnosis continuously record data and store it to be analyzed later. Lots of redundancy or noises are present initially when data are captured. For example, let us take a case of a patient having bone fracture. Initially doctors suggested for X-ray images which is 2D or 3D imaging but when they do not find it helpful in diagnosis then CT scan and MRI (magnetic resonance imaging) may be taken which gives more detailed results for further right diagnosis. Now assume that an analyst sits with all this data to analyze and identified all important variables/dimensions which have most significant information's and left all unwanted part of data. This is the problem of high unwanted dimension removal and needs treatment of dimension reduction. Dimension reduction is the process of reducing higher dimension data set into lesser dimension ensuring that final reduced data must convey equivalent information concisely.

Let's look at figure shown below. It shows two-dimension x and y which are measurement of several objects in cm(x_1) and inches(y_1), if you continue to use both dimensions in machine learning problems it will introduce lots of noise in system. So, it is better to just use one-dimension (z_1) and they will convey similar information.

A. Representation of measurement on two-dimensional space

B. Representation of measurement on one-dimensional space

Fig 1. Representation of data in two dimensional and one dimensional space

There are some common methods to perform dimensionality reduction:

3.2.1 Factor analysis

Some variables in given data are highly correlated. These variables can be grouped on the basis of their correlations. This means a particular group can have highly correlated variable but have low correlation with variables of other groups. Each group represents single inherent construct or factor. As compared to data having large number of dimensions, these factors are small in number, while these factors are difficult to find. There are two methods for doing factor analysis: *(i) Exploratory Factor Analysis, (ii) Confirmatory Factor Analysis*

3.2.2 Principal component analysis

A set of variables, which are linear combination of original set of variables, performs higher dimensional space mapped to lower dimensions in such a way that variance of data in lower dimensional space is maximum. These new set of variables are known as principle components.

Let's consider a situation of two-dimensional data set, there can be only two principal components, first principal component is the most possible variation of original data and second principal component is orthogonal to the first principal component, as shown in Fig. 2.
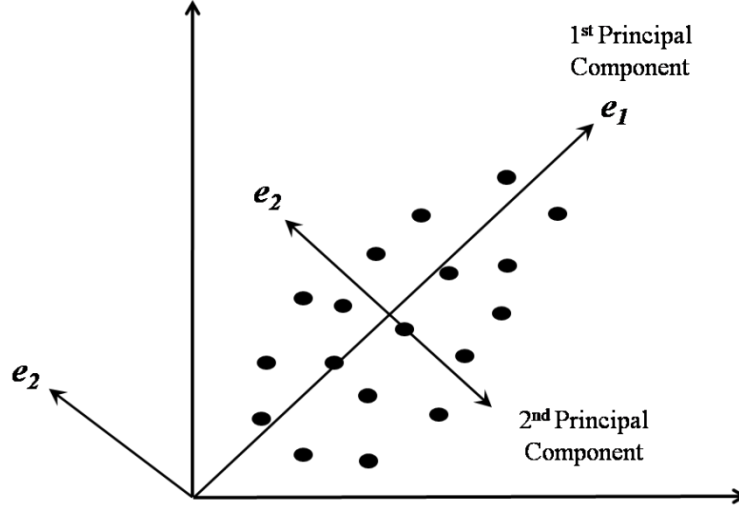
Fig. 2 Principle components on a two-dimensional data

In practice, a simple principal component analysis (PCA) can be used to construct the covariance or correlation matrix of the data and compute the eigenvectors. The eigenvectors correspond to largest eigenvalues (principal component) are used to reconstruct a large fraction of variance of original data. As result, it is left with lesser number of eigenvector and original space has been reduced. There might be chance of loss of data but it is retained by most important eigenvectors.

Consider a matrix $U(m)$ which stored empirical mean of input matrix R,

$$U(m) = \frac{1}{N}\sum_{n=1}^{N} R(m,n), where \text{ m } = (1,2,3\ldots\ldots\ldots M) \qquad (1)$$

Calculate a normalized matrix X, $X = (R - U_e)$, where e is unitary vector matrix of size N. Finally, the mean square error (E^2) is calculated in which smallest eigenvalues are removed,

$$E^2 = trace\big(cov(x)\big) - \sum_{i=1}^{L}\lambda_i = \sum_{i=\lambda+1}^{M}\lambda_i \qquad (2)$$

The *trace(A)* is the sum of all eigenvalues. Simple PCA is not capable of constructing nonlinear mapping, however, can implement nonlinear classification by using kernel techniques.

3.2.3 Kernel PCA

Kernel PCA is a nonlinear extension of conventional PCA which is designed for dimensionality reduction of nonlinear subspaces depending on magnitude of input data and problem setup. In medical image analysis, hybrid kernel PCA is frequently used to get better result in unsupervised deep learning training model. Fischer et al. (2017) proposed an unsupervised deep learning illumination invariant kernel PCA which is applied to each patch of respiratory signal extraction from X-ray fluoroscopy images leading to a set of low-dimensional embeddings.

A kernel PCA comprised a kernel matrix K and kernel function $k(.)$ is a Mercer kernel *(Minh et al. 2006),* defined as $K_{ij} = k(x^{(i)}, x^{(j)})$, such that $k(.)$ return dot product of

5

feature space. Now mapping of eigenvalue of the kernel matrix, the eigen decomposition and respected eigenvectors are computed as,

$$\lambda^{\{i\}} e^{\{i\}} = K\ e^{\{i\}} \tag{3}$$

$$D(x) = \frac{1}{e^{\{i\}}} \sum_{t=1}^{T} e^{\{i\}} k(x^{(t)}, x) \tag{4}$$

where $\lambda^{\{i\}}$ is eigenvalues and $e^{\{i\}}$ is eigenvectors of K; T is number of training sample x to the principal component "i". Fischer et al. (2017) analyzed different methods like PCA, KPCA and Multi-Resolution PCA to Diaphragm tracking correlation coefficient between different versions of same sequence and agreed that Multi-Resolution PCA produce best result among most of parameters. Principal component analysis network (PCANet) is simple network architecture and one of the benchmark frameworks (Chan et al. 2015) for the unsupervised deep learning in recent time. However, (Shi et al. 2017) propose an encoding as C-RBH-PCANet which is improved PCANet to effectively integrate color pattern extraction and random binary hashing method for learning feature from color histopathological images.

3.3 Clustering

Clustering is unsupervised classification of unlabeled data (patterns, data item or feature vectors) into similar groups (clusters) (Fig. 3). Cluster analysis is explanatory in nature to find structure in data (Jain, 2008). Some model of clustering includes semi-supervised clustering, ensemble clustering, simultaneous feature selection and large-scale data clustering were emerging as hybrid-clustering. It involves analysis of multivariant data and applied in various scientific domains where clustering technique is utilized, such as machine learning, image analysis, bioinformatics, pattern recognition, computer vision and so on.
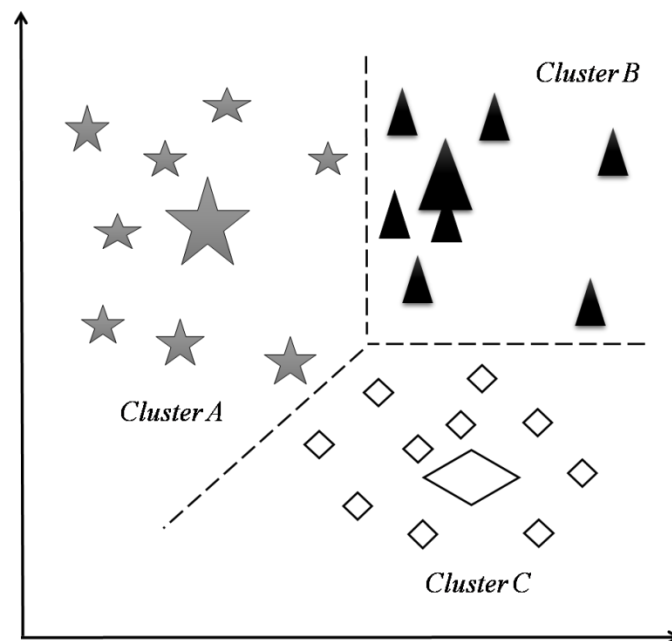
Fig. 3 An illustration of data clustering

Clustering algorithm can be broadly divided into two groups: *hierarchal clustering* and *partitional clustering*, as described below.

3.3.1 *Hierarchical clustering*

Hierarchical clustering algorithms find clusters recursively (using previously established cluster). These algorithms can be either in agglomerative mode (bottom-up) in which begin with each element as a separate cluster, merge the most similar pair of clusters successively into large clusters, or in divisive (top-down) mode which begin with all elements in one cluster, recursively divide into smaller clusters. A hierarchical clustering algorithm yields a dendrogram representing group of patterns and similarity level (Jain et al., 1999). A detailed discusion can be found in Jain et al. (1999).

