Journal of Theoretical and Applied Information Technology
31st May 2021. Vol.99. No 10

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

DEEP LEARNING-BASED CANCER CLASSIFICATION FOR MICROARRAY DATA: A SYSTEMATIC REVIEW

[1,*]NASHAT ALREFAI, [1,2]OTHMAN IBRAHIM
[1]School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[2]Azman Hashim International Business School, Universiti Teknologi Malaysia (UTM), Skudai, Johor 81310, Malaysia

E-mail: [1]amnashat@graduate.utm.my, [2]othmanibrahim@utm.my

ABSTRACT

Deep neural networks are robust techniques and recently used extensively for building cancer classification models from different types of data. Nowadays, microarray gene expression datasets consider an essential source of data that is used in cancer classifications. However, due to the small size of samples compared to the high dimensionality of microarray data, many machine learning techniques have failed to distinguish the most relevant and informatics genes. Therefore, deep learning is demand due to its ability to automatically discovering the complex relationship between features with significant accuracy and high performance.

The current study aims to reveal the state-of-the-art of deep neural network architectures and how it can utilize from microarray data. Therefore, several deep neural network architectures were built such as CNN, DNN, RNN, DBN, DBM and DAE to be compatible with the different learning processes (supervised, unsupervised and semi-supervised). As a result, CNN considers the most common neural network architecture used in the medical field due to its robustness and high performance in cancer classification. Results indicate that choosing suitable architecture of the deep neural network and its hyper-parameters is one of the most difficulties facing the researcher in designing models for cancer prediction and classification because there is no particular rule to ensure high prediction accuracy.

Keywords: *Deep Learning; Cancer Classification; Convolutional Neural Network; Transfer Learning; Microarray.*

1. INTRODUCTION

This guide provides details to assist authors in preparing a paper for publication in JATIT so that there is a consistency among papers. These instructions give guidance on layout, style, illustrations and references and serve as a model for authors to emulate. Please follow these specifications closely as papers which do not meet the standards laid down, will not be published.

Cancer is considering a common disease with a high death percentage among humans. It is considered the second reason of death in all continents of the world. With the early and accurate diagnosis for cancer, survival will increase significantly, the cancer death rate can be reduced in the case of early detected and diagnosis and treated [1]. Therefore, an accurate and reliable system is necessary for the early diagnosis of cancer.

There are over 200 categories of cancer [2]. Breast, colorectal, lung, and cervical consider the most popular types of cancer in women. Globally, around 1 in 6 deaths among humans is due to cancer. The most frequent cancers are lung (2.09 million cases), breast (2.09 million cases), colorectal (1.80 million cases) and prostate (1.28 million cases). With the early diagnosis of cancer, survival will increase from 56% to more than 86% [3]. **Figure 1** shows the percentage of cancer incidence, distributed on the continents of the world. Therefore, an accurate and reliable system is necessary for the early diagnosis and classification of cancer.

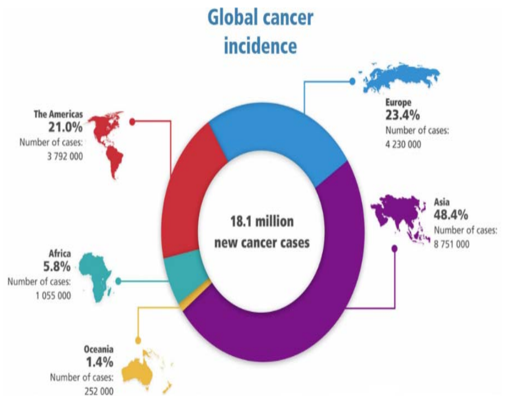


Figure 1: Percentage of cancer cases in different continents in 2018 [4]

Microarray dataset also is known as gene expression profile is usually constructed as a two dimensional array N×(M+1); where N is tissue samples represented as the number of rows or instances, and M is the gene expression level represented as the number of columns or features, one was added to the last column to present the class, usually, this class can be labeled, unlabeled or combination between them.

Microarray dataset analysis and classification procedure have proved strongly that it can offers an active and professional way for the classification and diagnosis of diseases, especially in cancer, many machine and deep learning methods have been used to tumors classifications using microarray datasets [5], [6]. A correct cancer diagnosis and classification is necessary for the successful application of particular therapies. Moreover, deep learning also proves the high performance when using gene expression microarray datasets, it has the ability to select the more informatics genes, which may boost the classification performance [7], [8].

For classification issues, a gradual change of research trends in disease diagnosis and classification from shallow handcrafted algorithms to deep learning [9]. The biggest pro of deep learning is the simultaneous training procedures of feature related tasks like selection, extraction, reduction and classification. Moreover, the reason for it is success is the capability to understand a complex non-linear relationships between features [10]. However, deep learning proves superior result in dealing with big data. On the other side, transfer learning used to overcome the small data problem and to decrease the computational complexity.

The transfer learning is an alternative technique to work with small dataset, and also reduce the training time by partially fine tuning some parameters. The knowledge obtained from large and normally public datasets is transferred to work with the small medical dataset [11]. Transfer learning also can utilize from the knowledge obtained through source tasks from irrelevant and relevant domains [12].

In this study, we answer many questions represent the relationship between the deep learning and various types of cancer to prove that deep learning achieves promising result in cancer classification comparing with traditional machine learning methods. As a result, we can determine the most common type of cancer used extensively in the recent researches and the most suitable deep neural network architecture used with microarray data.

2. METHODOLOGY

2.1 Research Questions

This systematic review aims to recognize different studies related to the cancer classification based on microarray medical datasets and deep learning methods for classification. To progress in a structured plan as a road map through this systematic review is done based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) method which was proposed by Moher [13], it is consider helpful method that can be following for writing systematic review properly with robust structure and easy tracing structure from top to down. In this systematic literature review paper, the research questions will define clearly and describe in terms of Participants, Interventions, Comparisons, Outcomes and Study design (PICOS). This systematic literature review paper aim to find the solution of the following research questions:

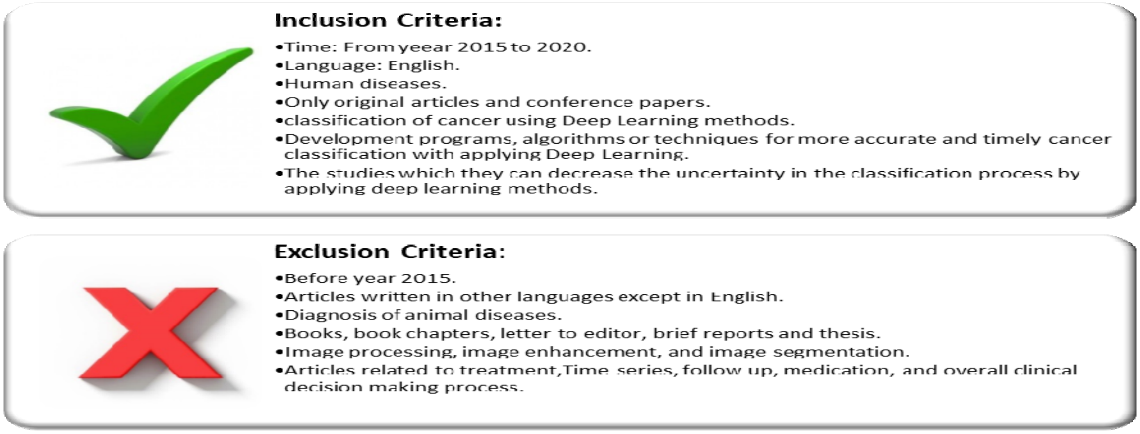
- What types of cancer were being used extensively to applying Deep Learning DL methods to improve cancer classification?
- What are the datasets used for the development of cancer classification based on deep learning DL?
- Which Deep Learning DL methods have been more used between experts in recent year (last five years)?
- What are the common methods currently applied for Deep Learning DL based on microarray medical datasets?

- What are the evaluation criteria used to evaluate the accuracy of cancer classification by using deep learning?
- Are Deep Learning DL methods can improve cancer classification?

In this systematic review we try to collect most appropriate evidence that match pre-specified eligibility criteria to find the answer of a particular research review questions. It uses unambiguous, systematic methods to reduce bias in the identification, choice, synthesis, and summary of studies[14]. When all steps are done correctly, that can provide robustness results that lead to a reliable conclusion. The main features of a systematic review are: a clear objective with a reusable methodology, tries to collect all studies that come across the eligibility criteria, validate the findings of the selected studies and present and synthesis the most important characteristics and findings of the obtained studies.

