MJFAS MALAYSIAN JOURNAL OF FUNDAMENTAL AND APPLIED SCIENCES
PRINT ISSN: 2289-5981 | ONLINE ISSN: 2289-599X

RESEARCH ARTICLE

The usage of deep learning algorithm in medical diagnostic of breast cancer

Kevin Nathanael Ramanto, Arli Aditya Parikesit*

Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta, Indonesia

* *Corresponding author:* arli.parikesit@i3l.ac.id

Article history
Received 2 July 2018
Revised 24 August 2018
Accepted 4 September 2018
Published Online 18 April 2019

Abstract

Diagnosis is a crucial step to identify the disease that experienced by the patient. Diagnosis includes information gathering, integration, and interpretation. However, diagnosis process is not an easy task. Diagnostic accuracy is depending on the experience and cognitive ability of diagnosticians. The new algorithm called deep learning that is developed by simulating the human visual mechanism has been implemented in medical diagnostics. One of the diseases that can be diagnosed by using deep learning algorithm is the breast cancer. Several studies showed that deep learning algorithm can be used for detecting and classifying lesions, detecting mitosis, and predicting specific gene status. In this review article, 16 research journals were reviewed and discussed. The limitations of each algorithm are provided. All of the journals showed that deep learning algorithm has high diagnostics accuracy in assisting the professional diagnosticians to determine diagnosis outcome accordingly.

***Keywords*:** Diagnosis, deep learning algorithm, breast cancer, medical

INTRODUCTION

Health diagnosis is a crucial step to identify the disease that affected the patient. The diagnostic process proceeds as follows: first experienced a health problem. The patient usually the first person to consider his or her symptoms and may choose to go to the Medical Centres. Once patient seeks health care for his or her symptoms, there is an iterative process of information gathering, information integration and interpretation, and determining a working diagnosis (Balogh, 2015); all of the processes are shown in Fig. 1. All the information regarding the disease can be collected in several ways such as performing clinical history and interview, physical exam, performing diagnostic testing and consulting with other clinicians (Balogh, 2015). The process of gathering information, integration, and interpretation involves hypothesis generation and updating prior probabilities as more information learned.

Throughout the diagnostic process, there is an ongoing assessment of whether the sufficient information has been collected. If the diagnostic team members are not satisfied that the necessary information has been collected or the information is not consistent with the diagnosis, the process of information gathering, integration, and interpretation continues (Balogh, 2015). When a diagnostic team has an accurate and timely explanation of the patient's health problem, they will communicate to the patient regarding the diagnosis.

Diagnosis accuracy is often depending on the experience and cognitive ability of diagnosticians. There are usually many different judgements regarding the information in each of the diagnosticians. One of the examples is interpreting the image of an ultrasound, in which different diagnostician may come with a different diagnosis (Li *et al*., 2018). An inaccurate diagnosis can result in harmful, inappropriate and unnecessary care. Therefore, giving a proper diagnostic to the patient is a challenging task.

Over the recent years, a new algorithm has been developed by simulating the human visual mechanism to help the diagnosis process (Memisevic, 2015). Computer vision is known to has the advantages of high detection speed and low cost. This technology is often used in the area of rapid intelligent image processing, such as image classification, object detection and object retrieval (Faust *et al*., 2018). This new algorithm is later known as deep learning.

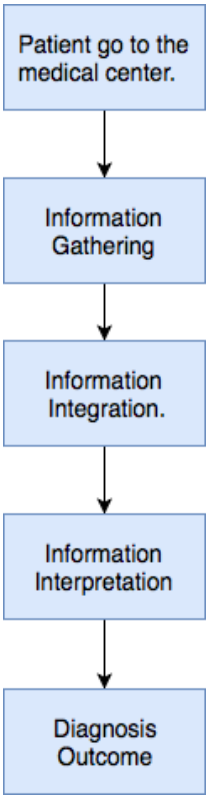

Fig. 1 Steps of diagnosis.

Deep learning can be defined as a subfield of machine learning with algorithm, inspired by the structure and function of the brain called artificial neural networks (Deng, 2013). Deep learning is considered as a new area of machine learning research, which has been introduced with the objective of moving the Machine learning closer to one of its origin goals: the Artificial Intelligence.

Deep learning belongs to the class of machine learning methods. It is a special form of representation-based learning, where a network learns and constructs features from each of successive layers in the Artificial Neural Network (ANN) structure (Oliver *et al*., 2018). The ANN algorithm models have the same function as biological brain. The model is created with the structure of input, hidden, and output layers; as shown in Fig. 2. Every neuron is connected to each neuron in the next layers through a connection link known as weights. A nerve cell is made up of axon (output), dendrites (input), a node (soma), nucleus (activation function), and synapses (weights). The activation function in the artificial neuron acts as the nucleus in a biological neuron whereas the input signals and it respective weights model the dendrites and synapses, respectively (Oliver *et al*., 2018).

