APPLICATION OF DEEP LEARNING IN HEALTH INFORMATICS: A REVIEW

Vinit Mehta[1], Noopur Shrivastava[2]
E-Mail Id: vinit741@gmail.com, noopur.shrivastava1@anandice.ac.in
[1]Department of Electrical Engineering, M.B.M. Engineering College, Jodhpur, India
[2]Department of Electrical Engineering Anand International Collge of Engineering Jaipur, India

Abstract- Today a variety of health care practices have been evolved to maintain and restore health by the latest prevention and best treatment. This implements biomedical sciences, biomedical research, genetics and medical technology to diagnose, treat, and prevent injury and disease, typically through pharmaceuticals or surgery, therapies as divers as psychotherapy, external splints   and traction, medical devices, biologics, and ionizing radiation. With advances in technology, the health sciences are constantly pushing toward more effective treatments and cures. With a massive influx of multimodality data, the role of data analytics in health informatics has grown rapidly in the last decade. This has also provoked increasing interests in the generation of analytical, data driven models based on machine learning in health informatics. Deep learning, a technique with its foundation in artificial neural networks, is emerging in recent years as a powerful tool for machine learning, promising to reform the future of artificial intelligence. Rapid improvements in computational power, fast data storage, and parallelization have also contributed to the rapid uptake of the technology in addition to its predictive power and ability to generate automatically optimized high-level features and semantic interpretation from the input data. This paper presents a comprehensive review of research employing deep learning in health informatics, providing a critical analysis of the relative merit, and potential pitfalls of the technique as well as its future outlook. The paper mainly focuses on key applications of deep learning in the fields of translational bioinformatics, medical imaging, pervasive sensing, medical informatics, and public health. Finally the limitations and challenges of deep learning in the field of health informatics have been discussed.
Keywords: machine learning, deep learning, health informatics, translational bioinformatics, medical imaging, pervasive sensing, medical informatics, and public health

1. INTRODUCTION

Deep learning has become a new trend in machine learning since recent years. The background foundations of deep learning are well rooted in the classical neural network (NN) theory. But unlike traditional NNs, deep learning algorithm uses many hidden neurons and layers, typically more than two that contribute an architectural advantage combined with new training paradigms. Deep learning is based on a set of algorithms that model high level abstractions in data. In a simple case, there might be two sets of neurons: one set that receives an input signal and one that sends an output signal. When the input layer receives an input it passes on a modified version of the input to the next layer. In a deep network, there are many layers between the input and the output, that allows the algorithm to use multiple processing layers, composed of multiple linear and non-linear transformations. Fig. 1 shows that each successive layer in a neural network uses features from the previous layer to learn more complex features. In this manner a deep neural network is capable of composing features of increasing complexity in each of its successive layers.
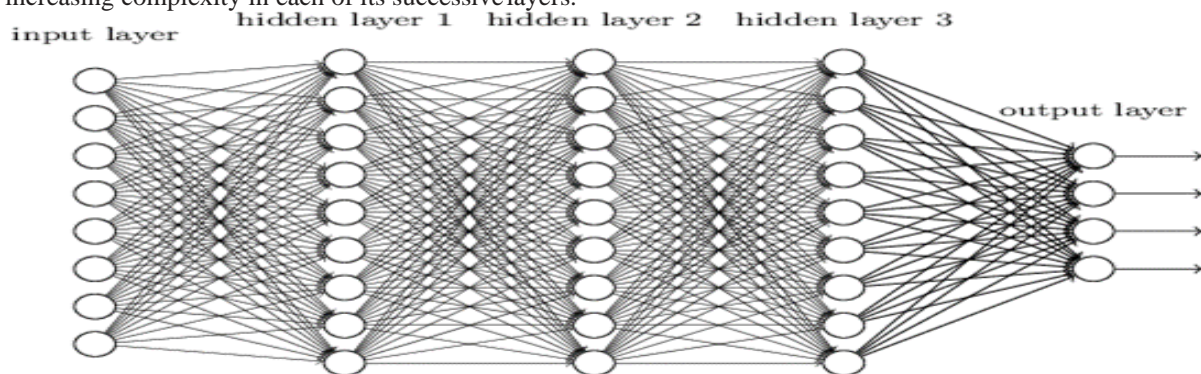