2.2 Inclusion and Exclusion Criteria

The studies which involved in this systematic review paper met the inclusion and exclusion criteria illustrated in **Figure 2**. Eligible articles that met the inclusion criteria are selected, and according to the exclusion criteria thesis, book chapters, letters and reports are excluded. Moreover, non-English, non-human diseases, image (processing, enhancement and segmentation) and all articles related to treatment, time series, follow up, medication, and overall clinical decision-making process are excluded. In addition, this study selects the related articles which are published last five years.



Inclusion Criteria:

- Time: From yeear 2015 to 2020.
- Language: English.
- Human diseases.
- Only original articles and conference papers.
- classification of cancer using Deep Learning methods.
- Development programs, algorithms or techniques for more accurate and timely cancer classification with applying Deep Learning.
- The studies which they can decrease the uncertainty in the classification process by applying deep learning methods.

Exclusion Criteria:

- Before year 2015.
- Articles written in other languages except in English.
- Diagnosis of animal diseases.
- Books, book chapters, letter to editor, brief reports and thesis.
- Image processing, image enhancement, and image segmentation.
- Articles related to treatment, Time series, follow up, medication, and overall clinical decision making process.

Figure 2: Inclusion and exclusion criteria for selecting eligible articles.

2.3 Search Criteria

To determine the state-of-the-art related with deep learning methods on cancer classification, the search procedure was done on six electronic databases to find the relative articles according to our research questions: Science Direct, PubMed, IEEE, Web of Science, Scopus and Scholar. Only the studies published from the year 2015 until November, 2020 and meeting the inclusion criteria were considered for this study. **Table 1** illustrates the keywords which were used in each database. Accordingly, based on search strategy, 278 records were retrieved. Each database results are shown in PRISMA diagram, **Figure 3** show the PRISMA diagram protocol for the identification, screening, eligibility and included articles. On the other hand, all citations were entered into mendeley bibliographic software, 24 duplicated studies were removed automatically by the software. As a result, 254 relative articles remained. After investigated titles and abstracts to decide which articles meet the criteria, hence, irrelevant articles were removed.

Table 1 Search strategy in various databases.

Database	Keywords	Number of articles
Science direct	Title, abstract, keywords: ("cancer classification" AND "Deep Learning" AND "microarray")	46
IEEE	((("Abstract":" deep learning ") AND "Abstract":" cancer classification") AND "Full Text & Metadata":microarray)	28
WOS	TS=(”microarray”) AND AB=((“deep learning” OR ”Transfer learning") AND "Cancer")) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article)	28
Scopus	***(TITLE-ABS-KEY (cancer) AND TITLE (("microarray" OR "gene expression")) AND TITLE-ABS-KEY ("deep learning"))***	35
Pubmed	"cancer classification" AND "Deep Learning"	34
Scholar	allintitle: "Deep Learning" "cancer classification"	107

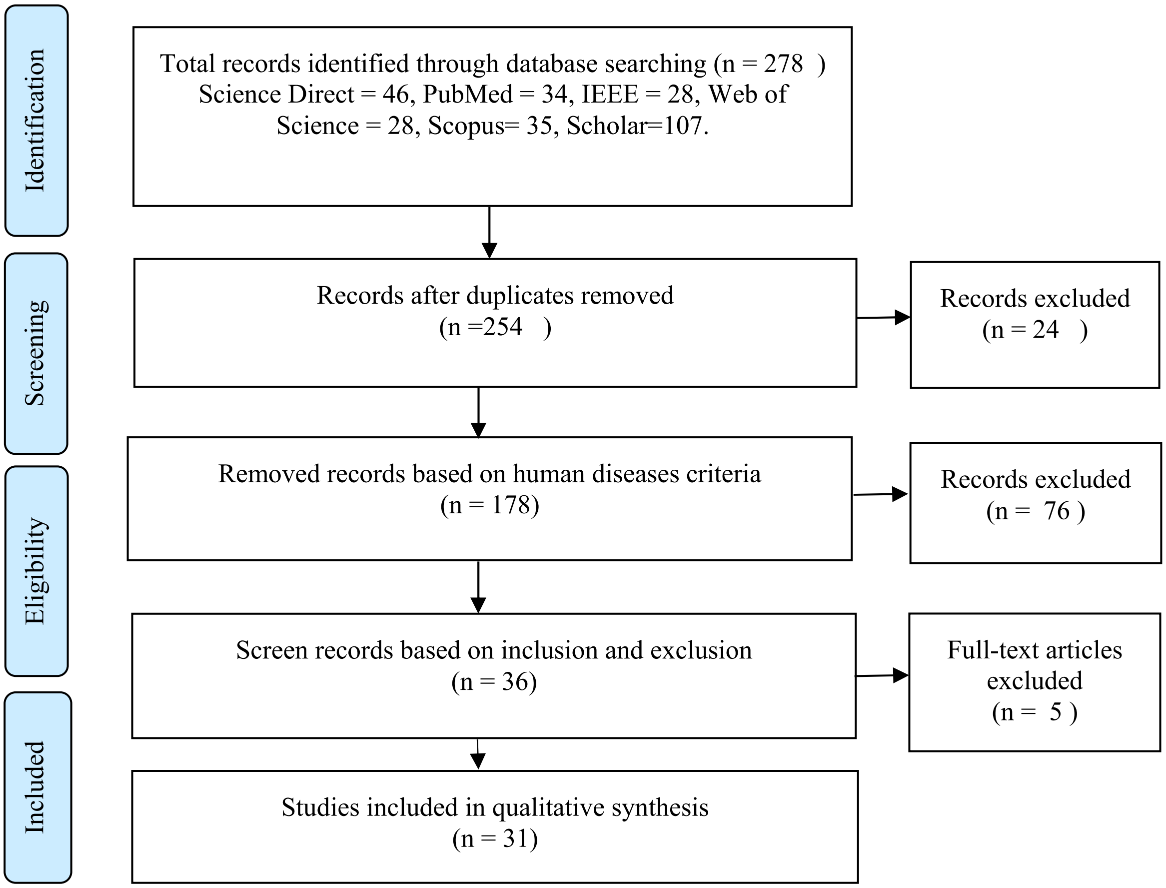


Figure 3. PRISMA diagram for the identification, screening, eligibility and included articles.

2.4 Data Extraction

The data extracted from the chosen articles after examined the full-text are illustrated in Tables 2 which presents the search results for Science Direct, IEEE, Web of Science, Scopus, PubMed and Scholar respectively. In the reviewed literature, **it was found that the breast cancer is the most popular sort of cancers have garnered the attention of the authors and the convolutional neural network is frequently method used in most of the articles.** All researches that met the criteria were studied in details and clarify the advantages, limitations, remarks, finding and results for each article. Furthermore, the most common used performance measures in the literature are: Accuracy (Acc), Sensitivity, Specificity, and Area Under the Curve (AUC).

3. Results

This systematic literature review reviewed the articles which related to cancer classification by using deep learning methods from 2015 till November 2020. Figure 4 indicates the rate of publications in this period of time. Furthermore, it was found that an intersection among common challenges which was faced the authors when using microarray dataset such as the high dimensional and the low sample size of this type of datasets. Alternatively, deep learning techniques need a lot of data for training procedure and more computational complexity for more accuracy. Hence, increasing the number of samples artificially and reduce the computational time and complexity was demand besides increase the accuracy of classification.

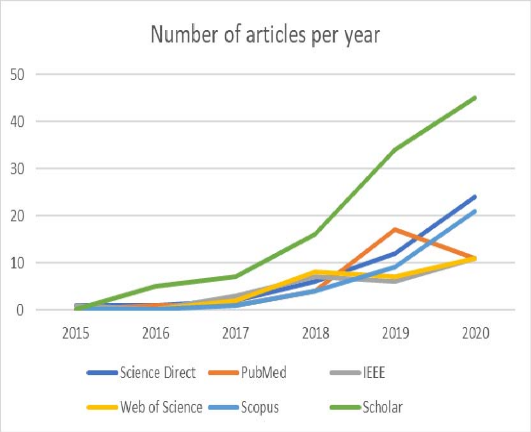


Figure 4 Number of papers published from 2015 to November 2020 using specified criteria.

In this systematic review our interest was collects all deep learning methods which are used for cancer classification based on microarray datasets. Moreover, this section illustrates the analysis of the result shown in **Table 2**.

3.1 Microarray

Genetic diseases are spread rapidly especially the cancer, and microarray considers one of the widest technologies used for diagnosis, prediction, survivability of these types of diseases. Microarray dataset contains thousands of gene expressions called features and a little number of rows called patients or instances. Every feature represents the level of activity of genes within a given tissue, so by comparing the normal and abnormal cancerous tissue can reveal a very specific details insight the cell. Hence, we can increase the reliability of diagnosis and predict the value of the unseen label.