Nevertheless, the ANN structure is receptive to translation and shift deviation that may adversely affect the classification performance. In order to eliminate these shortcomings, an extended version of ANN called as the Convolutional Neural Network (CNN) is developed. The CNN architecture ensures translation and shift invariance. CNN network structure is a feed forward network, which comprises of convolution, pooling, and fully connected layers (Deng, 2013). Convolution layer is where the input sample is convolved with a kernel in this layer. Pooling layer is a down sampling layer where the pooling operation is employed to reduce the spatial dimension of the input sample, while retraining the significant information (Oliver *et al*., 2018). Fully connected layer is fully connected, which signifies that each neuron in the previous layer is connected to all the neurons in the current layer.

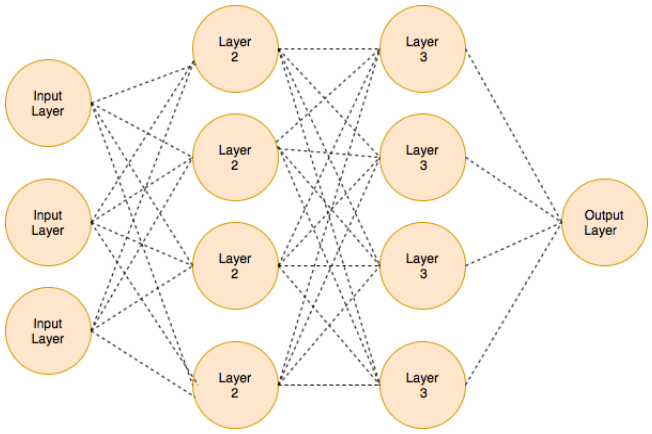

Fig. 2 ANN structure overview.

Convolution can be calculated by using equation (1), where *a* represents an input and *w* represents a weight. This calculation can determine the significant of output given in each layer.

$$a1.w1 + a2.w2 + \ldots\ldots = Output \quad (1)$$

A leaky rectifier linear unit is used as an activation function after the convolution layer. The purpose is to map the output to the input set and introduce non-linearity as well as sparsity to the network (Oliver *et al*., 2018). The CNN is trained with backpropagation and the hyperparameters, which may be tuned for optimal training performance.

Beside ANN and CNN, there are other deep learning algorithms that have been developed by researchers. Some of the deep learning algorithms are designed to diagnose specific diseases. One of the diseases that can be diagnosed with deep learning algorithm is breast cancer. The treatment of breast cancer usually depends on: type of breast cancer, stage of the cancer, sensitivity of hormones, etc. Therefore, an accurate diagnostic is crucial to determine an effective treatment to breast cancer patient.

Diagnostic test in breast cancer patient may include imaging and laboratory test, such as a mammogram, an ultrasound scan and an MRI scan (WHO, 2015). These results will be studied by diagnostician to obtain all the information regarding breast cancer. In recent years, several studies have used deep learning algorithm to diagnose breast cancer in different ways, such as: detecting and classifying lesions, detecting mitosis, predicting specific gene status, etc. These studies indicated that deep learning algorithm has higher diagnostic accuracy and can assist the professional diagnosticians to determine diagnosis outcome accordingly.

The main goal of this review article is to deliver a panoramic overview of how deep learning algorithm has helped to improve the accuracy of diagnosis in breast cancer. Throughout this review article, several topics regarding deep learning application in diagnosis process will be discussed.

METHODS

16 Journals article has been collected from PubMed and google scholar. The searching of journals article was conducted from 30th April 2018 until 20th June 2018 with "Deep learning in breast cancer diagnostic" as the keywords search. All journals that have been selected were published in 2016 - 2018. Each of these journals implemented the deep learning algorithm in medical diagnostic, especially in breast cancer diagnostics. From 16 journals article that have been found, only 10 of them that used deep learning algorithms to diagnose breast cancer. Table 1 shows the review summary of 16 journals. Each of algorithms will be discussed in this review article.

RESULTS AND DISCUSSION

FCNN used for the detection and quantification of intraretinal cystoid fluid in multivendor optical coherence tomography.

Fully Convolutional Neural Network (FCNN) is a deep learning algorithm that designed to diagnose age-related macular degeneration (AMD). This algorithm is created by Venhuizen *et al.* AMD is a complex multi-factorial retinal disease where genetic and environment play a big role in the development of disease. Since the disease is affected by a wide array of intrinsic and extrinsic factor and protective factors, a universal one-fits all treatments are arguably not the optimal solution.