Fig. 1.1 Block diagram of Deep Neural Network

Also the layer-by-layer pipeline of nonlinear combination of their outputs generates a lower dimensional projection of the input space. Every lower-dimensional projection corresponds to a higher perceptual level. This results in an effective high-level abstraction of the raw data or images if the network is optimally weighted.
In the past few years, deep learning has become an immediate tool to solve applications in computer vision, speech recognition, natural language processing, and other application areas of commercial interest. In areas such as health informatics, the generation of the deep learning feature set without human intervention has many advantages. In medical imaging, it can generate features that are more sophisticated and difficult to elaborate in descriptive means. Implicit features could determine fibroids and polyps [1], and characterize irregularities in tissue morphology such as tumors [2]. In translational bioinformatics, such algorithms may also determine nucleotide sequences that could bind a DNA or RNA strand to a protein [3]. Fig. 2 outlines a quick rush of interest in deep learning in recent years in terms of the number of papers published in sub-fields in health informatics including bioinformatics, medical imaging, pervasive sensing, medical informatics, and public

health. The publication statistics have obtained from Google Scholar; the search phrase is defined as the subfield name with the exact phrase deep learning and at least one of medical or health appearing, e.g., ‒public health‖ ‒deep learning‖medical OR health.

A excess of experimental works have implemented deep learning models for heath informatics, reaching similar performance or in many cases exceeding that of alternative techniques. Nevertheless, the application of deep learning to health informatics raises a number of challenges that need to be resolved.
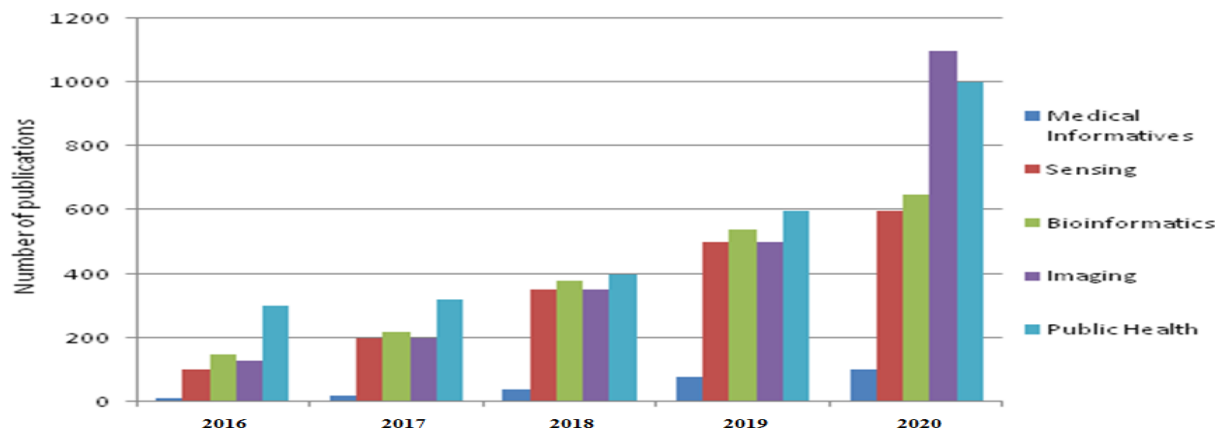

Fig. 2.1 Number of published papers that use deep learning in health informatic applications

In addition, deep learning requires extensive computational resources, without which training could become excessively time-consuming. Attaining an optimal definition of the network's free parameters can become a particularly laborious task to solve. Eventually, deep learning models can be affected by convergence issues as well as overfitting, hence supplementary learning strategies are required to address these problems [4]. In this paper, a review of recent health informatics studies has been done that employ deep learning to discuss its relative strength and potential pitfalls. Furthermore, their schemas and operational frameworks are described in detail to clarify their practical implementations, as well as expected performance.