Although microarray considered a rich source of information and can be helpful in diagnosis and prediction, it prone to the curse of high dimensionality which increases the computational time and complexity. Furthermore, the low number of samples puts it at the risk of overfitting. Feature selection and feature extraction approaches can be useful to solve the high dimensionality problem and data augmentation can be utilized to solve the low samples size[15], [43].

Feature selection methods can choose the most informative and relevant, and eliminate the redundant features. Feature extraction consider more general comparing with feature selection, the idea of feature extraction is transforming the dimensionality from distinct space to another one, but the medical data is very sensitive and there is a complex relationship between features. Hence, the new distinct space may have no physical meaning for interpretation. Although, the transformation may provide a better discriminatory ability [44].

3.2 Data Augmentation

Deep learning needs a big data; therefore, many researchers have found that extended the dataset by augmentation or generation procedure can enhance the classification performance. On the other side, the medical data consider very sensitive data that mean, any stochastic expansion in the dataset may lead to an unreliable result. Semi-supervised learning can generate data by adding the unlabeled data to the labeled data. Moreover, it has a technique to selected the most confidence unlabeled after labelled them, this procedure can produce significant enhancement in learning accuracy, the classification accuracy of semi-supervised was increased by adding unlabeled samples to the training set compared with using a small percentage of labeled data for training. However, the performance of semi-supervised learning never superior the supervised learning with the same amount of labeled training data [45].

Small data consider a challenge in machine and deep learning and cause numerous problems such as over-fitting and outliers. Therefore, there are many techniques used to address this problem in images datasets such as data augmentation (flips, crops and brightness), transformation including (zooming, translation, rotation and shearing), batching strategy by converting one sample to many samples.

Data generation is another technique for increase the number of samples such as Synthetic Minority Oversampling Technique (SMOTE) which utilize from oversampling technique, R.Gao et al.(2019) [46] used SMOTE for artificially over-sampling the data, hence, the classification accuracy have improved significantly. Generative Adversarial Networks (GAN) another technique was used by M. Frid-Adar et al.(2019) [47] for increased convolutional neural network Performance in the classification of Liver Lesion by increasing the data synthetically.

Another data generation technique is semi-supervised which utilize from labeled and unlabeled data together. Triguero [48] make an experimental study of the state-of-the-arts self-labeled techniques to find the best method and the best ratio of the labeled data, and used an open-source semi-supervised learning model for knowledge extraction based on Evolutionary Learning (KEEL) software tool [49].

Normally, the medical data different from natural data, any augmentation may be effect on the reliability, because there is an interrelationship between features. Jian Liu [15] proposes a new Sample Expansion (SE) method to handle the issue of lacking labeled data. Inspired from Denoising Autoencoder (DAE), a huge number of labeled data are collected by randomly cleaning partially corrupted input more than one time. These labeled data are used as the extra data. Then, mixed the extra data and unprocessed data into a matrix as the training data.

Trivizakis [25] Extending two dimensional Convolutional Neural Networks (CNN) to three dimensions for the classification of cancer where is working directly on whole three dimensions tomographic data without using any preprocessing techniques. As a result, significant tissue classification accuracy enhancement compared to two dimensions CNNs of various architectures also created for the appropriate clinical problem with the same dataset.

3.3 Deep Learning Architectures

Machine learning algorithms are suitable way for cancer classification, but due to Excessive increase in the amount of data the need for deep learning has appeared, because most traditional machine learning algorithms are not affected by the steady increase of data, it gives results of constant accuracy despite the increase in data **Figure 5**

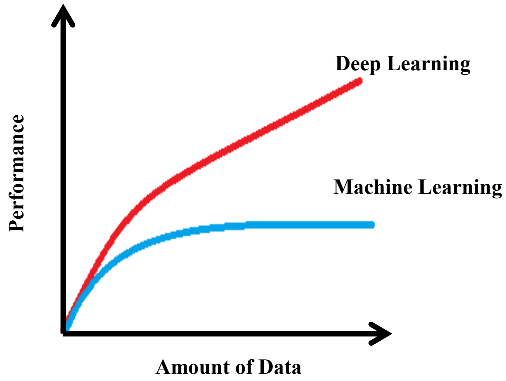
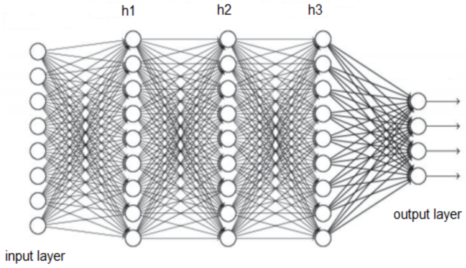
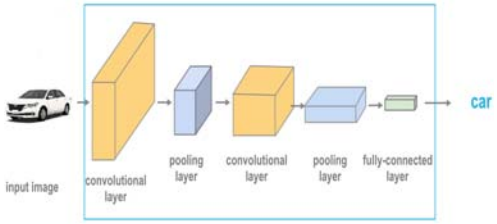
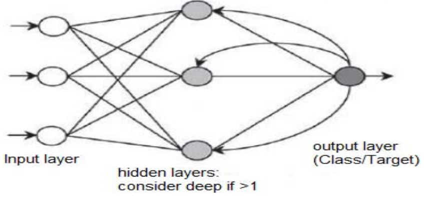
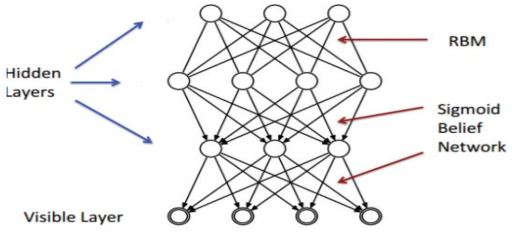


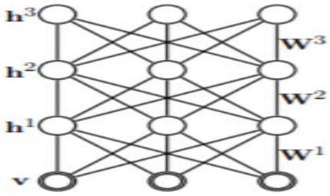
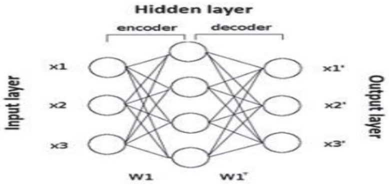
Figure 5 Comparison between the performance of deep learning algorithms and traditional machine learning algorithms according to the amount of data

There are different kinds of architectures of deep learning; each type has particular use case such as the type of data, whether it is supervised or unsupervised or semi-supervised and the type of task which we want to solve. Consequently, many deep learning methods take place such as: Deep Neural Network (DNN), Convolutional Neural Network (CNN), Deep Belief Network (DBN), Deep Boltzmann Machine (DBM), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Generative Adversarial Network (GAN), Auto Encoders (Denoising autoencoder, Stacked autoencoder, Variational autoencoder, Sparse autoencoder, Contractive autoencoder and Convolutional autoencoder).

Table 3 illustrate the most common deep learning architectures.

Table 3 Different Deep Learning Architectures

Architecture	Description
Deep Neural Network (DNN) 	Deep neural network composed of many hidden layers and can handle the non-linear interrelationship between them. Moreover, it is commonly used for regression and classification. Although it was used widely in many fields, the learning process still very slow due to forward and backward propagation.
Convolutional Neural Network (CNN) 	CNN mimics the brain process in detection and classification of the objects in deep way with the idea of multi-hidden layers as a stack, connectivity, weight sharing and pooling. The main building blocks of CNN are input layer, convolutional layer, pooling or subsampling layer, ReLU layer, dense layer, fully connected layer and output layer.
Recurrent Neural Network (RNN) 	RNN usually used with streams or sequential of data where the output depends on the previous computations, therefore, it has the ability of memory, such as DNA sequences, text and speech, which depend on time series and sequential memory. Moreover, RNN share the same weights through all steps which decrease the computational complexity.
Deep Belief Network (DBN)  [50]	DBN consist of many layers of stochastic, hidden variables. it allows unsupervised and supervised training of the network, consider a composition of restricted Boltzmann machine (RBM) where each sub-network's hidden layer serves as the visible layer for the next. Moreover, it has not directed connections expect at the top two layers. This type of network proposed layer by layer greedy learning strategy to initialize the network but this consider computationally expensive.