Fluid accumulation is best visualized on spectral domain optical coherence tomography (SD-OCT) imaging. SD-OCT provides a non-invasive, high resolution, three-dimensional visualization of the retina, where fluid is visible as a hypo reflective area (Venhuizen *et al*., 2018). A clear distinction can be made on SD-OCT imaging between intraretinal fluid (IRF) and subretinal fluid (SRF) based on the relative location in the retina, inside the sensory retina or below it.

FCNN can be used for the detection, segmentation and quantification of intraretinal cystoid fluid (IRC) in SD-OCT volumes from AMD patient. Venhuizen *et al.* conducted an experiment involving 221 SD-OCT volumes from 151 patients with varying presence of IRC by using FCNN. The deep learning algorithm could automatically detect, segment and quantify IRC in the entire SD-OCT volume. The algorithm produces IRC volume segmentation by processing every B-scan in an SD-OCT volume individually (Venhuizen *et al*., 2018).

In the end of his experiment, Venhuizen *et al* found out that the Pearson correlation coefficient is was equal to 0.936 between IRC and annotated ground truth volume. This indicated that the FCNN could serve as a reliable tool that would produce a fluid quantification that was similar to human performance in a fraction of time.

Faster R-CNN used to detect cancer

Faster R-CNN is the development of R-CNN, where Faster R-CNN achieves state of the art performance on pattern analysis, statistical analysis, statistical modelling and computational visual object classes

(Hailiang *et al*., 2018). Faster R-CNN is basically based on a convolutional neural network with additional components for detecting, localizing and classifying object as an image. However, Faster R-CNN needs an extra step such as selective search or Edge boxes to generate object proposal.

The research that conducted by Hailiang *et al.* (2018) has improved the Faster R-CNN for ultrasonic image detection. This algorithm was used to detect papillary thyroid carcinoma from 4670 ultrasound images. Improved Faster R-CNN used the conv3 layer and conv5 layer of ZF that were concatenated and normalized. Other than that, spatial constrained was also added before the output layer. In the end, both truth positive rate (TPR) and true negative rate (TNR) were calculated and compared with the original. TPR and TNR were increased 2.7% and 4.7% from the original. Not only that, about 93.5% papillary thyroid carcinoma region was detected automatically. This indicated that the improved Faster R-CNN could identify more correct samples, especially the negative sample.

Another experiment that conducted by Tóth *et al* has used the basic Faster R-CNN to detect and classify lesion in breast cancer.

The dataset was contained with 2620 digitized film-screen screening mammography exams, with pixel level ground truth annotation of lesion. In this experiment, VGG16 network was used, which was a 16-layer deep CNN. The final layer could detect 2 kinds of object in the images: benign or malignant lesions. Overall, Faster R-CNN could detect 90% of the malignant lesions in the dataset.

Epistasis and Heterogeneity Analysis with Deep Learning

Complex disease is caused by the defects in multiple genes and do not has a clear-cut pattern of inheritance (Pevsner, 2015). Not only genetic, environment also plays a big role in complex disease (Pevsner, 2015). Therefore, understanding the genetic mechanism of complex disease is remained as a big challenge. Existing methods usually neglect the heterogeneity phenomenon of complex diseases, resulting in lack of power or low reproducibility (Li *et al*., 2018).

Table 1 Review Summary.

Author	Sample	Diagnosis	Deep Learning Algorithm	Application	Result
Venhuizen *et al*., 2018	221 SD-OCT* volumes from 151 patients	Age-related macular degeneration	FCNN*	Automatic segmentation and quantification of intraretinal cystoid fluid in SD-OCT*	The Pearson correlation coefficient: 0.936 between intraretinal cystoid fluid and annotated ground truth
Hailiang *et al.*, 2018	4670 ultrasound images from 300 cases	Thyroi Papillary carcinoma	Modified Faster R-CNN	Detecting thyroid papillary cancer in ultrasound images	True positive rate was increased 2.7% and true negative rate was increased 4.7%, while 93.5% of papillary thyroid regions were detected automatically
Li *et al.,* 2018	10 pure and 10 heterogenous datasets consist of 100 SNPs	Complex diseases	DPEH*	Addressing epistasis and heterogeneity	Accuracy: 81%
Tóth *et al.*, 2018	2620 digitised film-screen from mammography exams	Breast Cancer	Faster R-CNN	Detecting and classifying lesions in mammogram	90% of the malignant lesions were detected in the dataset
Becker *et al.*, 2017	3228 mammography images from 143 patients	Breast Cancer	ANN*	Evaluating the accuracy of ANN in detecting breast cancer	Accuracy: 82%
Bejnordi *et al.*, 2017	399 whole-slide images	Breast Cancer	32 different algorithms	Evaluating the performance of different deep learning algorithms at detecting lymph node metastases	7 deep learning algorithms showed greater discrimination
Daniel *et al.,* 2017	71896 images from 14880 patients	Diabetic retinopathy, possible glaucoma, and age-related macular degeneration	DLS*	Evaluating the performance of DLS	Sensitivity :90.5 – 100% Specificity: 87.2 - 91.6%
Wang *et al.,* 2016	Camelyon 16 dataset consists of 400 slide images	Breast cancer	CNN*	Identifying metastatic breast cancer	Area under the curve :0.9250
Albarqouni et al., 2016	MICCAI-AMIDA13 dataset consists of histology images from 23 patients	Breast cancer	CNN	Detecting mitosis in histology images	Overall improvement score was 22.5% $\pm$ 6.8

Table 1 Review Summary (continue).