3.3.2 *Partitional (k-means) clustering*

One of the most popular partitioning clustering algorithms is k-means (Steinhaus, 1956). In spite of several clustering algorithm published in over 50 years, K-means is still widely used. The most frequently used function in partitional clustering is squared error criterion, which applied to isolate and compact clusters. Let $X = \{x_i : i = 1, 2, 3, ... n\}$ be the set of n d-dimensional elements clustered into set of K clusters as $C = \{c_k : k = 1, 2, 3, ...K\}$. to find partitions the squared error between empirical mean of a cluster and elements in the cluster is minimized. Let μ_k be the mean of cluster (c_k), the squared error between mean and elements in cluster is defined as:

$$E(ck) = \sum_{xi \in ck} \|xi - \mu k\|^2 \tag{5}$$

The main objective of K-means is to minimize the sum of squared error for all K cluster (<mark>Drineas et al., 1999</mark>).

$$E(c) = \sum_{k=1}^{k} \sum_{xi \in ck} \|xi - \mu k\|^2 \tag{6}$$

Minimizing objective function is to be an NP-hard problem even for $k = 2$. Thus, k-means is a greedy algorithm and only be expected to converses to local minimum.

4. Unsupervised deep learning models

The goal of this section is to introduce a formal introduction and definition of unsupervised deep learning concepts, techniques and architectures that we found in medical analysis. The unsupervised deep learning models can be classified as shown in Fig. 4.
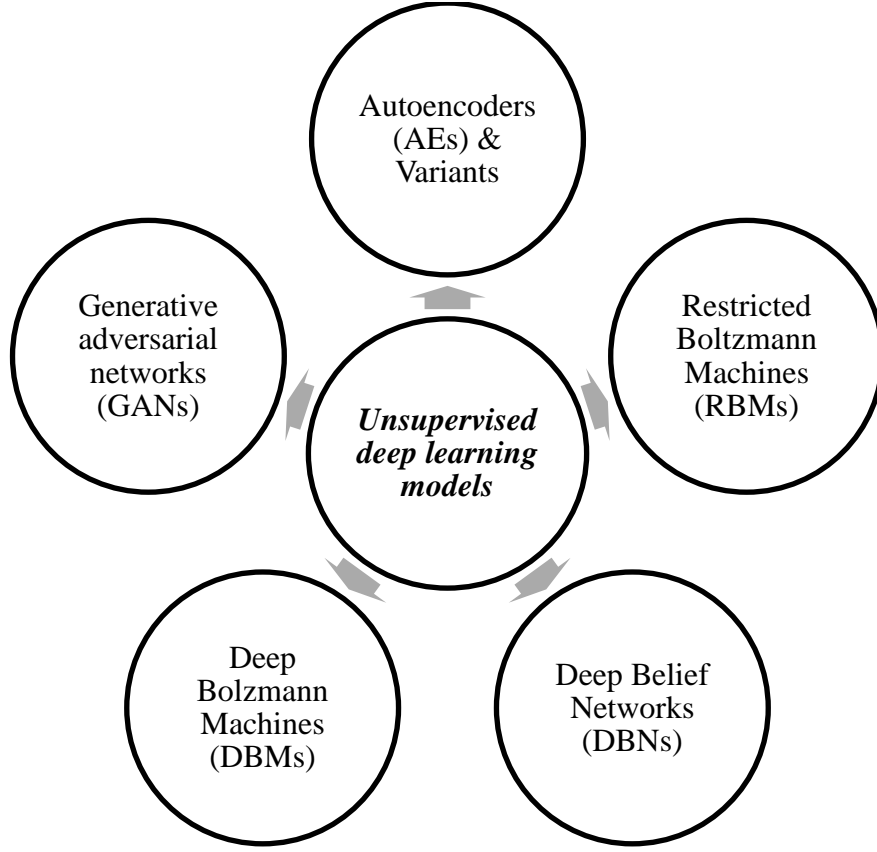
Fig. 4 Unsupervised deep learning models

4.1 Auto-encoders and its variants

In the literature, autoencoders and its several variants are reported and are being extensively applied in medical image analysis.

4.1.1 Autoencoders and Stacked autoencoder

Autoencoders (AEs) (Bourlard et al., 1988) are simple unsupervised learning model consist single-layer neural network that transforms the input into a latent or compressed representation by minimizing the reconstruction errors between input and output values of network. By constraining the dimension of latent representation (may be from different input) it is possible to discover relevant pattern from data. AEs framework defines a feature to extract function with specific parameters (Bengio et al., 2013). Basically, AEs are trained with specific function f_θ is called encoder and $h = f_\theta(x)$ is feature vector or representation from input x, another parameterized function g_θ called decoder, producing input space back from feature space. In short, basic AEs are trained to minimize reconstruction error in finding a value of parameter θ, given by,

$$T_{AE}(\theta) = \sum L\left(x, g_\theta [\![(f]\!]_\theta(x))\right) \qquad (7)$$

This minimization optionally followed by a non-linearity (most commonly used for encoder and decoder) as given by,

$$f_\theta(x) = S_f(W_x + b) \tag{8}$$

$$g_\theta(x) = S_g(W'_h + d) \tag{9}$$

where S_f and Sg are encoder and decoder activation function (normally, sigmoid or hyperbolic tangent or identity function) respectively, parameters of model $\theta = \{W, b, W', d\}$, where W and W' are encoder decoder weight matrices, and b and d are encoder and decoder bias vector, respectively. Moreover, regularization or sparsity constraints may be employed to enhance discovery process. If hidden layer had same as size of input and no any non-linearities is added, the model would learn simply identity function. Fig. 5(a) illustrates the basic structure of AE.

Stacked autoencoders (SAEs) are formed by placing AEs on top of each other also known as deep AEs. SAEs consists of multiple AEs stacked into multiple layers where output of each layer is wired to the inputs of the successive layers Fig. 5(b). To obtain good parameters, SAE uses greedy layer-wise training. The benefit of SAE is that it can enjoy the benefits of deep network which has grater expressive power. Furthermore, it usually captures a useful hierarchical grouping of the input (Shin et al., 2013).

4.1.2 Denoising autoencoder

Denoising autoencoder (DAEs) is another variant of auto-encoder. Denoising investigated as training criterion for learning to constitute better higher-level representation and extract useful features (Vincent et al. 2010). DAEs prevent the model from learning a trivial solution (Litjens G. et al., 2017) here model is trained to reconstruct a clean input from corrupted version from noise or another corruption. This is done by corrupting the initial input x into x̃ by using a stochastic function x̃ ~ q_D (x̃ | x). The corrupted input $\tilde{x}$ is mapped to a hidden representation y = $f_\theta(\tilde{x})$ = $s(W_{\tilde{x}} + b)$ and reconstruct $z = g_{\theta'}$ (y). A schematic representation is shown in Fig. 5(c). Parameter θ and θ' are initialized at random by stochastic gradient descent and trained to minimize average reconstruction error over a training set. The denoising autoencoders continue minimizes same reconstruction loss between clean *X* and reconstruction from *Y*. This continues maximizing a lower bond on the mutual information between input x and representation y, and difference are obtained by applying mapping f_θ to a corrupted input. Thus, such learning is more cleaver than the identity, and it extracts features useful for denoising.

Stack denoising autoencoder (SDAE) is a deep network utilizing the power of DAE (Bengio et al., 2007; Larochelle et al., 2009a) and RBMs in deep belief network (Hinton & Salakhutdinov, 2006; Hinton et al., 2006).

4.1.3   Sparse autoencoder

The limitation of autoencoders to have only small numbers of hidden units can be overcome by adding sparsity constraint, where large number of hidden units can be introduced usually more than one inputs. The aim of sparse autoencoder (SAE) is to make large number of neurons to have low average output so that neurons may be inactive most of the time.

Sparsity can be achive by introducing a loss function during training or manually zeroing few strongest hidden unit activations. A schematic representation of SAE is shown in Fig. 5(d).

If activation function of a hidden neurons is a_j, the average activation function of each hidden neurons j is given by

$$P_j = \frac{1}{m} \sum_{i=1}^{m} [a_j x_i]$$ (10)

The objective of sparsity constraints is to minimize P_j so that $P_j = P$, where P is sparsity constraint very close to 0 such as 0.05.

To enforce sparsity constraint, a penalty term is added to cost function which penalizes $\hat{P}_j$, deweighting significantly from P. The penalty term is the Kullback-Leibler (KL) divergence between Bernoulli random variables, can be calculated as (Ng, 2013; Makhzani & Frey, 2013),

$$Penalty\ term = \sum_{j=1}^{N_2} KL(P||\hat{P}_j)$$ (11)

where N_2 is number of neurons in the hidden layers, and index j is summing over the hidden units in the network.

$$KL(P||\hat{P}_j) = P\ log\frac{P}{\hat{P}_j} + (1-P)log\frac{1-P}{1-\hat{P}_j}$$ (12)

The property of penalty function is that $KL(P\ ||\ \hat{P}_j) = 0$, if $P_j = \hat{P}_j$, otherwise it increases gradually as $\hat{P}_j$ diverses for P.

