BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Journal of Human Physiology
https://ojs.bilpublishing.com/index.php/jhp

ARTICLE

Deep Learning in Medical Imaging and Drug Design

Surayya Ado Bala[1*] Shri Ojha Kant[2] Adamu Garba[3]

1. Department of Computer Science and Engineering, Sharda University , India
2. Research and Technology Development Center, Sharda University, India
3. School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, 212013, China

ARTICLE INFO

Article history
Received: 14 December 2020
Accepted: 18 January 2021
Published Online: 31 January 2021

Keywords:
Deep learning
Medical imaging
Drugs design
Cheminformatics

ABSTRACT

Over the last decade, deep learning (DL) methods have been extremely successful and widely used in almost every domain. Researchers are now focusing on the convergence of medical imaging and drug design using deep learning to revolutionize medical diagnostic and improvement in the monitoring from response to therapy. DL a new machine learning paradigm that focuses on learning with deep hierarchical models of data. Medical imaging has transformed healthcare science, it was thought of as a diagnostic tool for disease, but now it is also used in drug design. Advances in medical imaging technology have enabled scientists to detect events at the cellular level. The role of medical imaging in drug design includes identification of likely responders, detection, diagnosis, evaluation, therapy monitoring, and follow-up. A qualitative medical image is transformed into a quantitative biomarker or surrogate endpoint useful in drug design decision-making. For this, a parameter needs to be identified that characterizes the disease baseline and its subsequent response to treatment. The result is a quantifiable improvement in healthcare quality in most therapeutic areas, resulting in improvements in quality and life duration. This paper provides an overview of recent studies on applying the deep learning method in medical imaging and drug design. We briefly discuss the fields related to the history of deep learning, medical imaging, and drug design.

1. Introduction

Even though deep learning algorithms exist for over a decade, their applications to solve real-world problems were very slow due to limited data and hardware computational power. Its booms started in 2016 when a model built using a deep learning algorithm (Alpha Go) beat the world champion player of Go[1]. This rejuvenates the researchers' interest in using deep learning algorithms to solve their various domain problems, especially in medical imaging and drug design.

Deep learning is an extension of an artificial neural network (ANN) that has been around for over three decades. This network work is based on mimicking human brain neurons [1]. An ANN is a shallow network consisting of an input layer, a single hidden layer, and an output layer, as shown in Fig.1. Each node in the input layer corresponds to a feature that is sent to the hidden layers. Upon

**Corresponding Author:*
Surayya Ado Bala,
Department of Computer Science and Engineering, Sharda University , India;
Email: surayyaadob@gmail.com

DOI: https://doi.org/10.30564/jhp.v2i2.2683

receiving the features (data), the hidden layer performs some computation on the data and transfers them to the output layer, where results are generated. The difference between target output values and the desired output value is calculated using the back-propagation algorithm [2]. These errors are propagated back to the input, and weight adjustment is mostly made using stochastic gradient descent [3]; this process continues until the error is negligible. Deep learning has drastically increased machine learning algorithms' performance, primarily due to the reduction or complete elimination of feature engineering [4] as shown in Fig.2. The algorithms can extract features themselves and use those extracted features to either predict, classify, or cluster depending on the task performed. Recent development of computing capabilities, especially graphics processing units (GPUs) used in speeding up the complex calculations performed by the various algorithms, helps increase deep learning algorithms [5].

Advancement of technology, data availability, and improvement of existing algorithms are vital factors for the recent success of deep learning [4].