Author	Sample	Diagnosis	Deep Learning Algorithm	Application	Result
Cheng *et al.*, 2016	275 benigns and 245 malignant lesions from 520 patients Randomly selecting 700 benigns and 700 malignant nodules	Breast cancer and lung cancer	SDAE*	Differentiating breast ultrasound lesion and lung CT nodules.	Accuracy: 87.4% ± 1.9
Alakwaa *et al.*, 2017	Metabolomics data consists of 271 breast cancer samples	Breast cancer	ANN	Predicting estrogen receptor status in breast cancer metabolomics data	Area under the curve was 0.93 in classifying ER+/ER-
Han *et al.*, 2017	7909 histopathological images from 82 patients	Breast cancer	CSDCNN*	Multi-class breast cancer classification	Accuracy: 93.2%
Vandenberghe *et al.*, 2017	74 whole-slide images of breast tumour resection	Breast cancer	CNN	Predicting HER2 status in breast cancer	Accuracy: 83%
Araújo *et al.*, 2017	269 histology images	Breast cancer	CNN	Classifying breast cancer histology images	Accuracy for four classes: 77.8% Accuracy for carcinoma/ non-carcinoma: 83.3% Sensitivity: 95.6%
Cang *et al.*, 2017	223 mutation instances of membrane protein	Molecular simulation	multi-task multichannel topological convolutional neural network	Biomolecular property predictions	The Pearson correlation was improved 9.6%
Torng *et al.*, 2017	3890 protein families	Molecular simulation	3DCNN	Structure-based protein analysis	Accuracy: 42 – 67%

SD-OCT: spectral domain optical coherence tomography
FCNN: fully convolutional neural network
DPEH: deep learning method for epistasis heterogeneity analysis
ANN: artificial neural networks
DLS: deep learning system
CNN: convolutional neural network
SDAE: stacked denoising auto-encoder
CSDCNN: class structure-based deep convolutional neural network

Addressing heterogeneity when detecting epistatic single nucleotide polymorphism (SNPs) can enhance the power of association studies and improve prediction performance of complex diseases diagnosis.

Li *et al.* (2018) created a deep learning method for epistasis and heterogeneity analysis (DPEH). DPEH will detect epistasis and heterogeneity by using three stage frameworks, which are epistasis detection, clustering, and prediction. In this experiment, a pure and heterogenous dataset was used to evaluate DPEH. The accuracy of DPEH reached 81% in the evaluation. This indicated that DPEH was accurate to detect epistasis and heterogeneity.

Breast cancer detection

Deep learning algorithm is used to analyse any kind of imaging test result to diagnose breast cancer, such as: mammogram, ultrasound, CT scan, histological images, etc. Deep learning algorithm is created to predict and classify the breast cancer according to the samples. Most of the researchers use ANN or CNN and some of them may create a new algorithm.

A mammogram is an x-ray picture of the breast and used to check for breast cancer in women who have no signs or symptoms of the breast cancer (Breast Cancer, 2017). Becker *et al* conducted the cohort study that was implemented the ANN to detect breast cancer in mammogram images. There were 3228 mammogram images that obtained from 143 breast cancer patients. The samples were categorized based on aged, breast density, etc. The result of this experiment showed that the diagnostic accuracy of ANN reached 82%. This result was pretty high compared to experienced radiologist that only has 79% of accuracy. Therefore, the deep learning algorithm was designed for generic image analysis that could be trained to detect breast cancer on mammography data.

Ultrasound imaging uses sound waves to produce pictures of the patient body. It can assess the morphology, orientation, internal structure and margins of lesions from multiple planes with am high resolution both in predominantly fatty breast and in a dense glandular

structure (Guo *et al*., 2018). Ultrasound has been used to classify benign and malign lesion in the breast cancer. Cheng *et al* proposed a new deep learning algorithm called stacked denoising auto-encoder (SDAE) that would help to classify benign and malign in ultrasound images. The SDAE could automatically discover the diverse representative patterns from the data with the intrinsic data reconstruction mechanism. Comparing to other algorithms, SDAE has the highest accuracy, which was 82.4 $\pm$ 4.5 % (Cheng *et al*., 2016).