The k-sparse autoencoder (Makhzani & Frey 2013) is a form of sparse AE where k neurons having highest activation function are chosen and rest are ignored. The advantage of k-sparse AE is that they allow better exploration on a data set in terms of percentage activation of the network. Advantage of SAE is the sparsity constraints which penalize the cost function and as a result degrees of freedom are reduced. Hence, it regularizes and maintains the complexity of the network by preventing overfiting.

4.1.4   Convolutional autoencoder

The most popular and widely used network model in deep unsupervised architecture is stacked AE. Stacked AE requires layer-wise pre-training. When layers go deeper during pre-training process, it may be time consuming and tedious because of stacked AE is built with fully connected layers. Li (2017) propose first trial to train convolutional directly an end-to-end manner without pretraining. Guo et al., (2017) suggested convolutional autoencoder (CAE) that is beneficial to learn feature for images and preserving local structure of data and avoid distortion of feature space. A general architecture of CAE is depicted in Fig. 5(e).
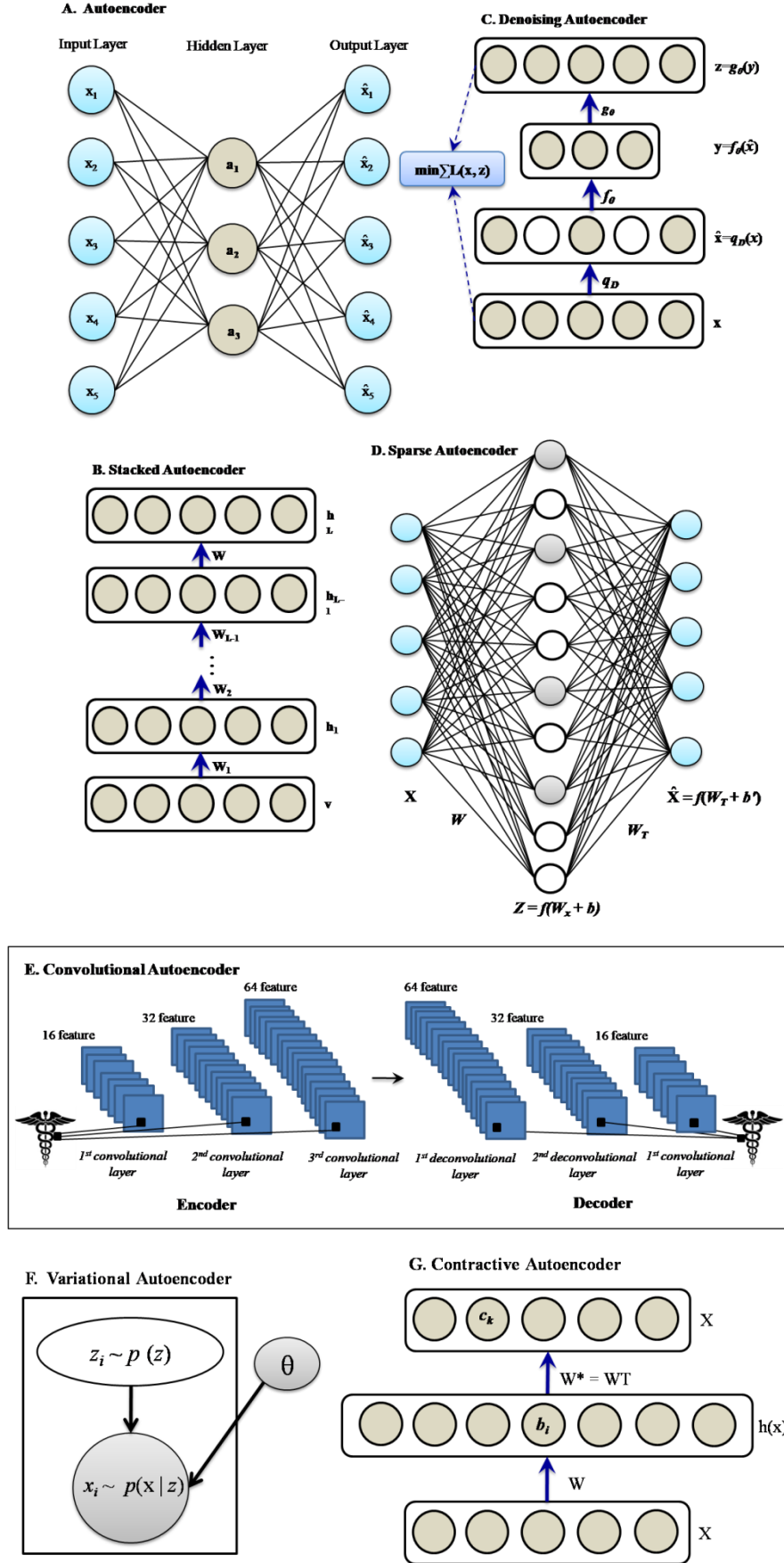
Fig. 5 (a)-(g) Diagrams showing networks of autoencoders and its different variants

4.1.5 Variational autoencoder

Another variant of autoencoder, called variational autoencoder (VAE), was introduced as a generative model (Kingma & Welling, 2013). A general architecture of VAE is given in Fig. 4(f). VAEs utilize the strategy of deriving a lower bond estimator from the directed graphical models with continuous distribution of latent variable. The generative parameter θ in the decoder (generative model) assist the learning process of the variational parameter, ϕ as encoder in the variational approximation model. VAEs use variational approach for latent representation learning in additional loss component training estimators called stochastic gradient variational Bayes (SGVB) and Auto-encoding variational Bayes (AEVB) (Kingma & Welling, 2013). Optimizes the parameter ϕ and θ for probabilistic encoder $q_\phi(z|x)$ which is an approximation to the generative model $p_\theta(x, z)$, where z is latent variable and x is continuous or discrete variable which aims to maximize the probability of each x in the training data set under the entire generative process. However, alternative configuration of generative latent variable modeling rises to give deep generative models (DGMs) instead of existing assumption of symmetric Gaussian posterior (Partaourides at el., 2017).

4.1.6 Contractive autoencoder

Rifai (2011) presented a novel approach for training deterministic autoencoder. Contractive autoencoder is addition of explicit regularizer in the objective function that enables the model to learn a function that has slight variations of input values. This additional regularizer corresponds to squared Forbenius norm of the Jacobian matrix of given activation with respect to the input. The contractive autoencoder is obtained with the regularization term in following equation yield final objective function,

$$f_{CAE}(\theta) = \sum_{x \epsilon D_n} \left(\text{L}(x, g(f(x)) + \lambda \left\| J_f(x) \right\|_F^2 \right) \tag{13}$$

The difference between contractive AE and DAE stated by (Vincent et al. (2010) as contractive AE explicitly encourage robustness of representation whereas DAE stressed on robustness of reconstruction this property make sense of contractive AE a better choice than DAEs to learn useful features extraction. Table1 presents a summary of autoencoders and its variants, and Table 2 presents its applications for medical images analysis.

Table 1. Summary of autoencoders and its variants

Types	Descriptions	References
Autoencoder	One of the simplest form which aims to learn a representation (encoding) for a set of data.	Ballard (1987); Bourlard & Kamp (1988); Hinton & Zemel1(994)
Stacking autoencoder	An autoencoder having multiple layers where outputs of each layers is wired to the inputs of the successive layer.	Zabalza et al. (2016)
Sparse autoencoder	Encourages hidden units to be zero or near zero	Goodfellow et al. (2009)
Denosing autoencoder	Capable to predict true inputs from noisy data	LeCun (1987); Gallinari et al. (1987); Vincet et al. (2008)

Convolutional autoencoder	Learn feature and preserving local structure of data and avoid distortion of feature space		Guo et al. (2017)	
Variational autoencoder	A generative model utilizing strategy of deriving a lower bond estimator from directed graphical models with continuous distribution of latent variable.		Kingma & Welling (2013)	
Contractive autoencoder	Forces encoder to take small derivatives		Rifai et al. (2011)	

Table 2 Applications of autoencoders and its variants for medical image analysis. Abbreviations - H&E: hematoxylin and eosin staining, AD:Alzheimer's disease, MCI: Mild cognitive impairment, fMRI: Functional magnetic resonance imaging, sMRI: Structural magnetic resonance imaging, rs-fMRI: Resting-state fMRI, DBN :Deep belief network, RBM : Restricted Boltzmann machine.