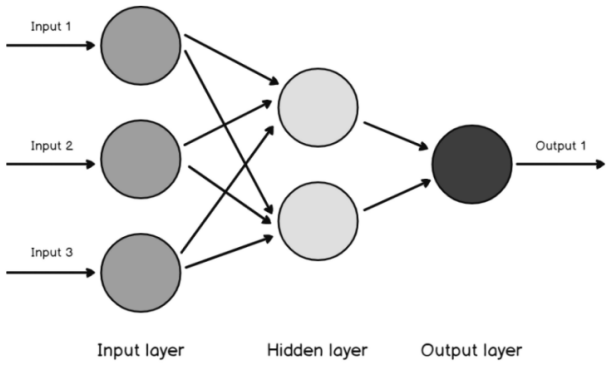

Figure 1. Artificial neural network

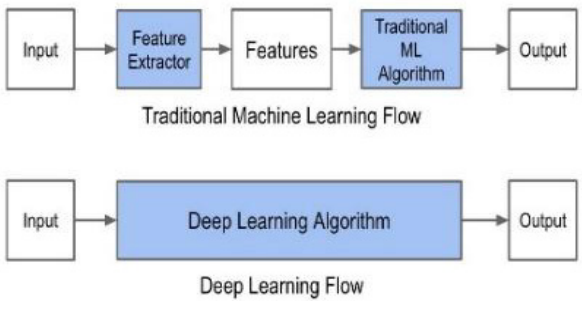

Figure 2. Machine learning vs Deep Learning

The ability to build a deep network with hundreds of hidden layers and millions of neurons give deep learning an edge over ANN, which mostly fails with more than four (4) hidden layers. Deep learning models have hyper-parameters such as activation function, learning rate, hidden layers, and many neurons which are essential factors to consider while designing your model. For medical images and drug discovery, the common activation function used is a rectified linear unit (RELU) [6]; in some instances, it variant like leaky RELU [7] is used. Model overfitting poses a significant challenge while training a deep network. Most of the models face this challenge where they learned all the features during the training but fails to perform the task during testing, i.e., using different datasets with the one used during the training. Regularization [8] and dropout [9] are commonly used techniques to reduce overfitting to solve this problem.

To test the robustness of deep learning, many researchers have built different models, one with machine learning and the other with deep learning, and used the same datasets to compare their performance. Various impressive results were obtained from the investigation conducted by [10-12] that deep learning models do not have superior performance over their machine learning counter fact without a large amount of training data. These experiments have shown that deep learning models performed better with a large number of training datasets.

Convolutional Neural Network (CNN) [13], Deep Belief Network (DBN) [14], Sparse and Variable Auto encoders [15] are among the commonly used deep learning algorithms in medical images and drug discovery. Selecting an algorithm depends on the task you want to perform. For medical image classification, different researchers used different architecture. The model proposed by [15] used CNN, and Auto encoders are used on the one proposed by [16]. Similarly, Auto encoders are used in drug discovery in the model proposed [17]. Generally, CNN is the most widely used architecture for image classification due to the robustness of pooling layers, and integrating dropout in the network has substantially decrease overfitting.

Even though Bayesian Network [18], Decision Tree [19], and Support Vector Machine (SVM) [20] are the most commonly used machine learning algorithms for drugs design and discovery, as shown in Fig. 3; recent trends of deep learning made it possible to make a substantial inroad to be among the algorithm used in drug discovery [1]. CNN, Autoencoders, Recurrent Neural Network (RNN), and other generative models are algorithms used in drug discovery and design [21-22].

 DOI: https://doi.org/10.30564/jhp.v2i2.2683

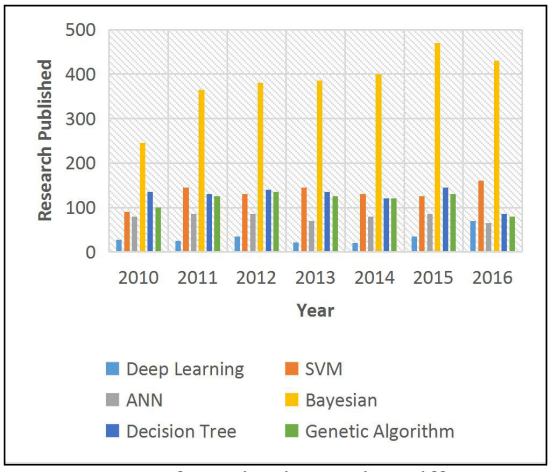

Figure 3. Analyzing the total number of publications using different machine learning algorithm for drugs design and discovery

2. Deep Learning in Medical Imaging

In a comparative analysis, deep learning models used in medical image analysis outperformed their counter fact of machine learning in anomaly detection, localization, image registration, segmentation, and diagnosis.

Most of the proposed medical image models are supervised learning based in nature; as such, annotated training data from experts are required to train them. Regrettably, obtaining training data in medical imaging is time-consuming, costly, and prone to errors. Furthermore, each medical image field has a separate template of the training process. This limitation has hindered the progress of the application of deep learning in medical images for years. The limitation was eliminated by the research performed by [23]. They proposed a general learning model that automatically extracts image pixel features; their model is flexible enough to be applied to any medical image field.

Clear separation of healthy and unhealthy body organs makes medical image segmentation one of the areas that received high research attention. Moreover, segmentation accuracy is the precursor of successful diagnosis and prevention. As such, various automated models were proposed to this effect [24]. The model proposed by [25] used automated image processing and clustering algorithms (K-means and expectation-maximization) to segment brain tumor. To improve spatial-temporal consistency of cardiac MRI segmentation, a hybrid spatio-temporal network (HST-Net) was proposed by [26]. Segmentation of children brain (from birth to age of 5) using MRI imaging is considered challenging to perform mainly due to noise increase, high volume effect, and minimization of tissue content [27], a CNN architecture model proposed by [28] to segment children brain solved the problems. The model proposed by [29] used both 2D and 3D CNN architecture to perform end-to-end volumetric segmentation of cardiac images.