Unlike mammogram and ultrasound, histopathological image or histology image is obtained by biopsy. This procedure will take out a piece of the tissue from the human body. Different structures of the tissues are coloured with different stains (Asawathy *et al*., 2017). Then, this image will be used to study the growth of tumours in various types of cancer, such as breast cancer. Unfortunately, the manual examination of histopathological images may require intense workload of highly pathologist. Therefore, deep learning algorithm has been implemented to diagnose breast cancer in the histopathological images. Araújo *et al.* (2017) classified histopathological images into 4 classes which were normal, benign, in *situ*, and invasive. This criterion was done by using CNN to increase the sensitivity of determined carcinoma classes (in *situ* and invasive) and non-carcinoma classes (normal and benign). The classification was obtained by different patch labels which were majority voting, maximum probabilities, and sum of probabilities. Majority voting was where the image label was selected as the most common patch label; maximum probability was where the patch with higher class probability decided the image label; sum of probabilities is was where the patch class probabilities sum and the class with the largest value were assigned. From 269 histopathological images, CNN could obtain 77.8% accuracy for class and 83.3% in carcinoma/non-carcinoma. The evaluation suggested that the CNN algorithm has 95.6% sensitivity to classify all the images.

Similar research was done by Han *et al.* in 2017. Instead of using basic CNN or ANN, they proposed a new deep learning algorithm called class structure-based deep convolutional neural network (CSDCNN). This algorithm was used to provide an accurate and reliable solution for breast cancer multi-classification through the histopathological images. The CSDCNN adopted the end-to-end training manner that could automatically learn semantic and discriminative hierarchical from low level to high level. 7909 images were classified into 8 classes, which were ductal carcinoma, lobular carcinoma, adenosis, fibroadenoma, phyllodes tumor, tubular tumor, tubular adenoma, muscinous carcinoma, and papillary carcinoma. CSDCNN achieved 93.2% of accuracy, which was higher than ordinary CNN algorithm that used by Araújo *et al*. This indicated that CSDCNN could provide an efficient algorithm than CNN in classifying breast cancer.

Deep learning algorithm can help to identify the metastatic breast cancer which is classified as stage 4 breast cancer. Metastatic breast cancer is breast cancer that has spreaded to other parts of the body (Redig *et al*., 2013). Breast cancer commonly spreads to the lungs, liver, bones and brain. In order to diagnose metastatic breast, diagnostician may evaluate breast sentinel lymph nodes under the microscope. However, CNN algorithm has provided a better method to identify cancer metastases from whole slide images of breast sentinel breast lymph nodes. This experiment was done by Wang *et al*. in which Camelyon 16 dataset consisted of 400 slide images were used as sample. This method obtained an area under the receiver operating curve of 0.925 for the task tumor localization task. This result suggested that CNN could reduce approximately 85 percent of human error rate.

Mitosis can also be calculated by counting mitotic figures in histopathological images. Mitotic figures appear as hyperchromatic objects without nuclear membrane and hairy extension of nuclear material (Veta *et al*., 2016). In the recent years, a pathology workflow that implemented deep learning algorithm has been developed. Research that was conducted by Albarqouni *et al.* (2016) used histopathological images from at least 23 breast cancer patients. This research showed that CNN algorithm achieved 22.5% $\mp$ 6.8 of improvement score from the original workflow.

Deep learning has also been applied to study the gene expression data in breast cancer. Therefore, determining the status of a certain gene expression will be possible. Alakwaa *et al.* (2017) and Vandenberghe *et al.* (2017) conducted the different experiments to predict a certain gene expression status within the breast cancer. Alakwaa *et al.* (2017) used ANN algorithm to predict the estrogen receptor (ER) status in metabolomics data and resulted in 0.93 area under the curve in classifying ER+/ER-. In the other hand, Vandenberghe *et al* used CNN to predict human epidermal growth factor receptor 2 (HER2) status and resulted in 83% of accuracy. Both of the experiments indicated that deep learning algorithm has predicted a certain gene expression may open more understanding in regarding how breast cancer could develop in the first place.

Deep learning for molecular simulation

The goal of molecular biology research is to determine the functions of genes and their products, allow them to be linked into pathways and provide a detailed understanding on how biological system works (Andrianantoandro, 2006). Some biological researches have provided the molecular simulation to understand protein structure. Determination on the structural and functional roles of individual amino acids within protein may provide information to alter the targeted proteins. Torng *et al.* (2017) proposed 3D convolutional neural network (3DCNN) to do structured based protein analysis. The algorithm was used to analyse local protein microenvironments surrounding 20 amino acids and predict which the amino acid was most compatible with environments within protein structure. 3DCNN achieved 42 -67% of accuracy.