Method	Task	Image type	Remarks	References
SAE	AD/MCI classification	MRI	stacked auto encoders with supervised fine tuning	Suk and Shen (2013)
SAE	AD/MCI/HC classification	MRI & PET	SAE for latent feature extraction on a large set of hand-crafted features from MRI and PET	Suk et al. (2015)
SAE	AD/MCI/HC classification	MRI	SAE used to pre-train 3D CNN	Payan & Montana (2015)
SAE	MCI/HC classification	fMRI	SAE used for feature extraction, HMM as a generative model on top	Suk et al. (2016)
SAE	Hippocampus segmentation	MRI	SAE for representation learning used for target/atlas patch similarity measurement	Guo et al. (2014)
SAE	Visual pathway segmentation	MRI	Learning appearance features from SAE for steering the shape model for segmentation	Mansoor et al. (2016)
SAE	Denoising DCE-MRI	MRI	Using an ensemble of denoising SAE (pre-trained with RBMs)	Benou et al. (2016)
SSAE	Nucleus detection	Digital pathology image	Nuclei detection on breast cancer histopathology images	Xu et al. (2016b)
SAE	Stain normalization	Digital pathology image	SAE is applied for classifying tissue and subsequent histogram Matching	Janowczyk et al. (2017)
SAE	Density classification	Mammography	Unsupervised CNN feature learning with SAE for breast density classification	Kallenberg et al. (2016)
SAE	Lesion classification	MRI	Learns features from multiple modalities, hierarchical random forest for classification	Zhu et al. (2017)
SAE	Heart, kidney, liver location detection	MRI	SAE to learn temporal/spatial features on 2D + time DCE-MRI	Shin et al. (2013)
SAE	Cell segmentation	Digital pathology image	Learning spatial relationships	Hatipoglu, N. 2017
SAE	Segmentation of the right ventricle from cardiac	MRI	SAE applied to obtain an initial right ventricle segmentation.	Avendi, M. 2017
SDAE	Cell segmentation	Digital pathology image	The SDAE trained with data and their structured labelsfor cell segmentation	Su. H. at el 2018
SDAE	Lesion classification	Ultrasound and CT	Detection of breast lesions in Ultrasound and pulmonary nodules in CT	Cheng et al. (2016a)
SSAE	AD	MRI	Stacked Sparse AE for Early diagnosis of Alzheimer disease from brain MRIs	Liu et al. (2014)
SDAE	AD	Ultrasound	Stacked Denoising AE for Diagnosis of breast nodules and lesions	Cheng et al. (2016)
SDAE	Patientclinical events	Patient clinical history	SDAE for Deep Patient: an unsupervised representation of patients to predict future clinical events	Miotto et al. (2016)

Method	Task	Image type	Remarks	References
SDAE		CT/MRI	Utilize multi-modal SDAE to pre-train the deep neural network.	Cheng et al. (2018)
DCAE	Modeling task fMRI	tfMRI	Deep Convolutional AE to model tfMRI.	Haung et al. (2018)
CAE	AD/MCI/HC classification	fMRI	CAE used to pre-train 3D CNN	Hosseini-Asl et al. (2016)
CAE	Nucleus detection	Digital pathology image	CAE detects and encodes nuclei in image patches	Hou et al. (2019)

4.2. Restricted Boltzmann Machines

Restricted Boltzmann Machines (RBMs) are a variant of Markov Random Field (MRF), constitute of single layer undirected graphical model with an input layer or visible layer x = $(x_1, x_2...... x_N)$ and a hidden layer h = $\{h_1, h_2, ... h_M\}$. The connection between nodes/unit are bidirectional, so each given input vector x can take the latent feature representation h and vice versa. As such an RBM is a generative model that learn probability distribution over the input space and generate new data point (Yoo, et al. 2014). Illustration of a typical RBM architecture is shown in Fig. 6(a). RBMs are restricted variant of Boltzmann machines, with the restriction that neurons must form an arrangement of bipartite graph. Due to this, pairs of nodes from each of the groups (i.e. visible and hidden units) have a symmetric connection between them, and there are no connections between nodes within a group. This restriction makes RBM more efficient training algorithm than general case of Boltzmann machine. Hinton et al. (2010) proposed a practical guide to train RBMs.

RBMs have been utilized for various aspects of medical image analysis such as detaction of variants in Alzheimer disease (Brosch, et al. 2013), image segmentation (Yoo et al. 2014), dimensonality reduction (Cheng J-Z, et al. 2016), feature learning (Pereira et al. 2018) and so on. A brief account for the application of RMBs in medical image analysis is shown in Table 3.

**Table 3.** Applications of RBM for medical image analysis

Method	Task	Image type	Remarks	References
RBM	AD	MRI	Manifold of brain MRIs to detect modes of variations in Alzheimer disease	Brosch et al. (2013)
RBM	sclerosis lesions	3DMRI	Segmentation of multiple sclerosis lesions in multi-channel 3D MRIs	Yoo et al. (2014)
RBM	AD/MCI/HC classification	MRI & PET	Deep Boltzmann Machines on MRI and PET modalities	Suk et al. (2014)
RBM	Mass detection in breast cancer	Mammography	RBM based method for oversampling and semi-supervised learning to solve classification of imbalanced data with a few labeled samples	Cao et al. (2015)
RBM	fMRI blind source separation	fMRI	RBM for both internal and functional interaction-induced latent sources detection	Huang et al. (2016)
RBM*	Vertebrae localization	CT, MRI	RBM determines position, orientation and label of vertebrae	Cai et al. (2016b)
RBM	Benign/Malignant classification	Ultrasound	Classification of benign vs. malignant with shear wave elastography	Zhang et al. (2016c)
RBM	Tongue contour	Ultrasound	Analysis of tongue motion	Jaumard-Hakoun et al.

| CRBM	Lung tissue classification and airway detection	CT	Convolutional RBM combines a generative and a discriminative learning to learn filters for training data and classification as well.	van Tulder& de Bruijne (2016)
	extraction		during speech, combines auto-encoders with RBMs	(2016)
RBM	Cardiac arrhythmias classification	ECG	Achieves average recognition accuracies of ventricular ectopic beats (93.63%) and supraventricular ectopic beats (95.57%)	Mathews et al. (2018)
RBM	Brain lesion segmentationC	MRI	RBM is used for feature learning, and a Random Forest as a classifier.	Pereira et al. (2018)
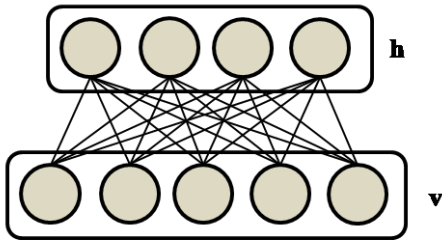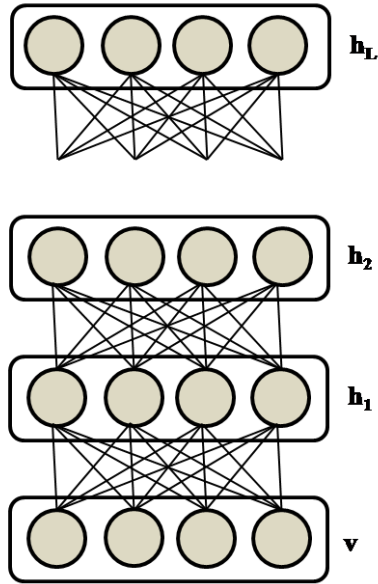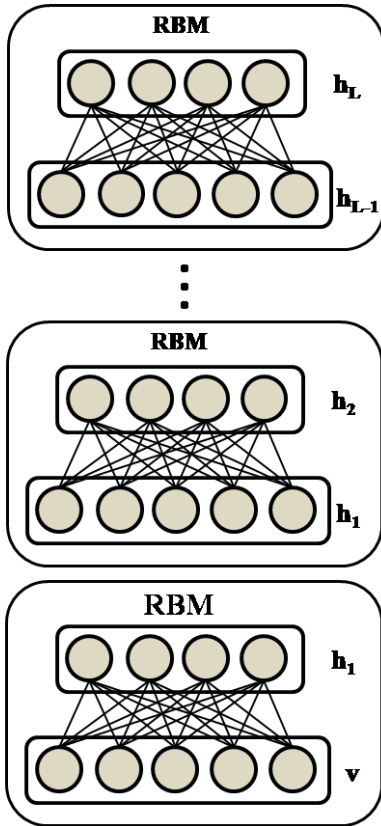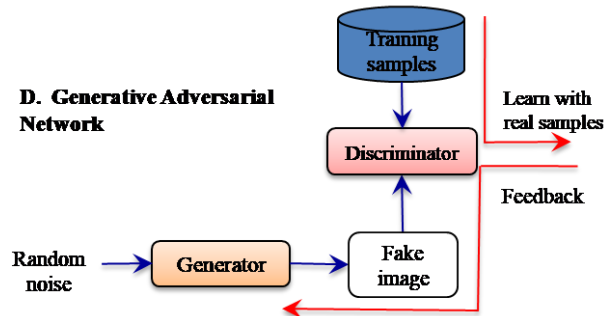
Fig. 6 (a)-(d) Diagrams showing various unsupervised network models

4.3.   Deep Belief Networks

Deep Belief Networks (DBN) is a kind of neural network proposed by Bengio (2009). It is a greedy layer wise unsupervised learning algorithm with several layers of hidden variables (Hinton et al., 2016). Layer-wise unsupervised training (Bengio 2007) help the optimization and weight initialization for better generalization. In fact, DBN is a hybrid single probabilistic generative model which is a typical RBM. In order to construct a deep architecture like SAEs where AEs layers are replaced by RBMs, DBN has one lowest visible layer **v**, representing state of input data vector and a series of hidden layers h^1, h^2, h^3, . . . h^L. When multiple RBMs stacked hierarchically, the top two layers form an undirected generative model and lower layers formed directed generative model. Fig. 6(b), illustrates the structure of DBN. The following function in DBN represent joint distribution of visible unit v, hidden layers h^1 (l = 1, 2…. L):

$$P(V, \; h^1, h^2 \dots h^L) = \left(\prod_{l=0}^{L-2} P(h^l | h^{l+1}) P\left(h^{L-1}, h^{(L)}\right) \right) \tag{14}$$

Hinton at el. (2006a) applied layer-wise training procedure, where lower layers learns low-level features and subsequently higher layers learns high-level features (Hinton at el. 1995). DBN are used to extract features from fMRI images (Plis et al., 2014), temporal ultrasound (Azizi et al. 2016), classify Autism spectrum disorders (Akhavan et al. 2018), and so on. Some of the applications of DBNs are presented in Table 4.