Lesion and abnormality detections are a source of concern to many researchers in medical imaging due to the images' misclassification. With deep learning, various classifiers are used to classify images using binary classification. The models proposed by [30-32] used DNN architecture to detect and classify coronary classification of vein artery, cerebral microbleeds and, healthy and unhealthy skins.

3. Deep Learning in Drug Design

Complex molecule structure makes drug discovery and designs a challenging task to perform by the researchers in pharmacology and cheminformatics. This complexity arises from the hidden features among the molecules and features extracted from the molecules with artificial neural network models like SVM, decision tree, and genetics algorithm, which is a further step toward simplifying the drug design and development process [33-34]. The recent advances in deep learning have reduced the complexity of discovering molecule structures and their relationship within a compound. The importance of revealing their compound structure in obtaining qualitative classifiers or quantitative structure-activity relationship (QSAR) models.

For years, earlier research in the drug design domain continued to use human engineering to handcrafted molecule features descriptors. Even though some successes were reported in [35-38], deep learning benefits, like directly learning high powerful features among the molecules, are missing. Researchers have question that remained unanswered for years; can molecule complexity and hidden features structure be resolve by shifting from human engineering features extraction to deep learning models? Hilton Group took the challenge and proposed the first deep learning model for drug design and won the Merck Kaggle challenge 2012 (https://www.kaggle.com/c/MerckActivity). Similarly, a collaborative work between Hilton Group and Google in the subsequent year led to many research papers on deep learning-based QSAR modeling using different DNN architecture to perform multiple tasks. To imitate compound and protein interaction, different weights are given to the compound and protein features. These features become the input of the first hidden layer; thus, the model training is accelerated. The amount of training data does not show any significant effect on model performance [39]. Using this technique, a DNN model for drug discovery called AtomNet was proposed

DOI: https://doi.org/10.30564/jhp.v2i2.2683

by [40]. This model was the first to utilize deep CNN architecture to extract compound features in drug analysis and discovery. The back-bone of drug design lies in the feature identification of compound-protein interaction. A model that can predict this interaction and protein sequence generation was proposed by [41]. The reason for the difficulty of adapting deep learning in drug discovery has to do with molecules structure. The model proposed by [42] used several features and chemical properties to predict molecule structure in drug discovery. Their model has shown robust performance compared to the rest of the model in the same category. The model proposed by [17] was the first that used unsupervised learning in drug design. They used a seven-layer Generative Adversarial Network (GAN) to screen a compound. To differentiate their model from the traditional compound screening methods using QSAR, they extracted features from input molecular fingerprints and generated new fingerprints that they used for training and testing their model. Their model outperformed all the traditional QSAR compound, screening models.

4. Conclusion and Promising Future

As we explained in this paper, different deep learning models have been applied to different medical imaging and drug discovery tasks as depicted in Fig.4 , which have achieved high performance with huge training data availability. The same performance was not obtained with a small amount of training data. This has shown that the success of any deep learning model depends on data. Some researchers [43] view that ANN models have the same performance precedence as deep learning models without enough training data. The question researchers continue to ask is how to quantify the amount of enough training data, and this has become a topic of discussion among them. Even though various techniques like transfer learning are developed to mitigate the scarcity of training data, its effectiveness varies across different domains and tasks to perform. Similarly, hyper-parameters tuning, the number of hidden layers, and the type of activation function used are also difficult to decide because each model performs best using different activation function and hyper-parameters values.

Even though deep learning has achieved almost the same accuracies as a human being, especially in image classification and segmentation using well-annotated datasets [44-45], their full adaptation in compound structures domain like drugs design and discovery [45] is problematic due to high constraint of the number of input features the models accept especially RNN, CNN and Restricted Boltzmann Machine [47] architectures. Simultaneously, high-performance models that perform various medical imaging tasks with limited training data are absent. This has continued to slow down deep learning models on lung cancer, liver, and spinal code injuries.

In the final analysis, although many successes have been recorded of using deep learning in both medical imaging, drug design, and discovery, development and improvements of models that will perform well with a limited amount of data are in dire need. Furthermore, deep learning models that can extract disease structure from the medical image is highly needed in our health care system. Additionally, to solve the problem associated with deep learning in compound structure domains like drug design and discovery, there is a need to optimize model architecture that will automatically extract useful molecule features and infer that the compound's relationship can be easily observed. Models that can perform these tasks will out rightly speedup drug design and clinical trials, reducing the time taken to produce drugs.