Another molecular simulation algorithm was proposed by Cang *et al.* in 2017. This algorithm was called as multi-task multichannel topological convolutional neural network (MM-TCNN). Unlike 3DCNN, MM-TCNN was focused on biomolecular properties prediction, such as protein-ligand binding affinities, folding free energy changes upon mutation, and membrane protein mutation. This experiment showed that deep learning algorithm has improved the predictions than previous methods.

Deep learning evaluation

Deep learning has high sensitivity and diagnostic accuracy. Therefore, in the recent years, several deep learning algorithms are developed in medical field. Since there are several deep learning algorithms that used to diagnose a certain disease, the evaluation is needed to prove the performance of deep learning algorithm.

In the experiment that conducted by Ting *et al.*, evaluation on deep learning system (DLS) was conducted for diagnosing diabetic retinopathy, possible glaucoma, and age-related macular degeneration. This experiment involved 71896 images from 14880 patients. The result showed the DLs has a sensitivity of 90.5% and specificity 91.6% for diabetic retinopathy. This implied that the DLS has a high sensitivity and specificity for identifying diabetic retinopathy and related eye diseases using retinal images from multi-ethnic populations with diabetes.

Another research that conducted by Bejnordi *et al.* (2017), evaluation on 32 different deep learning algorithms that used to detect lymph node metastases in breast cancer was conducted. From 32 algorithms, 7 of them showed a greater discrimination than a panel of 11 pathologists in a simulated time constrained diagnostic setting. These findings suggested the potential utility of deep learning algorithms for pathological diagnosis but assessment setting was further required.

The advantages and disadvantages of deep learning

Deep learning algorithms have demonstrated advantages in medical diagnostic. However, these methods have some limitations that should be considered. Each of the algorithms has different advantages and disadvantages as shown in table 2. The most common problems that found in deep learning algorithm are included the "black box" problem, the need for large training data sets, and the high computational cost of training (Mamoshina *et al*., 2016).

The "black box" problem is one of the major limitations of deep learning that related to quality control and interpretation (Mamoshina *et al*., 2016). Some of the deep learning algorithms are "black boxes"

that learnt by simple associations and co-occurrence. They have limited means with which to interpret the representation.

Table 2 Advantages and disadvantages of deep learning algorithm.

Deep Learning Algorithm	Advantages	Disadvantages
FCNN	Provide a large contextual window. Large modelling capacity.	Large objects cannot be efficiently detected because of the confined receptive field of FCNN. Only evaluated in a dataset containing intraretinal cystoid fluid as the result of Age-related macular degeneration.
CNN	Operate directly on a patch of images centered on the abnormal tissue. Highly parallelizable algorithm.	Designed for 2-D images. Do not model spatial dependencies.
ANN	Require less formal dataset training to develop. Flexible with respect to incomplete, missing and noisy data.	ANN structure is receptive to translation and shift deviation, which affected classification performance. Poor transparency which operated as “black blocks”
Faster R-CNN	Identify more correct samples, especially negative samples. Detect and localize objects on the image, regardless of the class of the object.	Cannot capture more local texture of object due to the Regions of Interesting (ROI) pooling mechanism. Difficult to extract local textures from low resolution images.
DLS	High sensitivity and specificity. Usually trained to detect related eye diseases.	Use multiple levels of representation to analyze each sample without showing the actual lesions. “Black box” issues may have an effect on physician’s acceptance for clinical used.
SDAE	Automatically discover the diverse representative patterns from the data with intrinsic data reconstruction mechanism. Address the issues of high variation in either shape or appearance of samples.	Circumvent the potential inaccurate image processing results can lead to unreliable features in the framework. Require big memory space.
CSDCNN	Automatically learn semantic and discriminative hierarchical features from low-level to high-level. Account the relation of features space among intra-class and inter-class.	The features space distance of samples from the same class may be larger than the samples from different classes. The intra-class variance is not preserved.
3DCNN	Reduce the tendency of over-fitting. Comprised of local spatial features	The grid voxel system is not rotationally invariant. The atomic details do not provide significant additional information.
DPEH	Classify samples into more precise categories. Effectively address the heterogeneity of complex disease.	The performance level is low when the sample size is small or epistatic pattern is simple. Require big memory space.
MM-TCNN	Exploit the relations among various structure-function predictions. Enhance the prediction for problems with small and noisy training data.	There is no obvious transferable property of the convolution filters along the convolution dimension. Do not use auxiliary features.

Therefore, most of the researchers use deep learning approaches as "black box" without the ability to explain why it provides good results or apply model modifications in the case of misclassification issues (Ravi *et al*., 2017). These make some of deep learning algorithms are unable to uncover complex casual and structural relationship in biology without human input.