Table 4. Applications of DBNs for medical image analysis

Method	Task	Image type	Remarks	References
DBN	AD/HCclassification	MRI	deep belief networks with convolutional RBMs for manifold learning	Brosch and Tam (2013)
DBN, Convolutional RBM	Manifold Learning	MRI	DBN with conv. RBM layers for modeling the variability in brain morphology and lesion distribution in MS	Brosch et al. (2014)
DBN		MRI	Deep belief networks evaluated on brain network estimation, Schizophrenia and Huntington's disease classification	Plis et al. (2014)
DBN	AD/MCI/HC classification	MRI	an ensemble of DBNs, with their votes fused using an SVM classifier	Ortiz et al. (2016)
DBN	Left ventricle segmentation	Ultrasound	DBN embedded in system using landmarks, non-rigid registration, and patches steers multi-atlas segmentation process	Carneiro et al. (2012); Carneiro and Nascimento (2013); Nascimento and Carneiro (2016)
DBN	Schizophrenia/NH classification	MRI	DBN pre-training followed by supervised fine-tuning	Pinaya et al. (2016)
DBN	Lesion classification	Ultrasound	DBN learns features from temporal ultrasound to classify prostate lesions benign/malignant	Azizi et al. (2016)
DBN	Left ventricle segmentation	MRI	DBN is used to initialize a level set framework	Ngo et al. (2017)
DBN	Cardiac arrhythmias classification	ECG	Achieves average recognition accuracies of ventricular ectopic beats (93.63%) and supraventricular ectopic beats (95.57%)	Mathews et al. (2018)

| DBN | Autism spectrum disorders classification | rs-fMRI, sMRI | Classifies Autism spectrum disorders in young children using combination of rs-fMRI and sMRI data | Akhavan et al. (2018) |

4.4. Deep Boltzmann Machine

Deep Boltzmann machine (DBM) is robust deep learning model proposed by Salakhutdinov et al. (2009) and Salakhutdinov et al. (2012). They stacked multiple RBMs in hierarchal manner to handle ambiguous input robustly. Fig. 6(c) represents the architecture of DBM as a composite model of RBMs which clearly shows how DBM differ from DBN. Unlike DBNs, all layers in DBMs form undirected generative model combining information from both lower and upper layers which improves the representation power of DBMs. Training of layer-wise greedy algorithm for DBM (Salakhutdinov R. et al. 2015; Goodfellow et al 2013b) is calculated by modifying in procedure of DBN.

Dinggang et al. (2017) and Salakhutdinov et al. (2015) presented a three-layer DBM with given values in the neighboring layer(s) and the probability of visible or hidden units computed by logistic sigmoid function to learn the parameters $\Theta = \{W^1, W^2\}$. The derivative of the log likelihood of the observation with respect to the model parameter is computed as,

$$\frac{\partial}{\partial w(l)} \ln P(V:\theta) = E_{data}[h^{l-1}(h^l)^T] \sim E_{model}[h^{l-1}(h^l)^T] \tag{15}$$

Where $E_{data}[.]$ denote data-dependent obtained from visible units and $E_{model}[.]$ denote data-independence obtained from the model. Some of the applications of DBMs are shown in Table 5.

Table 5. Applications of DBMs for medical image analysis

Method	Task	Image type	Remarks	References
DBM	Heart motion tracking	MRI	Using three-layered Deep Boltzmann Machine to guide frame-by-frame heart segmentation during radiation therapy of cancer patient on cine MRI images.	Wu, et al., (2018)
DBN	AD/HCclassification	MRI	deep belief networks with convolutional RBMs for manifold learning	Brosch and Tam (2013)
RBM	Breast-image classification	MRI	Restricted Boltzmann machine with backpropagation have been used for hhistopathological breast-image classification	Nahid et al., (2018)
DBM	Medical image retrieval	Multi digital image	deep Boltzmann machine-based multimodal learning model to learn the joint density model	Cao, et al., (2014)

4.5. Generative Adversarial Network (GAN)

Generative Adversarial Network (GAN) (Goodfellow, et al. 2014) is one of recent promising technique for building flexible deep generative unsupervised architecture. Goodfellow et al. (2014) proposed two models generative model G and Discriminative model D, where G capture data distribution (p_g) over real data t, and D estimates the probability of a sample coming from training data (m) not from *G*. In every iteration backpropagation generator and discriminator competing with each other. The training procedure the probability of D is maximized. This framework functions like a minimax two-player game. The value function V(G, D) establishes following two-player minimax game is given by,

$$\underset{G}{\min}\,\underset{D}{\max}\ V(G,D) = E_{t \sim p_{data}}[logD(t)] + E_{m \sim p_m(m)}\left[\log\left(1 - D\big(G(m)\big)\right)\right] \qquad (16)$$

Where $D(t)$ represents the probability of **t** from data m and p_{data} is distribution of real-world data. This model seems to be stable and improved as $p_g = p_{data}$. A typical architecture of GAN is depicted in Fig. 6(d). In fact, these two adversaries, *Generator* and *Discriminator*, continuously battle during the processing of training. GAN have been applied to generate samples of photorealistic images to visualize new designs. Some of the applications of GAN for medical image analysis is presented in Table 6.

Table 6. Applications of GAN for medical image analysis

Method	Task	Image type	Remarks	References
GAN	Synthesis of retinal images	Retinal images	MI-GAN generates precise segmented image for the application of supervised analysis of medical images.	Iqbal and Ali (2018)
GAN	chest X-Ray	X-Ray	GAN use to to produce photorealistic images that retain pathological quality	Canas, et al, (2018)
Dual GAN-FCN	Segmentation of regions of interest (ROIs)	---	Improve the GAN using dual-path adversarial learning for Fully Convolutional Network (FCN)-Based Medical Image Segmentation	Bi, et al., (2018)
GAN	Simulation of B-mode ultrasound images	Ultrasound	conditional generative adversarial networks used to simulate ultrasound images at given 3D spatial locations	Hu et al., (2018)
GAN	Treatment of lymphomas and lung cancer	PET	multi-channel generative adversarial networks used to synthesize PET data	Bi, et al., (2017)

5. Discussion, opportunities and challenges

Medical imaging and diagnostic techniques are one of the most widely used for early detection, diagnosis and treatment of complex diseases. After significant advancement in machine learning and deep learning (both supervised and unsupervised), there is paradigm shift from manual interpretation of medical images by human experts such as radiologists and

physicians to an automated analysis and interpretation, called computer-assisted diagnosis (CAD). As unsupervised learning algorithms can derive insights directly from data, use them for data-driven decisions making, and are more robust, hence they can be utilized as the holy grail of learning and classification problems. Furthermore, these models are also utilized for other important tasks including compression, dimensionality reduction, denoising, super resolution and some degree of decision making.

Unsupervised learning and CAD, both being in its infancy, researchers and practitioners have much opportunity in this area. Some of them are: *(i)* Allow us to perform exploratory analysis of data *(ii)* Allow to be used as preprocessing for supervise algorithm, when it is used to generate a new representation of data which ensure learning accuracy and reduces memory time overheads. (*iii*) Recent development of cloud computing, GPU-based computing, parallel computing and its cheaper cost allow big data processing, image analysis and execute complex deep learning algorithm very easily.

Some of the challenges and research directions are:

(i) Difficult to evaluate whether algorithm has learned anything useful: Due to lack of label in unsupervised learning, it is nearly impossible to quantify its accuracy. For instance, how can we access whether K-means algorithm found the right clusters? In this direction, there is need to develop algorithms which can give an objective performance measure in unsupervised learning.

(ii) Difficult to select right algorithm and hardware: Selection of right algorithm for a particular type of medical image analysis is not a trivial task because performances of the algorithm are highly dependent on the types of data. Similarly, hardware requirement also varies from problem to problem.

(iii) "Will unsupervised learning work for me?" It is mostly asked question but its answer totally depends on the problem at hand. In image segmentation problem, clustering algorithm will only work if images do fit into naturals groups.