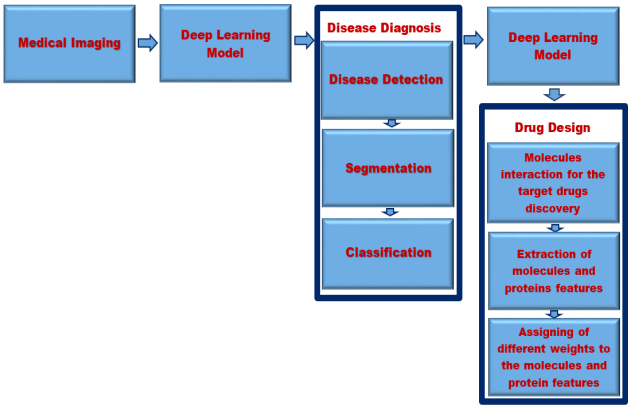

Figure 4. Application of Deep Learning from Medical Imaging to Drug Design

References

[1] Y. Jing, Y. Bian, Z. Hu, L. Wang, X.-Q. S. Xie. Deep learning for drug design: an artificial intelligence paradigm for drug discovery in the big data era, The AAPS Journal, 2018, 20: 58.

[2] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors, Cognitive modeling, 1988, 5: 1.

[3] Y. LeCun, Y. Bengio, G. Hinton. Deep learning, Nature, 2015, 521: 426.

[4] Bala S.A, Kant S., Kumar K. Impact of deep learning in medical imaging: a systematic new proposed model. International Journal of recent technology and engineering, 2019: 112-118.

[5] B. J. Erickson, P. Korfiatis, T. L. Kline, Z. Akkus, K. Philbrick, A. D. Weston. Deep learning in radiology: does one size fit all?, Journal of the American Col-

 DOI: https://doi.org/10.30564/jhp.v2i2.2683

lege of Radiology, 2018, 15: 521-526.

[6] X. Glorot, A. Bordes, Y. Bengio. Deep sparse rectifier neural networks, in Proceedings of the fourteenth international conference on artificial intelligence and statistics, 2011: 315-323.

[7] D.-A. Clevert, T. Unterthiner, S. Hochreiter. Fast and accurate deep network learning by exponential linear units (elus), arXiv preprint arXiv:1511.07289, 2015.

[8] D. A. Winkler, T. C. Le. Performance of deep and shallow neural networks, the universal approximation theorem, activity cliffs, and QSAR, Molecular informatics, 2017, 36: 1600118.

[9] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting, The Journal of Machine Learning Research, 2014, 15: 1929-1958.

[10] B. Gao, Y. Xu. Univariant approximation by superpositions of a sigmoidal function, Journal of mathematical analysis and applications, 1993, 178: 221-226.

[11] G. B. Goh, N. O. Hodas,A. Vishnu. Deep learning for computational chemistry, Journal of computational chemistry, 2017, 38: 1291-1307.

[12] S. J. Capuzzi, R. Politi, O. Isayev, S. Farag, A. Tropsha. QSAR modeling of Tox21 challenge stress response and nuclear receptor signaling toxicity assays, Frontiers in Environmental Science, 2016, 4: 3.

[13] K. Simonyan, A. Zisserman. Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv, 2014, 1409.1556.

[14] G. E. Hinton, S. Osindero, Y.-W. Teh. A fast learning algorithm for deep belief nets, Neural computation, 2006, 18: 1527-1554.

[15] A. Ng. Sparse autoencoder, CS294A Lecture notes, 2011, 72: 1-19.

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

[17] A. Kadurin, A. Aliper, A. Kazennov, P. Mamoshina, Q. Vanhaelen, K. Khrabrov, et al. The cornucopia of meaningful leads: Applying deep adversarial autoencoders for new molecule development in oncology, Oncotarget, 2017, 8: 10883.

[18] P. Domingos, M. Pazzani. On the optimality of the simple Bayesian classifier under zero-one loss, Machine learning, 1997, 29: 103-130.

[19] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, B. P. Feuston. Random forest: a classification and regression tool for compound classification and QSAR modeling, Journal of chemical information and computer sciences, 2003, 43: 1947-1958.

[20] C. Cortes, V. Vapnik. Support-vector networks, Machine learning, 1995, 20: 273-297.

[21] S. Ekins. The next era: Deep learning in pharmaceutical research, Pharmaceutical Research, 2016, 33: 2594-2603.

[22] E. Gawehn, J. A. Hiss, G. Schneider. Deep learning in drug discovery, Molecular informatics, 2016, 35: 3-14.