The requirement of large training data set becomes another challenge in deep learning algorithm. Large training data set is required to train a reliable and effective model (Ravi *et al*., 2017). However, the availability of disease specific data is often limited. A small data set may arise an overfitting, which occurred when the number of parameters in the network is proportional to the total number of samples in training data set (Mamoshina *et al*., 2016). The network will be able to memorize training examples, but cannot generalize to new samples that has not already observed. In this case, the error rate in training data set is very low compared to new data that has high error rate. To overcome overfitting problem and improve generalization, the dropout is used. Dropout is temporal removal of a random subset of units with their connections (Mamoshina *et al*., 2016). This technique is usually exploited during training and used to reduce conspiracy between units.

Since the training data sets are large, then deep learning algorithm requires high performance computing platform for practical applications (Li *et al*., 2018). The training process is usually time consuming, computationally intensive, and requires programming knowledge for graphics processing units (GPU). Requirements for GPU will even be more demanding in molecular modelling of breast cancer biomarkers (Parikesit, 2018). Parallelization through GPU acceleration, cloud computing, and multicore processing are created to overcome this limitation and enable deep learning algorithm to be recognized as significant intelligence (Mamoshina *et al*., 2016).

Future perspectives in breast cancer diagnostics

Deep learning algorithm has a potential to help the diagnostician in determining breast cancer. The applications of deep learning algorithm in diagnose breast cancer are various such as determining lesion, metastatic, mitosis, and gene expression. One major impact of deep learning algorithm in diagnosis process is decreasing human error rate that can cause inaccurate diagnostics. The development of deep learning algorithm delivers improvement of diagnosis accuracy in medical field compared to manual diagnosis. Thus, various types of deep learning algorithms have been developed to fulfil specific objective in breast cancer diagnostics.

This review article showed the development of deep learning algorithm in medical diagnostic especially in breast cancer. All of them showed a high accuracy in diagnosis process. This proved that by implementing deep learning algorithm, inaccurate diagnostic could be avoided. However, researcher needs to consider several limitations in deep learning algorithm, if not, it will return low accuracy of diagnosis.

In the future, deep learning algorithm may be used to determine stage of the breast cancer. However, it will require many tuning parameters to properly train the model for this task. The hyperparameters include dropout rate, kernel functions, and learning rate (Kim *et al*., 2018). A slight modification of these parameters can lead to drastic model with verifying performances. In the recent years, these parameters have relied on human experts. Therefore, further research is needed to find optimal hyper parameters for deep learning algorithm.

CONCLUSION

The development of deep learning algorithm grows in all fields of science, such as medical diagnostics. The combination of deep learning algorithm and diagnosticians has potentially improved the accuracy of diagnosis process in breast cancer. Despite the limitations, deep learning algorithm will be essential for diagnosis of breast cancer in this era.

ACKNOWLEDGEMENT

The authors would like to thank to Institute for Research and Community Services of Indonesia International Institute for Life Sciences (i3l) for their heartfelt support. Thanks also goes to Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jendera Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia for providing Hibah Penelitian Berbasis Kompetensi DIKTI/KOPERTIS III 2018 No. 049/KM/PNT/2018 and Hibah Penelitian Dasar DIKTI No.T/140/E3/RA.00/2019.

REFERENCES

Alakwaa, F. M., Chaudhary, K., Garmire, L. X. (2017). Deep learning accurately predicts estrogen receptor status in breast cancer metabolomics data. *Journal of Proteome Research, 17*(1), 337-347.

Albarqouni, S., Baur, C., Achilles, F., Belagiannis, V., Demirci, S., Navab, N. (2016). AggNet: Deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE Transactions on Medical Imaging, 35*(5), 1313-1321.

Andrianantoandro, E., Basu, S., Karig, D. K., Weiss, R. (2006). Synthetic biology: new engineering rules for an emerging discipline. Molecular Systems Biology, *2*, 2006.0028.

Araújo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., . . . Campilho, A. (2017). Classification of breast cancer histology images using Convolutional Neural Networks. *Plos One,12*(6).

Asawathy, M., Jagannath, M. (2017). Detection of breast cancer on digital histopathology images: Present status and future possibilities. *Informatics in Medicine Unlocked, 8*, 74-79.

Balogh, R., Wood, J., Lunsky, Y., Isaacs, B., Ouellette-Kuntz, H., Sullivan, W. (2015). Care of adults with developmental disabilities: Effects of a continuing education course for primary care providers. *Canadian Family Physician Medecin de Famille Canadien, 61(7)*, e316-23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/26380855

Becker, A. S., Marcon, M., Ghafoor, S., Wurnig, M. C., Frauenfelder, T., Boss, A. (2017). Deep Learning in Mammography. *Investigative Radiology, 52*(7), 434-440.

Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., … Venâncio, R. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA - Journal of the American Medical Association.