Deep Boltzmann Machine (DBM)  [51]	DBM is a special BM where the hidden units are organized in a deeply layers manner, have the ability to holds undirected connections among all layers of the network; moreover, it is normally used stochastic maximum likelihood algorithm to maximize the lower bound of the likelihood. The main drawback of this type of network is the parameter optimization especially when the dataset is large and the time complexity for inference. As a result of incorporates top-down feedback.
Deep Autoencoder (DAE) 	Deep Autoencoder is commonly used in case of dimensionality reduction and feature extraction. The main characteristics of DAE that number of input and output nodes are the same, furthermore, it is considered unsupervised learning method, so, there is no need for lebeled data, and it is aim to create the input vector. The main drawback of DAE is the requirement of a pre-training stage. Some common variations have been proposed (Sparse, Denoising, Contractive, convolutional and stacked autoencoder).

Recently, deep learning used intensively in health informatics such as: medical information, bioinformatics, imaging, sensing and public health. In this study we will concentrate on using deep learning in cancer classification, **figure 6** explain the number of publications which are related to deep learning in cancer classification last five years. Significantly, **figure 7** show the number of publications related to cancer disease diagnosis, prediction and classification based on different deep learning network architectures.

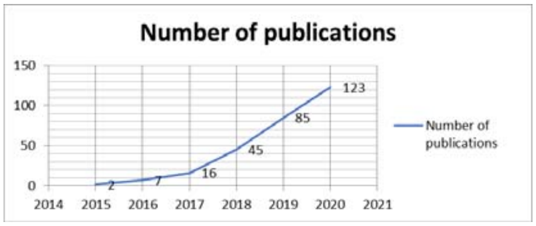


Figure 6 number of publications which are related to using a deep learning in cancer classification last five years.

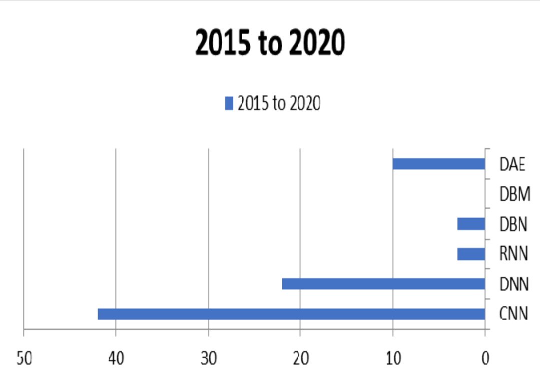


Figure 7 number of publications for different architecture of deep learning in cancer last five years (microarray and images dataset).

The most popular programming languages used in deep learning are Python, R, Matlab, C++ and Java. However, Python consider the most popular language in 2018 and 2019 according to the IEEE Spectrum statistical report [52]. Recently many frameworks and tools were developed to facilitate dealing with deep learning network architectures such as: Tensorflow, Keras, Theano, Caffe, Torch, MXNet, OpenNN, GPUMLip and Deeplearning4j. Figure 8 explains the most common framework for each language.

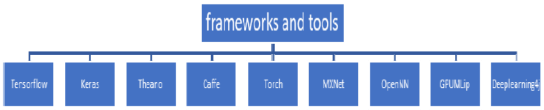


Figure 8 Common frameworks and tool for deep learning

3.4 Evaluation of Metrics

The dataset normally splits into training and testing data samples, usually, 80% for training and 20% for testing or 75% and 25% for training and testing respectively. The training part used to train the network after applying some pre-processing techniques and the testing part used to fed the neural network for classification issues the result will evaluated to assess the performance of the neural network, there are several ways for assessing the network performance. For cancer diagnosis and classifications, commonly, the positive samples represent malignant/cancerous class label and the negative samples represent benign/non-cancerous class label.

For evaluation metrics most researchers repeatedly used accuracy, sensitivity or recall, specificity, precision, F1 score and AUC Area under the Receiver Operating Characteristic curve (ROC), which is a curve of the true positive rate versus the false positive rate. The value of AUC can be between 0 and 1, the value consider better as prediction or diagnosis result when it is closer to one [53]. The formulas for the previous performance measurements can be extracted from confusion matrix Table 4. Where TP: true positive, TN: true negative, FP: false positive, FN: false negative.

Table 4 Confusion Matrix

		Predicted Class	
		Negative Class (noncancerous)	Positive Class (cancerous)
Actual Class	Negative Class (noncancerous)	True Negative (TN)	False Positive (FP)
	Positive Class (cancerous)	False Negative (FN)	True Positive (TP)

In deep learning the following metrics are repeatedly used for prediction, diagnosis, classification and survivability to evaluate the performance of the various neural network architectures:

$$Sensitivity(Recall) = \frac{TP}{FP + TN} \quad (1)$$

$$Specfity = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Overall\ accuracy = \frac{TP + TN}{TN + TP + FN + FP} \quad (4)$$

$$F1\ score = 2 * \frac{precision * sensitivity}{precision + sensitivity} \quad (5)$$

3.5 Datasets

Deep learning can be used in different area such as: bioinformatics, medical imaging, medical information and public health. Moreover, DL used in different applications related to the previous area such as: cancer (diagnosis, prediction, detection and classification). Medical dataset related to cancer disease can be collected in different ways to use it in deep learning; it can be sequence data (DNA sequence, RNA sequence), MRI images, Cryo-EM images, florescence microscopy images, protein contact map, CT images, PET images, Microarray gene expression[54].

3.6 Challenges and Issues

3.6.1 High dimensionality and reducing the computational complexity and training time

Machine learning and deep learning techniques prone to the curse of high dimensionality which cause the disaster of overfitting. As a result, the Dimensionality reduction techniques are demand, Dimensionality reduction techniques map high-dimensional data to a lower-dimensional space [55]. The main purpose of dimensionality reduction is to find the most useful and informative subspace that can reduces the computational complexity and training time for both machine and deep learning [56]. In some condition when the dataset is microarray, normally the number of features is often more than the number of instances [57].

There are two main aspects in decreasing the dimensionality: feature selection and feature extraction. Feature selection is the selection of the most informative portion of the original dataset, while feature extraction is extracting a small subset from original dataset with new representation [58], in other words it transform the search process from original search space to another, however, in medical data that may effect on the reliability especially with strong and various relationships between features.

3.6.2 Low sample size and overfitting

Due to the small number of samples overfitting can be occurred and make the dataset unstable. Therefore, many researchers proposed different data augmentations techniques to override this challenge such as : data augmentation (flips, crops and brightness), transformation including (zooming, translation, rotation and shearing), batching strategy by converting one sample to many samples and some researchers added noise to increase the number of samples [59]. On the other hand, the medical data consider very sensitive data that mean, any stochastic expansion in the dataset may lead to an unreliable result. Therefore, data generation is another robustness technique which considers suitable for this situation to extend the number of samples such as: Synthetic Minority Oversampling Technique (SMOTE), Generative Adversarial Networks (GAN) and semi-supervised methods.

3.6.3 Data pre-processing, Class imbalance data, Missing values and Outliers detections.

Data pre-processing step consider very essential step to guarantee that the dataset is clean from missing data and not contain outlier values because treat the previous problems can enhance the performance of the model and increase the accuracy and avoid poor result, at the same time, the class should be checked if it was imbalance. Class imbalance problem, where the main class of interest is rare. That is, the data set distribution reflects a significant majority of the negative class and a minority positive class [53], [60].

An outlier is an observation that lies an abnormal distance from other values in a population that have an effect on the model performance and lead to Deterioration in the accuracy. Therefore, several techniques were proposed to solve this problem [61]–[63]. And also, for missing data there are many techniques proposed to deal with this problem [64]–[66].

4 Conclusion

Recently deep learning achieves promising results in medical field especially in cancer diagnosis, prediction and classification. Therefore, several deep neural network architectures were built such as CNN, DNN, RNN, DBN, DBM and DAE to be compatible with the different learning processes (supervised, unsupervised and semi-supervised). Moreover, medical data can be in different types and formats and deep learning need a big amount of data in the training process to create a high performance and accurate model.

In this paper, we reveal different architectures of deep neural network used in cancer disease by using different types of data. As a result, CNN considers the most common neural network architecture used in this field due to its robustness and high performance in cancer classification. On the other hand, microarray considers one of the most common technologies used for diagnosis, prediction, classification and survivability of cancer diseases because it contains thousands of gene expressions which consider a rich source of information. However, dealing with this type of dataset have many challenges to the researchers, this systematic lecturer review tried to focus on the main challenges such as the high dimensionality and low sample size which can cause overfitting. Furthermore, reducing the computational complexity and training time. In addition, Class imbalance data, Missing values and Outliers detections.