[23] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, D. Shen. Unsupervised deep feature learning for deformable registration of MR brain images, in International Conference on Medical Image Computing and Computer-Assisted Intervention, 2013: 649-656.

[24] Bala S.A., Kant S. Deep learning-based model architectures for cardiac MRI segmentation: A Survey. International Journal of Innovative Science, Engineering & Technology, 2020: 129-135.

[25] S. M. Anwar, S. Yousaf, M. Majid. Brain tumor segmentation on Multimodal MRI scans using EMAP Algorithm, in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018: 550-553.

[26] Bala S.A, Kant, S. A hybrid Spatio-temporal network for cardiac MRI segmentation. International Journal of advanced trends in computer science and engineering, 2020: 9089-9097.

[27] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. Benders, I. Išgum. Automatic segmentation of MR brain images with a convolutional neural network, IEEE transactions on medical imaging, 2016, 35: 1252-1261.

[28] H. Xue, L. Srinivasan, S. Jiang, M. Rutherford, A. D. Edwards, D. Rueckert, et al. Automatic segmentation and reconstruction of the cortex from neonatal MRI, Neuroimage, 2007, 38: 461-477.

[29] M. Prastawa, J. H. Gilmore, W. Lin, G. Gerig. Automatic segmentation of MR images of the developing newborn brain, Medical image analysis, 2005, 9: 457-466.

[30] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, et al. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation, NeuroImage, 2015, 108: 214-224.

[31] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger. 3D U-Net: learning dense volumetric segmentation from sparse annotation, in International conference on medical image computing and computer-assisted intervention, 2016: 424-432.

[32] J. M. Wolterink, T. Leiner, B. D. de Vos, R. W. van Hamersvelt, M. A. Viergever, I. Išgum. Automatic coronary artery calcium scoring in cardiac CT angi-

DOI: https://doi.org/10.30564/jhp.v2i2.2683

ography using paired convolutional neural networks, Medical image analysis, 2016, 34: 123-136.

[33] Q. Dou, H. Chen, L. Yu, L. Zhao, J. Qin, D. Wang, et al. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks, IEEE transactions on medical imaging, 2016, 35: 1182-1195.

[34] D. Sheet, S. P. K. Karri, A. Katouzian, N. Navab, A. K. Ray, J. Chatterjee. Deep learning of tissue specific speckle representations in optical coherence tomography and deeper exploration for in situ histology, in 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), 2015: 777-780.

[35] G. Schneider. Adaptive systems in drug design: CRC Press, 2002.

[36] G. E. Hinton, J. A. Anderson. Parallel models of associative memory: updated edition: Psychology press, 2014.

[37] B. F. Begam and J. S. Kumar. A study on cheminformatics and its applications on modern drug discovery, Procedia engineering, 2012, 38: 1264-1275.

[38] F. K. Brown. Chemoinformatics: what is it and how does it impact drug discovery, Annual reports in medicinal chemistry, 1998, 33: 375-384.

[39] V. Arulmozhi, R. Rajesh. Chemoinformatics-a quick review, in 2011 3rd International Conference on Electronics Computer Technology, 2011: 416-419.

[40] R. G. Bone, M. A. Firth, R. A. Sykes. SMILES extensions for pattern matching and molecular transformations: Applications in chemoinformatics, Journal of chemical information and computer sciences, 1999, 39: 846-860.

[41] J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl, V. Svetnik. Deep neural nets as a method for quantitative structure–activity relationships, Journal of chemical information and modeling, 2015, 55: 263-274.

[42] I. Wallach, M. Dzamba, A. Heifets. AtomNet: a deep convolutional neural network for bioactivity prediction in structure-based drug discovery, arXiv preprint arXiv, 2015, 1510.02855.

[43] F. Wan, J. Zeng. Deep learning with feature embedding for compound-protein interaction prediction, bioRxiv, 2016: 086033.

[44] A. Mayr, G. Klambauer, T. Unterthiner, S. Hochreiter. DeepTox: toxicity prediction using deep learning, Frontiers in Environmental Science, 2016, 3: 80.

[45] J. Schmidhuber. Deep learning in neural networks: An overview, Neural networks, 2015, 61: 85-117.

[46] H. Chen, O. Engkvist, Y. Wang, M. Olivecrona, T. Blaschke. The rise of deep learning in drug discovery, Drug discovery today, 2018, 23:1241-1250.

[47] Y. Wang, J. Zeng. Predicting drug-target interactions using restricted Boltzmann machines, Bioinformatics, 2013, 29: i126-i134.

 DOI: https://doi.org/10.30564/jhp.v2i2.2683