Breast cancer. (2017, December 21). Retrieved June 28, 2018, from http://www.cancerresearchuk.org/about-cancer/breast-cancer/advanced/tests-diagnose/ct-scan-brain

Cang, Z., Wei, G. (2017). TopologyNet: Topology based deep convolutional and multi-task neural networks for biomolecular property predictions. *PLOS Computational Biology,13*(7).

Cheng, J., Ni, D., Chou, Y., Qin, J., Tiu, C., Chang, Y., . . . Chen, C. (2016). Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in us images and pulmonary nodules in CT Scans. *Scientific Reports,6*(1).

Deng, L., Li, J., Huang, J.-T., Yao, K., Yu, D., Seide, F., … Acero, A. (2013). Recent advances in deep learning for speech research at Microsoft. *In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 8604–8608). IEEE.

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*, *161*, 1-13.

Guo, R., Lu, G., Qin, B., Fei, B. (2018). Ultrasound imaging technologies for breast cancer detection and management: A Review. Ultrasound in Medicine & Biology, *44*(1), 37-70.

Han, Z., Wei, B., Zheng, Y., Yin, Y., Li, K., Li, S. (2017). Breast cancer multi-classification from histopathological images with structured deep learning model. *Scientific Reports,* 7(1).

Kim, J., Hong, J., Park, H. (2018). Prospects of deep learning for medical imaging. *Precision and Future Medicine*, *2*(2), 37-52.

Li, H., Weng, J., Shi, Y., Gu, W., Mao, Y., Wang, Y., . . . Zhang, J. (2018). An improved deep learning approach for detection of thyroid papillary cancer in ultrasound images. *Scientific Reports*, *8*(1).

Li, X., Liu, L., Zhou, J., Wang, C. (2018). Heterogeneity analysis and diagnosis of complex diseases based on deep learning method. *Scientific Reports*, *8*(1).

Mamoshina, P., Vieira, A., Putin, E., Zhavoronkov, A. (2016). Applications of deep learning in biomedicine. *Molecular Pharmaceutics*, *13*(5),1445-1454.

Memisevic, R. (2015). Deep learning: Architectures, algorithms, applications. 2015 IEEE Hot Chips 27 Symposium (HCS). 22-25 August. Cupertino, CA, USA.

Oliver, A., Odena, A., Raffel, C. A., Cubuk, E. D., & Goodfellow, I. (2018). *Realistic Evaluation of Deep Semi-Supervised Learning Algorithms*. Retrieved from http://papers.nips.cc/paper/7585-realistic-evaluation-of-deep-semi-supervised-learning-algorithms

Parikesit, A. A. (2018). The construction of two- and three-dimensional molecular models for the miR-31 and its silencer as the triple negative breast cancer biomarkers. OnLine Journal of Biological Sciences, *18*(4), 424–431.

Pevsner, J. (2015). *Bioinformatics and Functional Genomics*. Chichester: Wiley Blackwell.

Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., Yang, G. (2017). Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*, *21*(1), 4-21.

Redig, A. J., Mcallister, S. S. (2013). Breast cancer as a systemic disease: A view of metastasis. *Journal of Internal Medicine*, *274*(2), 113-126.

Ting, D. S. W., Cheung, C. Y. L., Lim, G., Tan, G. S. W., Quang, N. D., Gan, A., … Wong, T. Y. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA - Journal of the American Medical Association.

Torng, W., Altman, R. B. (2017). 3D deep convolutional neural networks for amino acid environment similarity analysis. *BMC Bioinformatics*,*18*(1).

Tóth, D., Varga, Z., Tóth, J., Árkosy, P., & Sebő, É. (2018). Short- and Long-Term (10-year) Results of an Organized, Population-Based Breast Cancer Screening Program: Comparative, Observational Study from Hungary. World Journal of Surgery, *42*(5), 1396–1402.

Vandenberghe, M. E., Scott, M. L., Scorer, P. W., Söderberg, M., Balcerzak, D., Barker, C. (2017). Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. *Scientific Reports*, *7*, 45938.

Venhuizen, F. G., Ginneken, B. V., Liefers, B., Asten, F. V., Schreur, V., Fauser, S., . . . Sánchez, C. I. (2018). Deep learning approach for the detection and quantification of intraretinal cystoid fluid in multivendor optical coherence tomography. *Biomedical Optics Express*,*9*(4), 1545-1569.

Veta, M., Diest, P. J., Jiwa, M., Al-Janabi, S., Pluim, J. P. (2016). Mitosis counting in breast cancer: Object-level interobserver agreement and comparison to an automatic method. *Plos One*, *11*(8).

Wang, D., Khosla, A., Gargeya, R., Irshad, H., Beck, A. H. (2016). Deep learning for identifying metastatic breast cancer. The International Symposium on Biomedical Imaging. arXiv:1606.05718

WHO. (2015). WHO | Breast cancer: prevention and control.