Choosing the deep neural network architecture is one of the most difficulties facing the researcher in designing models for cancer prediction and classification because there is no specific rule to guarantee high prediction accuracy. Hence, trial and error are the most strategies that follow by most researchers in the previous literature to determine the number of hidden layers and neurons in the deep neural network. In future work, designing an automatic method can be suggested to select the suitable number of the hidden layers and the values of the hyperparameters used in the neural network architectures.

ACKNOWLEDGEMENTS

Funding
This study did not receive external funding.
Conflict of interest
The authors declare that they have no conflict of interest.

REFERENCES

[1] M. Montazeri, "Machine learning models in breast cancer survival prediction," *Technol. Heal. Care*, vol. 24, no. 1, pp. 31–42, Jan. 2016, doi: 10.3233/THC-151071.

[2] M. Prabukumar, L. Agilandeeswari, and K. Ganesan, "An intelligent lung cancer diagnosis system using cuckoo search optimization and support vector machine classifier," *J. Ambient Intell. Humaniz. Comput.*, vol. 10, no. 1, pp. 267–293, 2019, doi: 10.1007/s12652-017-0655-5.

[3] "Cancer," *World Health Organization*. Nov. 22, 2018, [Online]. Available: http://www.who.int/news-room/fact-sheets/detail/cancer.

[4] "New Global Cancer Data: GLOBOCAN 2018 | UICC," 2018. https://www.uicc.org/new-global-cancer-data-globocan-2018 (accessed Dec. 08, 2018).

[5] H. Alshamlan, G. Badr, and Y. Alohali, "mRMR-ABC: A Hybrid Gene Selection Algorithm for Cancer Classification Using Microarray Gene Expression Profiling.," *Biomed Res. Int.*, vol. 2015, p. 604910, Apr. 2015, doi: 10.1155/2015/604910.

[6] M. Maniruzzaman *et al.*, "Statistical Characterization and Classification of Colon Microarray Gene Expression Data using Multiple Machine Learning Paradigms," *Comput. Methods Programs Biomed.*, vol. 176, pp. 173–193, 2019, doi: 10.1016/j.cmpb.2019.04.008.

[7] A. M. Sarhan, "Cancer Classification Based on Microarray Gene Expression Data Using DCT and ANN," *J. Theor. Appl. Inf. Technol.*, pp. 208–216, 2009, Accessed: Jun. 19, 2019. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7881563/.

[8] M. Panda, "Elephant search optimization combined with deep neural network for microarray data analysis," *J. King Saud Univ. - Comput. Inf. Sci.*, 2018, doi: 10.1016/j.jksuci.2017.12.002.

[9] L. Zheng, Y. Yang, and Q. Tian, "SIFT Meets CNN: A Decade Survey of Instance Retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 5, pp. 1224–1244, May 2018, doi: 10.1109/TPAMI.2017.2709749.

[10] R. K. Sevakula, V. Singh, N. K. Verma, C. Kumar, and Y. Cui, "Transfer Learning for Molecular Cancer classification using Deep Neural Networks," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, pp. 1–1, 2018, doi: 10.1109/TCBB.2018.2822803.

[11] Y. Yuan *et al.*, "Prostate cancer classification with multiparametric MRI transfer learning model," *Med. Phys.*, vol. 46, no. 2, pp. 756–765, Feb. 2019, doi: 10.1002/mp.13367.

[12] R. K. Samala, H.-P. Chan, L. Hadjiiski, M. A. Helvie, C. D. Richter, and K. H. Cha, "Breast Cancer Diagnosis in Digital Breast Tomosynthesis: Effects of Training Sample Size on Multi-Stage Transfer Learning Using Deep Neural Nets," *IEEE Trans. Med. Imaging*, vol. 38, no. 3, pp. 686–696, Mar. 2019, doi: 10.1109/TMI.2018.2870343.

[13] D. Moher *et al.*, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement," *Syst. Rev.*, vol. 4, no. 1, p. 1, Dec. 2015, doi: 10.1186/2046-4053-4-1.

[14] E. M. Antman, J. Lau, B. Kupelnick, F. Mosteller, T. C. Chalmers, and H. Chalmers, "A Comparison of Results of Meta-analyses of Randomized Control Trials and Recommendations of Clinical Experts Treatments for Myocardial Infarction," 1992. Accessed: Oct. 18, 2019. [Online]. Available: http://jama.jamanetwork.com/.

[15] J. Liu, X. Wang, Y. Cheng, and L. Zhang, "Tumor gene expression data classification via sample expansion-based deep learning," *Oncotarget*, vol. 8, no. 65, pp. 109646–109660, 2017, doi: 10.18632/oncotarget.22762.

[16] B. Wei, Z. Han, X. He, and Y. Yin, "Deep learning model based breast cancer histopathological image classification," *2017 2nd IEEE Int. Conf. Cloud Comput. Big Data Anal. ICCCBDA 2017*, pp. 348–353, 2017, doi: 10.1109/ICCCBDA.2017.7951937.

[17] R. K. Samala, H.-P. Chan, L. M. Hadjiiski, M. A. Helvie, K. Cha, and C. Richter, "Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms," *Phys. Med. Biol.*, Oct. 2017, doi: 10.1088/1361-6560/aa93d4.

[18] I. Arvidsson *et al.*, "Generalization of

prostate cancer classification for multiple sites using deep learning," *Proc. - Int. Symp. Biomed. Imaging*, vol. 2018-April, no. Isbi, pp. 191–194, 2018, doi: 10.1109/ISBI.2018.8363552.

[19] M. Nishio *et al.*, "Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning," *PLoS One*, vol. 13, no. 7, p. e0200721, Jul. 2018, doi: 10.1371/journal.pone.0200721.

[20] Shallu and R. Mehra, "Breast cancer histology images classification: Training from scratch or transfer learning?," *ICT Express*, vol. 4, no. 4, pp. 247–254, Dec. 2018, doi: 10.1016/j.icte.2018.10.007.

[21] M. Nawaz, A. A. Sewissy, and H. A. Soliman, "Multi-Class Breast Cancer Classification using Deep Learning Convolutional Neural Network," 2018. Accessed: Jun. 22, 2019. [Online]. Available: www.ijacsa.thesai.org.

[22] Q. Liao, Y. Ding, Z. L. Jiang, X. Wang, C. Zhang, and Q. Zhang, "Multi-task deep convolutional neural network for cancer diagnosis," *Neurocomputing*, no. xxxx, 2018, doi: 10.1016/j.neucom.2018.06.084.

[23] K. Lan, T. Li, S. Fong, J. A. L. Marques, and R. K. Wong, "Multi-view Convolution Neural Network with Swarm Search Based Hyperparameter Optimization for Enhancing Heart Disease and Breast Cancer Detection," in *2018 5th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, 2018, pp. 140–146, doi: 10.1109/ISCMI.2018.8703249.

[24] X. Zhong *et al.*, "Deep transfer learning-based prostate cancer classification using 3 Tesla multi-parametric MRI," *Abdom. Radiol.*, vol. 44, no. 6, pp. 2030–2039, Jun. 2019, doi: 10.1007/s00261-018-1824-5.

[25] E. Trivizakis *et al.*, "Extending 2-D Convolutional Neural Networks to 3-D for Advancing Deep Learning Cancer Classification With Application to MRI Liver Tumor Differentiation," *IEEE J. Biomed. Heal. Informatics*, vol. 23, no. 3, pp. 923–930, May 2019, doi: 10.1109/JBHI.2018.2886276.

[26] B. Xu *et al.*, "An End-to-End Deep Learning Histochemical Scoring System for Breast Cancer TMA," *IEEE Trans. Med. Imaging*, vol. 38, no. 2, pp. 1–1, 2019, doi: 10.1109/tmi.2018.2868333.

[27] S. Khan, N. Islam, Z. Jan, I. Ud Din, and J. J. P. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognit. Lett.*, vol. 125, pp. 1–6, 2019, doi: 10.1016/j.patrec.2019.03.022.

[28] C. Xia, Y. Xiao, J. Wu, X. Zhao, and H. Li, "A Convolutional Neural Network Based Ensemble Method for Cancer Prediction Using DNA Methylation Data," pp. 191–196, 2019, doi: 10.1145/3318299.3318372.

[29] S. H. Shah, M. J. Iqbal, I. Ahmad, S. Khan, and J. J. P. C. Rodrigues, "Optimized gene selection and classification of cancer from microarray gene expression data using deep learning," *Neural Comput. Appl.*, 2020, doi: 10.1007/s00521-020-05367-8.