(iv) Not a common choice for medical image analysis: Unsupervised learning is not a common choice for medical image analysis. However, from literature it is revealed that these (autoencoders and its variants, DBN, RBM, etc.) are mostly used to learn hierarchy level of features for classification tasks. It is expected that unsupervised learning will play pivotal role in solving complex medical imaging problems which are not only scalable to large amount of unlabeled data but also suitable for performing unsupervised and supervised learning tasks simultaneously (Yi, et al., 2018).

(v) Development of patient-specific anatomical and organ model: Anatomical skeletons play crucial role in understanding diseases and pathology. Patient-specific anatomical model are frequently used for surgery and interventions. They helps to plan procedure, perform measurement for device surging, and predict the outcome of post-surgery complexities. Hence, algorithm needs to be developed to construct patient-specific anatomical and organ model from medical images.

(vi) Heterogeneous images data: In the last two to three decades, more emphasis was given to well defined medical images analysis applications where developed algorithms were validates on well-defined types of images with well-defined acquisition protocol. The algorithms are required which can work on more heterogeneous data.

(vii) Semantic segmentation of images: Semantic segmentation is task of complete scene understanding, leading to knowledge inference from imagery. Scene understanding is a core of computer vision problems which has several applications, including human-computer interaction, self-driving vehicles, virtual reality, and medical image analysis. The semantic segment of medical images with acceptable accuracy is still challenging.

(viii) Medical video transmission: Enabling 3D video in recently adopted telemedicine and U-healthcare applications result in more natural viewing conditions and better diagnosis. Also, remote surgery can be benefited from 3D video because of additional dimensions of depth. However, it is crucial to transmit data-hungry 3D medical video stream in real-time through limited bandwidth channels. Hence, efficient encoding and decoding techniques for 3D video data transmission is required.

(ix) Need extensive interorganizational collaborations: Interprofessional and interorganizational collaboration are important for better functioning of the healthcare system, eliminating some of the pitfalls such as limited resources, lack of expertise, aging populations, and combat chronic diseases (Karam et al., 2017). Medical image based CAD needs extensive interorganizational collaborations among doctors, radiologists, medical image analysts, and computational data analysts.

(x) Need to capitalize big medical imaging market: According to IHS Markit report (https://technology.ihs.com.), medical imaging market has total global revenue of $21.2 billion in 2016, which is forecasted to touch $24.0 billion by 2020. According to WHO, global population will rise from 12% to 22% from 2015 to 2050. Population aging lead to increased rate of chronic diseases globally, and hence there is a need to capitalize a big medical imaging market worldwide.

(xi) Black-box and its acceptance by health professionals: Machine learning algorithms are boon which solves the problems earlier thought to be unsolvable, however, it suffers from being "black-box", i.e., how output arrives from the model is very complicated to interpret. Particularly, deep learning models are almost uninterpretable and but still being used for complex medical image analysis. Hence, its acceptance by health professionals is still questionable.

(xii) Will technology replace radiologists? For the processing of medical images, deep learning algorithms help select and extract important features and construct new ones, leading to new representation of images, not seen before. For image interpretation side, deep learning helps identify, classify, quantify disease patterns, allow measure predictive targets, and make predictive models, and so on. So, will technology "replace radiologists", or migrate to "virtual radiologist assistant" in near future? Hence, following slogan is quite relevant in this context: "*Embrace it, it will make you stronger; reject it, it may make you irrelevant*".

In a nutshell, unsupervised learning is very much open topic where researchers can make contributions by developing a new unsupervised method to train how network (e.g. Solve a puzzle, generate image patterns, image patch comparison, etc.) and re-thinking of creating great unsupervised feature representation such as (e.g. What is object and what is background), nearly analogous to human visual system.

6. Conclusion

Medical imaging is one of the important techniques for early detection, diagnosis and treatment of complex diseases. Interpretation of medical images is usually performed by human experts such as radiologists and physicians. After the success of machine learning techniques, including deep learning, availability of cheap computing infrastructure through cloud computing, there has been a paradigm shift in the field of computer-assisted diagnosis (CAD). Both supervised and unsupervised machine learning approaches are widely applied in medical image analysis, each of them with their own pros and cons. Due to the fact that human supervisions are not always available or inadequate or biased, therefore, unsupervised learning algorithms, including its deep architecture, give a big hope with lots of advantages.

Unsupervised learning algorithms derive insights directly from data, and use them for data-driven decisions making. Unsupervised models are more robust and they can be utilized as the holy grail of learning and classification problems. These models are also used for other tasks including compression, dimensionality reduction, denoising, super resolution and some degree of decision making. Therefore, it is better to construct a model without knowing what tasks will be at hand and we would use representation (or model) for. In a nutshell, we can think of unsupervised learning as preparation (preprocessing) step for supervised learning tasks, where unsupervised learning of representation may allow better generalization of a classifier.

Conflict of Interest Statement

Authors declare that there is no any conflict of interest in the publication of this manuscript.

References

Aghdam, M. A., Sharifi, A., & Pedram, M. M. (2018). Combination of rs-fMRI and sMRI Data to Discriminate Autism Spectrum Disorders in Young Children Using Deep Belief Network. *Journal of digital imaging*, 1-9. https://doi.org/10.1007/s10278-018-0093-8

Anil K. Jain. 2010. Data clustering: 50 years beyond K-means. Pattern Recogn. Lett. 31, 8 (June 2010), 651-666. https://doi.org/10.1016/j.patrec.2009.09.011

Arora, S. and Zhang, Y. Do gans actually learn the distribution? an empirical study. CoRR, abs/1706.08224, 2017. URL http://arxiv.org/abs/1706.08224.

Avendi, M. R., Kheradvar, A., &Jafarkhani, H. (2017). Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach. Magnetic Resonance in Medicine, 78(6), 2439–2448. doi:10.1002/mrm.26631

Azizi, S., Imani, F., Ghavidel, S., Tahmasebi, A., Kwak, J.T., Xu, S., Turkbey, B., Choyke, P., Pinto, P., Wood, B., Mousavi, P., Abolmaesumi, P., 2016. Detection of prostate cancer using temporal sequences of ultrasound data: a large clinical feasibility study. Int. J. Comput. Assist. Radiol. Surg. 11 (6), 947–956. doi: 10. 1007/s11548- 016- 1395- 2.

Ballard, D. H. (1987, July). Modular Learning in Neural Networks. In *AAAI* (pp. 279-284).

Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. IEEE Trans Pattern Anal Mach Intell2013;35:1798–828

Bengio, Y. ,Lamblin, P. , Popovici, D. , Larochelle, H. , 2007. Greedy layer-wise training of deep networks. In: Proceedings of the Advances in Neural Information Processing Systems, pp. 153–160.

Benou, A. , Veksler, R. , Friedman, A. , Raviv, T.R. , 2016. De-noising of contrast-en- hanced MRI sequences by an ensemble of expert deep neural networks. In: Proceedings of the Deep Learning in Medical Image Analysis (DLMIA). In: Lecture Notes in Computer Science, 10 0 08, pp. 95–110 .

Bi, L., Feng, D., Kim, J. Dual-Path Adversarial Learning for Fully Convolutional Network (FCN)-Based Medical Image Segmentation (2018) Visual Computer, 34 (6-8), pp. 1043-1052. DOI: 10.1007/s00371-018-1519-5

Bi, L., Kim, J., Kumar, A., Feng, D., Fulham, M. Synthesis of positron emission tomography (PET) images via multi-channel generative adversarial networks (GANs) (2017) Lecture Notes in Computer Science, 10555 LNCS, pp. 43-51. DOI: 10.1007/978-3-319-67564-0_5

Bishop CM. Pattern Recognition and Machine Learning (Information Science and Statistics), 1st edn. 2006. corr. 2nd printing edn. Springer, New York, 2007.

Bourlard H, Kamp Y. 1988. Auto-association by multilayer perceptrons and singular value decomposition. *Biol. Cybern.* 59:291–94

Brosch, T., Tam, R., 2013. Manifold learning of brain MRIs by deep learning. In: Proceedings of the Medical Image Computing and Computer-Assisted Interven- tion. In: Lecture Notes in Computer Science, 8150, pp. 633–640. doi: 10.1007/ 978- 3- 642- 40763- 5 _ 78 .

Bulletin of acad. polon. sci., IV(C1. III), 801–804.

Canas, K., Liu, X., Ubiera, B., Liu, Y. Scalable biomedical image synthesis with GAN (2018) ACM International Conference Proceeding Series, art. no. a95. DOI: 10.1145/3219104.3229261

Cao, P., Liu, X., Bao, H., Yang, J., & Zhao, D. (2015). Restricted Boltzmann machines based oversampling and semi-supervised learning for false positive reduction in breast CAD. *Biomedical materials and engineering*, *26*(s1), S1541-S1547.

Cao, Y., Steffey, S., He, J., Xiao, D., Tao, C., Chen, P., Müller, H. Medical image retrieval: A multimodal approach (2014) Cancer Informatics, 2014, pp. 125-136. DOI: 10.4137/CIN.S14053

Carneiro, G., Nascimento, J.C., 2013. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. IEEE Trans. Pattern Anal. Mach. Intell. 35, 2592–2607. doi: 10.1109/TPAMI. 2013.96 .