[30] C. Kandaswamy, L. M. Silva, L. A. Alexandre, and J. M. Santos, "High-Content Analysis of Breast Cancer Using Single-Cell Deep Transfer Learning," *J. Biomol. Screen.*, vol. 21, no. 3, pp. 252–259, Mar. 2016, doi: 10.1177/1087057115623451.

[31] M. Panda, "Elephant search optimization combined with deep neural network for microarray data analysis," *J. King Saud Univ. - Comput. Inf. Sci.*, Dec. 2018, doi: 10.1016/j.jksuci.2017.12.002.

[32] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, "A deep learning-based multi-model ensemble method for cancer prediction," *Comput. Methods Programs Biomed.*, vol. 153, pp. 1–9, Jan. 2018, doi: 10.1016/j.cmpb.2017.09.005.

[33] A. Hajieskandar, J. Mohammadzadeh, M. Khalilian, and A. Najafi, "Molecular cancer classification method on microarrays gene expression data using hybrid deep neural network and grey wolf algorithm," *J. Ambient Intell. Humaniz. Comput.*, no. Young 2000, 2020, doi: 10.1007/s12652-020-02478-x.

[34] H. S. Basavegowda and G. Dagnew, "Deep learning approach for microarray cancer data classification," *CAAI Trans. Intell. Technol.*, vol. 5, no. 1, pp. 22–33, 2020, doi: 10.1049/trit.2019.0028.

[35] P. Guillen and J. Ebalunode, "Cancer Classification Based on Microarray Gene Expression Data Using Deep Learning," in

Proceedings - 2016 International Conference on Computational Science and Computational Intelligence, CSCI 2016, Dec. 2017, pp. 1403–1405, doi: 10.1109/CSCI.2016.0270.

[36] E. Deniz, A. Şengür, Z. Kadiroğlu, Y. Guo, V. Bajaj, and Ü. Budak, "Transfer learning based histopathologic image classification for breast cancer detection," *Heal. Inf. Sci. Syst.*, vol. 6, no. 1, p. 18, Dec. 2018, doi: 10.1007/s13755-018-0057-x.

[37] D. Zhang, L. Zou, X. Zhou, and F. He, "Integrating Feature Selection and Feature Extraction Methods with Deep Learning to Predict Clinical Outcome of Breast Cancer," *IEEE Access*, vol. 6, pp. 28936–28944, 2018, doi: 10.1109/ACCESS.2018.2837654.

[38] K. Adem, S. Kiliçarslan, and O. Cömert, "Classification and diagnosis of cervical cancer with stacked autoencoder and softmax classification," *Expert Syst. Appl.*, vol. 115, pp. 557–564, Jan. 2019, doi: 10.1016/j.eswa.2018.08.050.

[39] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, "A semi-supervised deep learning method based on stacked sparse auto-encoder for cancer prediction using RNA-seq data," *Comput. Methods Programs Biomed.*, vol. 166, pp. 99–105, Nov. 2018, doi: 10.1016/j.cmpb.2018.10.004.

[40] X. Zhang *et al.*, "Deep Learning Based Analysis of Breast Cancer Using Advanced Ensemble Classifier and Linear Discriminant Analysis," *IEEE ACCESS*, vol. 8, pp. 120208–120217, 2020, doi: 10.1109/ACCESS.2020.3005228.

[41] K. Adem, "Diagnosis of breast cancer with Stacked autoencoder and Subspace kNN," *Physica A*, vol. 551, p. 124591, 2020, doi: 10.1016/j.physa.2020.124591.

[42] S. Kilicarslan, K. Adem, and M. Celik, "Diagnosis and classi fi cation of cancer using hybrid model based on ReliefF and convolutional neural network," *Med. Hypotheses*, vol. 137, no. December 2019, p. 109577, 2020, doi: 10.1016/j.mehy.2020.109577.

[43] Q. Wu *et al.*, "Deep learning for predicting disease status using genomic data," vol. 9, no. 5, 2018, doi: 10.7287/peerj.preprints.27123v1.

[44] Zena M. Hira and D. F. Gillies, "A Review of Feature Selection and Feature Extraction Methods Applied on Microarray Data," *Adv. Bioinformatics*, vol. 2015, no. 1, 2015, doi: 10.1155/2015/198363.

[45] Z. Shi, "SEMI-SUPERVISED ENSEMBLE LEARNING METHODS FOR ENHANCED PROGNOSTICS AND HEALTH MANAGEMENT," 2018.

[46] R. Gao, J. Peng, L. Nguyen, Y. Liang, S. Thng, and Z. Lin, "Classification of Non-Tumorous Facial Pigmentation Disorders using Deep Learning and SMOTE," *2019 IEEE Int. Symp. Circuits Syst.*, pp. 1–5, 2019, doi: 10.1109/iscas.2019.8702334.

[47] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification," 2018. Accessed: Jul. 11, 2019. [Online]. Available: https://arxiv.org/pdf/1803.01229.pdf.

[48] I. Triguero, S. García, and F. Herrera, "Self-labeled techniques for semi-supervised learning: Taxonomy, software and empirical study," *Knowl. Inf. Syst.*, vol. 42, no. 2, pp. 245–284, 2015, doi: 10.1007/s10115-013-0706-y.

[49] J. Alcalá-Fdez *et al.*, "KEEL: a software tool to assess evolutionary algorithms for data mining problems," vol. 13, pp. 307–318, 2009, doi: 10.1007/s00500-008-0323-y.

[50] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A Fast Learning Algorithm for Deep Belief Nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006, doi: 10.1162/neco.2006.18.7.1527.

[51] Z. Chen, S. Deng, X. Chen, C. Li, R. V. Sanchez, and H. Qin, "Deep neural networks-based rolling bearing fault diagnosis," *Microelectron. Reliab.*, vol. 75, pp. 327–333, 2017, doi: 10.1016/j.microrel.2017.03.006.

[52] "The Top Programming Languages 2019 - IEEE Spectrum," 2019. https://spectrum.ieee.org/computing/software/the-top-programming-languages-2019 (accessed Oct. 31, 2019).

[53] J. Han, M. Kamber, and J. Pei, "Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)," 2011. Accessed: Dec. 01, 2018. [Online]. Available: http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-

Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf.

[54] Y. Li, C. Huang, L. Ding, Z. Li, Y. Pan, and X. Gao, "Deep learning in bioinformatics: introduction, application, and perspective in big data era," 2019, doi: 10.1101/563601.

[55] K. Q. Weinberger and L. K. Saul, "An Introduction to Nonlinear Dimensionality Reduction by Maximum Variance Unfolding," 2006. Accessed: Nov. 16, 2019. [Online]. Available: www.aaai.org.

[56] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, no. 1–2, pp. 273–324, Dec. 1997, doi: 10.1016/s0004-3702(97)00043-x.

[57] N. Armanfard, J. P. Reilly, and M. Komeili, "Local Feature Selection for Data Classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 6, pp. 1217–1227, 2016, doi: 10.1109/TPAMI.2015.2478471.

[58] S. Pölsterl, S. Conjeti, N. Navab, and A. Katouzian, "Survival analysis for high-dimensional, heterogeneous medical data: Exploring feature extraction as an alternative to feature selection," *Artif. Intell. Med.*, vol. 72, pp. 1–11, 2016, doi: 10.1016/j.artmed.2016.07.004.

[59] J. F. Ramirez Rochac, N. Zhang, L. Thompson, and T. Oladunni, "A Data Augmentation-Assisted Deep Learning Model for High Dimensional and Highly Imbalanced Hyperspectral Imaging Data," *2019 9th Int. Conf. Inf. Sci. Technol.*, pp. 362–367, 2019, doi: 10.1109/icist.2019.8836913.

[60] A. Ali, S. M. Shamsuddin, and A. L. Ralescu, "Classification with class imbalance problem: A review," *Int. J. Adv. Soft Comput. its Appl.*, vol. 7, no. 3, pp. 176–204, 2007.

[61] D. Chakraborty, V. Narayanan, and A. Ghosh, "Integration of deep feature extraction and ensemble learning for outlier detection," *Pattern Recognit.*, vol. 89, pp. 161–171, May 2019, doi: 10.1016/j.patcog.2019.01.002.

[62] A. Tiba, Z. Bartik, H. Toman, and A. Hajdu, "Detecting outlier and poor quality medical images with an ensemble-based deep learning system," *Int. Symp. Image Signal Process. Anal. ISPA*, vol. 2019-Septe, pp. 99–104, 2019, doi: 10.1109/ISPA.2019.8868911.

[63] A. G. C. Pacheco, A.-R. Ali, and T. Trappenberg, "Skin cancer detection based on deep learning and entropy to detect outlier samples," 2019. Accessed: Nov. 17, 2019. [Online]. Available: https://www.isic-archive.com.

[64] J. M. Jerez *et al.*, "Missing data imputation using statistical and machine learning methods in a real breast cancer problem," *Artif. Intell. Med.*, vol. 50, no. 2, pp. 105–115, 2010, doi: 10.1016/j.artmed.2010.05.002.

[65] M. Maniruzzaman *et al.*, "Accurate Diabetes Risk Stratification Using Machine Learning: Role of Missing Value and Outliers," *J. Med. Syst.*, vol. 42, no. 5, May 2018, doi: 10.1007/s10916-018-0940-7.

[66] G. Jeon, A. K. Sangaiah, Y.-S. Chen, and A. Paul, "Special issue on Machine learning approaches and challenges of missing data in the era of big data," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 10, pp. 2589–2591, Oct. 2019, doi: 10.1007/s13042-019-01010-8.

Table 2 Related work on cancer classification using deep learning.

Reference	Year	Method	Cancer type	Advantages	Limitations	Remarks	Finding and results
[15]	2017	CNN	Breast cancer Leukemia Colon cancer	Gene expression datasets normally, have insufficient number of samples for training. The author proposed two models for sample expansion that mimic the idea of Denoising Autoencoder (DAE). The models are: Sample Expansion-Based SAE (SESAE) and Sample Expansion-Based 1DCNN (SE1DCNN).	The description of SE method from a biological view is not enough.	The author used three datasets and The performance of 1DCNN and SE1DCNN prove that 1DCNN is a robust method when performing tumor classification.	The accuracy of SE1DCNN for Breast cancer =95.33% Leukemia =57.87% Colon cancer= 84.9% The accuracy of SESAE for Breast cancer =87.33% Leukemia =49.79% Colon cancer= 84.49%
[16]	2017	CNN	Breast cancer	The author used DCNN to design novel method named BiCNN for classify breast cancer histopathological image, to address the two-class breast cancer classification on the pathological image.	The accuracy cannot generalize, because a lot of labeled data are gathered by randomly cleaning partially corrupted input many times. Moreover, rotation, scaling and mirror which are used by the author may be not reliable with medical data.	The author used data augmentation and trnsfer learning to avoid overfitting when he has lack of data.	Accuracy =97%
[17]	2017	CNN	Breast Cancer Mammograms	The author used multi-task transfer learning DCNN for breast cancer diagnosis. SFMs and DMs datasets was used for train multi-task model, and k-fold cross validation used for training and parameter optimization. The research addresses a limited number of samples by using transfer learning.	Lack of generalization capabilities due to the limited number of samples. Hence, it is recommended to raise the number of data by augmentation techniques.	The p-value for multi-task transfer learning DCNN p=0.007 which significantly considers higher performance than the single-task transfer learning DCNN.	P=0.007
[18]	2018	CNN	Prostate Cancer	The author used CNN-autoencoder for down-sampling of the data. For prostate cancer classification and generalize the procedure	Algorithms trained to perform well on a particular dataset might implement extensively poorer on different datasets.	Using three different preprocessing approaches to extract more information than regular one. The accuracy 81% and 95% for the two types of dataset can be enhancing.	Accuracy =95%
[19]	2018	DCNN	lung nodule cancer	The author proposed a new method for classification among benign nodules, primary lung cancer, and metastatic lung	The author ignored some particular features, for	Using the transfer learning can enhance the accuracy of	Accuracy with transfere learning 60.7%,64.7%,6

				cancer and assessed the following: the advantage of the DCNN, the efficiency of transfer learning and the impact of image size as input to the DCNN.	instance nodule type and size. Using 2D-CNN because it is not easy to perform 3D-CNN on medical images with transfer learning. Also ignoring large image size to avoid computational cost.	DCNN.	8% Accuracy without transfere learning 55.9%,68%,62.4%
[20]	2018	CNN	Breast cancer	The author analyzes the efficacy of transfer learning when it was compared with the fully-trained network for the breast cancer classification using histopathological imaging and MR modality and to decide which pre-trained neural network achieve better accuracy for this application VGG16 or VGG19 or ResNet50.	Unbalanced and limited data size and overfitting are challenging the author should take it in considerations.	The study and the result are significant. However initialize the weight randomly in the pre-train model may increase the computational complexity.	VGG has the best result accuracy =92.6%. AUC=95.95%
[21]	2018	CNN	Breast cancer	The author proposed transfer learning with CNN for multi-class breast cancer classification. The proposed approach goals to categorize the breast tumors in benign or malignant, moreover to predict the subclass of the tumors like Lobular carcinoma, Fibroadenoma, etc.	The performance can be enhancing while using larger dataset.	Using the pretraining model DenseNet decrease the complexity of the training process and solve the overfitting problem	Accuracy = 95.4% by using DenseNet CNN model.
[22]	2018	CNN	Twelve types of cancers.	The gene expression samples of a specific type of cancer might be rare. The author proposed a new multi-task deep learning method (MTDL) to classify multiple cancers concurrently and improve the performance of the classification of every cancer by utilizing the knowledge throughout shared layers.	Lack of tissue samples during experiments, dimensionality problem of feature spaces.	The experimental outcomes reveal that MTDL significantly enhances the performance of diagnosing each sort of cancer when it learns from the12 types of cancers together.	The author makes comparison between MTDL, DNN and Sparse Auto-encoder. MTDL has superior result for the 12 datasets.
[23]	2018	PSO-CNN		The author used swarm-based hyperparameter optimizer (PSO) and multi view to optimaize the parameters of CNN, hyperparameters of the model to be optimized are the biases and the weights of every layer, to classifies heart disease and breast cancer.	It suggested to performe more validations because the author compared his resullts with only naive version of CNN. Also, diverse types of feature fusion techniques should be investigated in the multi-view CNN in collaboration with swarm-based	Using PSO for feature selection can guarantee cover all search space and select the more informatics features that can support classification process later.	---

					hyperparameter optimization and various metaheuristics algorithms in order to enhance the validation process.		
[24]	2019	CNN		The author suggests a deep transfer learning model to differentiate indolent from clinically significant prostate cancer lesions and make a comparison between deep transfer learning model with a deep learning without transfer learning.	The model needs to be validated by using large datasets to additional evaluate it is clinical service.	The Deep Transfer Learning model have higher accuracy comparison with the deep learning model without using transfer learning.	P=0.89
[25]	2019	CNN	Liver cancer	The author proposed and evaluate a novel extended (2-D) to three-dimensional (3-D) convolutional neural network (CNN) intended for tumor classification in medical imaging and employed for distinguishing between primary and metastatic liver tumors from diffusion-weighted MRI (DW-MRI) data.	Cancer classification working immediately on whole 3-D tomographic data without any data mining techniques such as preprocessing, but this effect on the classification performance.	The auther used recall and precision to calculate the accuracy of Liver Tumor classification which is 83% without preprocessing.	83% Vs 69.6 for 3D and 65.2% fro 2D.
[12]	2019	CNN	Breast cancer	The author developed a (DCNN) for the digital breast tomosynthesis (DBT) classification by using a multi-stage transfer learning method, where a pre-trained CNN is first fine-tuned on non-medical images to a related task in medical imaging domain before moving to fine-tuned to the target task.	It needs to verify the generalization ability of the trained CNN with independent unknown problems.	The author used transfer deep learning on digital breast tomosynthesis (DBT) which is consider new kind of data.	AUC=0.91±0.03
[26]	2019	CNN	Breast cancer	The author developed a deep learning framework for automatic a histochemical score (H-Score) evaluation for breast cancer tissue microarray.	It needs more validation and test to prove that H-score has a major effect on prediction accuracy.	The author observed that the main variances between pathologists and machine predictions happened in images that will have a high H-Score value.	---
[11]	2019	CNN	prostate cancer	The author develops a new multi-parametric magnetic resonance transfer learning (MPTL) method to automatically categorize prostate cancer. First build a deep convolutional neural network (CNN) with three-part architectures, which transfer the pre-trained model (ImageNet) to measure features from	The result may be validated on different types of datasets and diseases.	The outcomes have shown the possible benefits of transfer learning from natural images to the medical field when the training datasets are limited.	Accuracy=86.92%