Carneiro, G., Nascimento, J.C., Freitas, A., 2012. The segmentation of the left ven- tricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. IEEE Trans. Image Process 968–982. doi: 10. 1109/TIP.2011.2169273 .

Cheng J-Z, Ni D, Chou Y-H, et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. Sci Rep 2016;6:24454

Denton, E. L., Chintala, S., Szlam, A., and Fergus, R. Deep generative image models using a laplacian pyramid of adversarial networks. CoRR, abs/1506.05751, 2015. URL http://arxiv.org/abs/1506.05751.

Diederik P. Kingma and Max Welling. 2013. Auto-encoding variational bayes. *CoRR*abs/1312.6114 (2013). Retrieved from http://arxiv.org/abs/1312.6114

Fischer, P. Pohl, T. Faranesh, A., Maier, A. and Hornegger, J. (2017). Unsupervised Learning for Robust Respiratory Signal Estimation From X-Ray Fluoroscopy, *IEEE Transactions on Medical Imaging*, 36(4), 865-877. doi: 10.1109/TMI.2016.2609888.

Goodfellow, I. J., Mirza, M., Courville, A., and Bengio, Y. (2013b). Multi-prediction deep Boltzmann machines. In NIPS'2013.

Guo, X., Liu, X., Zhu, E., & Yin, J. (2017, November). Deep clustering with convolutional autoencoders. In *International Conference on Neural Information Processing* (pp. 373-382). Springer, Cham.

Hatipoglu, N., &Bilgin, G. (2017). Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships. Medical & Biological Engineering & Computing, 55(10), 1829–1848. doi:10.1007/s11517-017-1630-1

Hinton G, Dayan P, Frey B,NealR. 1995. The "wake–sleep" algorithm for unsupervised neural networks.Science 268:1158–61.

Hinton GE, Salakhutdinov RR. 2006. Reducing the dimensionality of data with neural networks. Science 313:504–7.

Hinton, G. , 2010. A practical guide to training restricted boltzmann machines. Momentum 9 (1), 926

Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A fast learning algorithm for deep belief nets. Neural Comput. 18, 1527–1554. doi: 10.1162/neco.2006.18.7.1527.

Hosseini-Asl, E., Gimel'farb, G., El-Baz, A., 2016. Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network. arxiv: 1607.00556 .

Hou, L., Nguyen, V., Kanevsky, A. B., Samaras, D., Kurc, T. M., Zhao, T., ... & Saltz, J. H. (2019). Sparse Autoencoder for Unsupervised Nucleus Detection and Representation in Histopathology Images. Pattern Recognition, 86: 188-200. https://doi.org/10.1016/j.patcog.2018.09.007

Hu, Y., Gibson, E., Lee, L.-L., Xie, W., Barratt, D.C., Vercauteren, T., Noble, J.A. Freehand ultrasound image simulation with spatially-conditioned generative adversarial networks (2017) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10555 LNCS, pp. 105-115. DOI: 10.1007/978-3-319-67564-0_11

Huang, H., Hu, X., Han, J., Lv, J., Liu, N., Guo, L., Liu, T., 2016. Latent source mining in FMRI data via deep neural network. In: Proceedings of the IEEE International Symposium on Biomedical Imaging, pp. 638–641. doi: 10.1109/ISBI.2016.74 9334 8 .

Huang, H., Hu, X., Zhao, Y., Makkie, M., Dong, Q., Zhao, S., ... & Liu, T. (2018). Modeling task fMRI data via deep convolutional autoencoder. IEEE transactions on medical imaging, 37(7), 1551-1561.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, SherjilOzair, Aaron Courville, and YoshuaBengio. 2014. Generative adversarial nets. In Advances in Neural Information Processing Systems. Curran Associates, 2672–2680.

Iqbal, T., Ali, H. Generative Adversarial Network for Medical Images (MI-GAN) (2018) Journal of Medical Systems, 42 (11), art. no. 231, DOI: 10.1007/s10916-018-1072-9.

Jabeen, A., Ahmad, N., & Raza, K. (2018). Machine Learning-Based State-of-the-Art Methods for the Classification of RNA-Seq Data. In *Classification in BioApps* (pp. 133-172). Springer, Cham.

J. Shi, J. Wu, Y. Li, Q. Zhang and S. Ying, "Histopathological Image Classification WithColor Pattern Random Binary Hashing-Based PCANet and Matrix-Form Classifier," in *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 5, pp. 1327-1337, Sept. 2017. doi: 10.1109/JBHI.2016.2602823

Janowczyk, A. , Basavanhally, A. , Madabhushi, A. , 2017. Stain normalization using sparse autoencoders (STANOSA): application to digital pathology.. Comput. Med. Imaging Graph 57, 50–61

Jaumard-Hakoun, A., Xu, K., Roussel-Ragot, P., Dreyfus, G., Denby, B., 2016. Tongue contour extraction from ultrasound images based on deep neural network. arxiv: 1605.05912 .

Kallenberg, M., Petersen, K., Nielsen, M., Ng, A., Diao, P., Igel, C., Vachon, C., Hol- land, K., Karssemeijer, N., Lillholm, M., 2016. Unsupervised deep learning ap- plied to breast density segmentation and mammographic risk scoring. IEEE Trans. Med. Imaging 35, 1322–1331. doi: 10.1109/TMI.2016.2532122

Karam, M., Brault, I., Van Durme, T., & Macq, J. (2017). Comparing interprofessional and interorganizational collaboration in healthcare: A systematic review of the qualitative research. *International journal of nursing studies*.

Karras, T., Aila, T., Laine, S., and Lehtinen, J. Progressive Growing of GANs for Improved Quality, Stability, and Variation. ArXiv e-prints, October 2017.

Li, F., Qiao, H., Zhang, B., Xi, X.: Discriminatively boosted image clustering with fully convolutional auto-encoders. arXiv preprint arXiv:1703.07980 (2017).

Liu S, Liu S, Cai W, et al. Early diagnosis of Alzheimer's disease with deep learning. In: International Symposium on Biomedical Imaging, Beijing, China 2014, 1015–18.

Makhzani, A. & Frey, B. (2013). k-Sparse Autoencoders. arxiv: preprint : 1312.5663.

Mansoor, A., Cerrolaza, J., Idrees, R., Biggs, E., Alsharid, M., Avery, R., Linguraru, M.G., 2016. Deep learning guided partitioned shape model for anterior visual path- way segmentation. IEEE Trans. Med. Imaging 35 (8), 1856–1865. doi: 10.1109/ TMI.2016.2535222 .

Mathews, S. M., Kambhamettu, C., & Barner, K. E. (2018). A novel application of deep learning for single-lead ECG classification. *Computers in biology and medicine*, 99:53-62. doi: 10.1016/j.compbiomed.2018.05.013

Metz, L., Poole, B., Pfau, D., and Sohl-Dickstein, J. Unrolled generative adversarial networks. CoRR, abs/1611.02163, 2016. URL http://arxiv.org/ abs/1611.02163.

Minh H.Q., Niyogi P., Yao Y. (2006) Mercer's Theorem, Feature Maps, and Smoothing. In: Lugosi G., Simon H.U. (eds) Learning Theory. COLT 2006. Lecture Notes in Computer Science, vol 4005. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11776420_14

Miotto R, Li L, Kidd BA, et al. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. Sci Rep 2016;6:26094.

Miyato, T. and Koyama, M. cGANs with projection discriminator. In International Conference on Learning Representations, 2018. URL https:// openreview.net/forum?id=ByS1VpgRZ.

Nahid, A.-A., Mikaelian, A., Kong, Y. Histopathological breast-image classification with restricted Boltzmann machine along with backpropagation (2018) Biomedical Research (India), 29 (10), pp. 2068-2077.

Ng, A. (2013). Sparse autoencoder lecture notes. Retrieve from: web.stanford.edu/class/cs294a/sparseAutoencoder.pdf

Ngo, T.A., Lu, Z., Carneiro, G., 2017. Combining deep learning and level set for the au- tomated segmentation of the left ventricle of the heart from cardiac cine mag- netic resonance. Med. Image Anal. 35, 159–171. doi: 10.1016/j.media.2016.05.009 .

Odena, A., Buckman, J., Olsson, C., Brown, T., Olah, C., Raffel, C. & Goodfellow, I.. (2018). Is Generator Conditioning Causally Related to GAN Performance?. Proceedings of the 35th International Conference on Machine Learning, in PMLR 80:3849-3858.

Odena, A., Olah, C., and Shlens, J. Conditional Image Synthesis With Auxiliary Classifier GANs. ArXiv e-prints, October 2016.

Ortiz, A., Munilla, J., Górriz, J.M., Ramírez, J., 2016. Ensembles of deep learning ar- chitectures for the early diagnosis of the Alzheimer's disease. Int. J. Neural Syst. 26, 1650025. doi: 10.1142/S0129065716500258 .

Partaourides, Harris; Chatzis, Sotirios P (2017). "Asymmetric deep generative models". *Neurocomputing*. **241**: 90. doi:10.1016/j.neucom.2017.02.028.