				multiparametric MRI images.			
[27]	2019	CNN	Breast cancer	The author used a deep learning framework for breast cancer detection and classification by using transfer learning for feature extractions and transfers the knowledge. The author used three CNN architectures (VGGNet, GoogLeNet, and ResNet) which are joined using the idea of transfer learning for enhancing the accuracy of classification. The three CNN architectures trained on imagNet dataset to transfer the knowledge to the new dataset.	If the features engineer used along with CNN features used, it will improve the classification accuracy.	The suggested framework delivers great results concerning accuracy without training from scratch which enhances classification efficienc.	Accuracy =97.67%
[28]	2019	CNN-ensemble	Cancer using methylation data.	The author proposed a convolutional neural network (CNN) based multi-model ensemble method for cancer prediction using DNA methylation data.	The study used the t-test for feature reduction can enhance with another feature selection method. The author used stochastic gradient descent(SGD) which prone to stuck in minima, also the validation not reliable because no comparison with others works.	DNA methylation data consider big challenge due to it is dimensionality and low size of sample, but the author solve this problem by work as stack. And the auther used CNN based ensemble method which can learn the relationship between the classifiers automatically and achieve better prediction.	Accuracy = CNN based ensemble for LUAD (99.39%) for LIHC (98.83%) for KIRC (99.58%)
[29]	2020	CNN	Many types of cancers	The author used convolutional neural network with laplacian score as hybrid method named (LS-CNN), laplacian score used for feature selection. The method used to classify cancer's by using 10 microarray datasets.	There are no criteria for choosing the optimal number of features after ranking.	The author proves that CNN can be used with CSV extension files, not as known that it is designated for unstructured data, also structure data such as CSV file can convert to the 2D matrix and used as input to the CNN.	The average accuracy ranges from (90 to 100) for all datasets.
[30]	2016	DNN	Breast cancer	The author used deep neural networks (DNNs), a cutting-edge machine learning method, to the classification of compounds in chemical mechanisms of action (MOAs). The compound procedure done by image-based profiling combined with feature reduction methods such as PCA or factor analysis. Then used deep transfer learning (DTL) to accelerate the training process and improve the accuracy.	high-throughput and high-dimensional data and computational speed, are the main challenges and limitations faced the author during this research.	The result is significant, 30% speedup and a 2% accuracy improvement by using transfer learning.	Accuracy= 77%

[31]	2018	DNN	cancers	The author proposed the elephant search (ESA) based optimization, to select the optimal gene expressions from the high dimensionality of microarray data. Moreover, the author used Firefly Search (FFS) to understand the efficiency of the Elephant search technique in the feature selection procedure.	The method should be checked in a dataset that contains a lot number of samples and high dimensionality in the future. Nevertheless, obtaining the optimal feature or gene from gene expression profiles is considering a difficulty for receiving high accuracy in prediction.	The number of genes after feature selection are reduce significantly, at the same time the accuracy increase in parallel.	10 datasets
[10]	2018	DNN	cancer	The author used transfer learning process for classification of cancer, which uses normalization techniques and feature selection in combination with stacked sparse auto-encoders on gene expression data. To decrease the complexity of computational of learning the (DNN) for gene expression data.	The problem of optimal network architectures could be address and use of other feature rank techniques.	The result statistically outperforms the state-of-the-art molecular cancer classification approaches considered in author research.	many
[32]	2018	DNN	Stomach Lung Breast	Majority voting in ensemble learning is too simple to identify complex information from various classifiers, and only studies the linear relationships between classifiers. Therefore, the author proposed deep learning with an ensemble method, by using DNN instead of majority voting to classify cancer.	The DESeq technique can reduce the dimensionality and select the most informatics features, but it cannot handle the outlier problem.	The author used the feature selection technique in the data preprocessing phase, which considerably decreases the running time and enhances the accuracy of prediction at the same time.	Accuracy for LUAD =0.994% For STAD =0.995% For BRCA=0.996%
[33]	2020	DNN	Stomach Lung Breast	The author used grey wolf algorithm for feature extraction and DNN for cancer classification	Another meta-heuristic method can achieve better results than the grey wolf.	The author proves that DNN with 15-layers can be better or at the same level of ability as machine learning for microarray classification.	Accuracy for LUAD =99.89% For STAD =99.37% For BRCA=99.19%
[34]	2020	DNN	Many types of cancer	The author develops a deepforward method to classify microarray cancer datasets	The author split the data 60% for training and 40% for testing, but this structure of division not generalized for all datasets. Moreover the author should increase the	The author used a binary class dataset, but the multi-class microarray dataset not examined by the researcher.	The average accuracy ranges from (95 to 100) for all datasets.

					percentage of training due to the low number of samples.		
[35]	2017	MLP	11_Tumors	Microarray data for cancer normally includes a small number of instances and has a high number of gene expression levels as features, nevertheless, it makes the classification pretty challenging. The author used deep learning techniques based on a multilayer perceptron (MLP), to prove that MLP can be used for cancer classification.	The result not validate, because there is no comparison with other results. Moreover, the architecture of MLP not clears in the research.	The author shows that MLP could learn well, and also, can classify the various samples with high performance, and finally, a high accuracy is achieved in the task of predicting the various classes.	MSE =0.15 R^2=0.9864
[36]	2018	CNN-SVM	Breast cancer	CNNs usually suffer from tuning a vast number of parameters that carry a lot of challenges and complexity to the network. Furthermore, the first value of the weights and bias of CNN is considering another challenge that requires to be handling correctly. The author proposed transfer learning (TL) and deep feature extraction methods. Which adapt a pretrained CNN model used for cancer detection. Vgg16 and AlexNet models are considered for feature extraction and fine-tuning. The selected features are then classified by support vector machines (SVM).	The author did not used data augmentation with transfer learning. Moreover, using the machine learning classifier (SVM) for classification consider step back.	The contribution in the study is significant, the complextiy was reduced by fine-tuning the parameters and given initial values to the weights and bias in the CNN by using transfere learning also reduce the complexity and training time.	Accuracy=91.37%
[37]	2018	Autoencoder neural network	breast cancer	The author presents an unsupervised feature learning framework by incorporating PCA algorithm and auto-encoder neural network to find dissimilar features from gene expression profiles. Then the ensemble classifier based on the AdaBoost algorithm (PCA-AE-Ada) was built to predict clinical results in breast cancer.	Although the model has produced excellent results, the generalization ability needs to be more enhanced with more public datasets. Classifiers based on selected gene signatures are so weak and customized for the study.	Experimental outcomes show that the performance of the proposed method which using deep learning techniques is better than others.	Acc1=0.65 Acc2=0.72 Acc3=0.77 Acc4=0.75 Acc5=0.85
[38]	2019	Stack autoencoder		The author proposed a model automatic diagnosis of cervical cancer. by using stacked autoencoder and SoftMax classification. stacked autoencoder has been used to reduce the dimensionality and SoftMax layer was used for classification.	The training time of the method, which used in the study is consider the worse compared with other methods because of a lot of consuming time lost during the dimensionality reduction	dimensionality reduction curse, yet, this challenge can be solved easily with deep learning.	Mean Acc=96.95 Acc=97.25 Acc=96.53 Acc=96.60

					process.		
[39]	2018	stacked sparse autoencoder	Cancers	The author proposed a semi-supervised deep learning technique, the stacked sparse auto-encoder (SSAE) based classification, for cancer prediction using RNA-seq data, by using greedy layer-wise pre-training and sparsity.	Semi-supervised data not available at the web site and the author not determine how the semi-supervised data collected. And also, the result need validation.	Empirical outcomes show that the proposed method using SSAE technique better performance than others.	Accuracy for LUAD =99.89% For STAD =98.15% For BRCA=96.23%
[40]	2020	LDA-AE	Breast cancer	The author proposed a combination between Auto Encoder (AE) Neural Network and Linear Discriminant Analysis (LDA) by ensemble learning to classify microarray breast cancer dataset.	The author should use more than one public dataset for generalization issues and should measure the stability by using the Jaccard index for example for more validation.	Using the deep learning in the claccification stage can discover the non-linear relationship between features.	Accuracy= 98.27%
[41]	2020	SAE	Breast cancer	The author proposed an approach by using KNN and stack autoencoder for cancer diagnosis based on a microarray dataset	For more validation, the author should use more than one dataset rather than one. Moreover, the author reduces the number of features to 100 without discussing the criteria for choosing this number.	The author used various classifiers within the softmax layer in the SAE to enhance the performance of the model.	Accuracy=91.24
[42]	2020	CNN	Central Nervous System. Leukemia Ovarian	The author proposed a hybrid method for feature selection by using Relief and SAE and for microarray cancer classification, the author utilized CNN and support vector machine.	The number of selected features needs more clarification and the hyperparameters of CNN should be obtained by using intelligent optimization.	Using deep learning in both stages of feature selection and classification gives more reliable results especially in medical data.	Acc1=98.6% Acc2=99.86% Acc3=83.95%