Payan, A., Montana, G., 2015. Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. arxiv: 1502.02506 .

Pereira, S., Meier, R., McKinley, R., Wiest, R., Alves, V., Silva, C. A., & Reyes, M. (2018). Enhancing interpretability of automatically extracted machine learning features: application to a RBM-Random Forest system on brain lesion segmentation. *Medical image analysis*, *44*, 228-244. https://doi.org/10.1016/j.media.2017.12.009

Pinaya, W.H.L., Gadelha, A., Doyle, O.M., Noto, C., Zugman, A., Cordeiro, Q., Jack- owski, A.P., Bressan, R.A., Sato, J.R., 2016. Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia. Nat. Sci. Rep. 6, 38897. doi: 10.1038/srep38897

Plis, S.M., Hjelm, D.R., Salakhutdinov, R., Allen, E.A., Bockholt, H.J., Long, J.D., John- son, H.J., Paulsen, J.S., Turner, J.A., Calhoun, V.D., 2014. Deep learning for neu- roimaging: a validation study. Front. Neurosci. doi: 10.3389/fnins.2014.00229 .

Radford, A., Metz, L., and Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. CoRR, abs/1511.06434, 2015. URL http://arxiv.org/abs/1511.06434.

Ruslan Salakhutdinov and Geoffrey Hinton. 2009. Deep Boltzmann machines. In Artificial Intelligence and Statistics.PMLR, 448–455.

Ruslan Salakhutdinov and Geoffrey Hinton. 2012. An efficient learning procedure for deep Boltzmann machines. Neural Computation 24, 8 (2012), 1967–2006.

Salah Rifai, Pascal Vincent, Xavier Muller, Xavier Glorot, and YoshuaBengio. 2011. Contractive auto-encoders: explicit invariance during feature extraction. In Proceedings of the 28th International Conference on International Conference on Machine Learning (ICML'11), LiseGetoor and Tobias Scheffer (Eds.). Omnipress, USA, 833-840.

Salakhutdinov R, Mnih A, Hinton G. Restricted Boltzmann machines for collaborative filtering. In: Proceedings of the 24th International Conference on Machine Learning, 2007, 791–8.

Salakhutdinov R. 2015. Learning deep generative models. Annu. Rev. Stat. Appl. 2:361–85

Shen, Dinggang& Wu, Guorong& Suk, Heung-Il. (2017). Deep Learning in Medical Image Analysis. Annual review of biomedical engineering. 19. 10.1146/annurev-bioeng-071516-044442.

Shin, H. C., Orton, M. R., Collins, D. J., Doran, S. J., & Leach, M. O. (2013). Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. *IEEE transactions on pattern analysis and machine intelligence*, *35*(8), 1930-1943.

Shin, H.-C., Orton, M.R., Collins, D.J., Doran, S.J., Leach, M.O., 2013. Stacked autoen- coders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. IEEE Trans. Pattern Anal. Mach. Intell. 35, 1930–1943. doi: 10.1109/TPAMI.2012.277 .

Socher, Eric H. Huang, Jeffrey Pennington, Andrew Y. Ng, and Christopher D. Manning. 2011. Dynamic pooling and unfolding recursive autoencoders for paraphrase detection. In Advances in Neural Information Processing Systems, Vol. 24. Neural Information Processing Systems Foundation, 801–809.

Steinhaus, 1956 STEINHAUS, H. 1956. Sur la division des corpmaterielsen parties.

Su H., Xing F., Kong X., Xie Y., Zhang S., Yang L. (2018). Robust Cell Detection and Segmentation in Histopathological Images Using Sparse Reconstruction and Stacked Denoising Autoencoders. Med Image ComputComput Assist Interv. 2015 Oct;9351:383-390. Epub 2015 Nov 18. DOI: 10.1007/978-3-319-24574-4_46.

Suk, H.-I., Lee, S.-W., Shen, D., 2014. Hierarchical feature representation and multi- modal fusion with deep learning for AD/MCI diagnosis. Neuroimage 101, 569–582. doi: 10.1016/j.neuroimage.2014.06.077

Suk, H.-I., Shen, D., 2013. Deep learning-based feature representation for AD/MCI classification. In: Proceedings of the Medical Image Computing and Computer- Assisted Intervention. In: Lecture Notes in Computer Science, 8150, pp. 583–590. doi: 10.1007/978- 3- 642- 40763- 5 _ 72

Suk, H.-I., Wee, C.-Y., Lee, S.-W., Shen, D., 2016. State-space model with deep learn- ing for functional dynamics estimation in resting-state FMRI. Neuroimage 129, 292–307. doi: 10.1016/j.neuroimage.2016.01.005

T. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng and Y. Ma, "PCANet: A Simple Deep Learning Baseline for Image Classification?," in *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5017-5032, Dec. 2015.doi: 10.1109/TIP.2015.2475625

T. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng and Y. Ma, "PCANet: A Simple Deep Learning Baseline for Image Classification?," in *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5017-5032, Dec. 2015. doi: 10.1109/TIP.2015.2475625

van Tulder, G., & de Bruijne, M. (2016). Combining generative and discriminative representation learning for lung CT analysis with convolutional restricted boltzmann machines. *IEEE transactions on medical imaging*, *35*(5), 1262-1272. https://doi.org/10.1109/TMI.2016.2526687

Vincent P, Larochelle H, Lajoie I, et al. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. J Mach Learn Res

Wani, N., & Raza, K. (2018). Multiple Kernel-Learning Approach for Medical Image Analysis. In *Soft Computing Based Medical Image Analysis* (pp. 31-47).

Wu, J., Ruan, S., Mazur, T.R., Daniel, N., Lashmett, H., Ochoa, L., Zoberi, I., Lian, C., Gach, H.M., Mutic, S., Thomas, M., Anastasio, M.A., Li, H. Heart motion tracking on cine MRI based on a deep Boltzmann machine-driven level set method (2018) Proceedings - International Symposium on Biomedical Imaging, 2018-April, pp. 1153-1156. DOI: 10.1109/ISBI.2018.8363775

Xi Cheng, Li Zhang & Yefeng Zheng (2018) Deep similarity learning for multimodal medical images, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 6:3, 248-252, DOI: 10.1080/21681163.2015.1135299

Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., Madabhushi, A., 2016b. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. IEEE Trans. Med. Imaging 35, 119–130. doi: 10.1109/TMI.2015.2458702

Yi, W., Tsang, K. K., Lam, S. K., Bai, X., Crowell, J. A., & Flores, E. A. (2018). Biological plausibility and stochasticity in scalable VO 2 active memristor neurons. *Nature Communications*, *9*(1), 4661.

Yoo Y, Brosch T, Traboulsee A, et al. Deep learning of image features from unlabeled data for multiple sclerosis lesion segmentation. In: International Workshop on Machine Learning in Medical Imaging, Boston, MA, USA, 2014, 117–24.

Zabalza, J., Ren, J., Zheng, J., Zhao, H., Qing, C., Yang, Z., ... & Marshall, S. (2016). Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. *Neurocomputing*, *185*, 1-10.

Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., and Metaxas, D. StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks. ArXiv e-prints, December 2016.

Zhang, Q., Xiao, Y., Dai, W., Suo, J., Wang, C., Shi, J., Zheng, H., 2016c. Deep learning based classification of breast tumors with shear-wave elastography. Ultrasonics 72, 150–157. doi: 10.1016/j.ultras.2016.08.004 .

Zhu, Y., Wang, L., Liu, M., Qian, C., Yousuf, A., Oto, A., Shen, D., 2017. MRI Based prostate cancer detection with high-level representation and hierarchical classification. Med. Phys. 44 (3), 1028–1039. doi: 10.1002/mp.12116 .

Bibliography

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., ... & Xie, W. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface*, *15*(141), 20170387.

Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: a review. *Computer methods and programs in biomedicine*.

Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep learning applications in medical image analysis. *IEEE Access*, *6*, 9375-9389.

Liu, J., Pan, Y., Li, M., Chen, Z., Tang, L., Lu, C., & Wang, J. (2018). Applications of deep learning to mri images: a survey. *Big Data Mining and Analytics*, *1*(1), 1-18.

Mahmud, M., Kaiser, M. S., Hussain, A., & Vassanelli, S. (2018). Applications of deep learning and reinforcement learning to biological data. *IEEE transactions on neural networks and learning systems*, *29*(6), 2063-2079.

Razzak, M. I., Naz, S., & Zaib, A. (2018). Deep Learning for Medical Image Processing: Overview, Challenges and the Future. In *Classification in BioApps* (pp. 323-350). Springer, Cham.

Samira Pouyanfar, Saad Sadiq, Yilin Yan, Haiman Tian, Yudong Tao, Maria Presa Reyes, Mei-Ling Shyu, Shu-Ching Chen, and S. S. Iyengar. 2018. A Survey on Deep Learning: Algorithms, Techniques, and Applications. ACM Comput. Surv. 51, 5, Article 92 (September 2018), 36 pages. https://doi.org/10.1145/3234150

Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual review of biomedical engineering*, *19*, 221-248.