FROM PERCEPTRON TO DEEP LEARNING Perceptron is a bio inspired algorithm for binary classification and consist of an input layer directly connected to an output node, emulate this biochemical process through an activation function (also referred to as a transfer function) and a few weights. Specifically, it can learn to classify linearly separable patterns by adjusting these weights accordingly.

To solve more complex problems, NNs with one or more hidden layers of perceptrons have been introduced [5]. To train these NNs, many stages or epochs are usually performed where each time the network is presented with a new input sample and the weights of each neuron are adjusted based on a learning process called delta rule. The delta rule is used by the most common class of supervised NNs during the training and is usually implemented by exploiting the back-propagation routine [6]. Specifically, without any prior knowledge, random values are assigned to the network weights. Through an iterative training process, the network weights are adjusted to minimize the difference between the network outputs and the desired outputs. The most common iterative training method uses the gradient descent method where the network is optimized to find the minimum along the error surface. This method requires the activation functions to be always differentiable.

Adding more hidden layers to the network allows a deep architecture to be built that can express more complex hypothesis as the hidden layers confine the nonlinear relationships. These NNs are known as Deep Neural Networks (DNN).

Deep learning has provided new sophisticated approaches to train DNN architectures. In general, DNNs can be trained with unsupervised and supervised learning methodologies. In supervised learning, labeled data are used to train the DNNs and learn the weights that minimize the error to predict a target value for classification or regression, whereas in unsupervised learning, the training is performed without requiring labeled data. Unsupervised learning is usually used for clustering, feature extraction or dimensionality reduction. For some applications it is common to combine an initial training procedure of the DNN with an unsupervised learning step to extract the most relevant features and then use those features for classification by exploiting a supervised learning step.

For many years, hardware limitations have made DNNs impractical due to high computational demands for both training and processing, especially for applications that require real-time processing. Recently, these limitations have been partially overcome and have enabled DNNs to be recognized as a significant breakthrough in artificial intelligence.

2. APPLICATIONS

2.1 Medical Informatics

Medical Informatics focus on the analysis of large, aggregated data in healthcare fields with the aim to enhance and develop clinical decision support systems or assess medical data both for quality assurance and accessibility of health care services. Deep learning approaches have been designed to scale up well with big and distributed datasets. The success of DNNs is largely due to their ability to learn novel features/patterns and understand data representation in both an unsupervised and supervised hierarchical manners. DNNs have also proven to be

efficient in handling multi information since they can combine several DNN architectural components. Therefore, it is unsurprising that deep learning has quickly been adopted in medical informatics research. In the field of medical informatics, many authors have made several documentations. Shin et al. [7] have presented a combined text image CNN to identify semantic information that links radiology images and reports from a typical picture archiving and communication system. Liang et al. [8] have used a modified version of DNN as an effective training method for large scale datasets on hypertension. Putin et al. [9] have applied DNNs for identifying markers that predict human chronological age based on simple blood tests. Nie et al. [10] have proposed a deep learning network for automatic disease inference, which requires manual gathering the key symptoms or questions related to the disease.

In another study, Mioto et al. [11] have showed that a stack of autoencoders can be used to automatically infer features from a large scale EHR database and represent patients without requiring additional human effort. The authors demonstrated the ability of their system to predict the probability of a patient developing specific diseases, such as diabetes, schizophrenia and cancer. Furthermore, Futoma et al. [12] have compared different models in their ability to predict hospital readmissions based on a large EHR database. DNNs have significantly higher prediction accuracies than conventional approaches, such as penalized logistic regression, though training of the DNN models were not straightforward. Mehrabi et al. [13] have proposed the use of DBN to discover common temporal patterns and characterize disease progression. The authors have highlighted the ability to interpret the newly discovered patterns requires further investigation.

The motivations behind these studies are to develop general purpose systems to accurately predict length of stay, future illness, readmission, and mortality with the view to improve clinical decision making and optimize clinical pathways. Early prediction in health care is directly related to saving patient's lives. Furthermore, the discovery of novel patterns can result in new hypotheses and research questions. In computational phenotype research, the goal is to discover meaningful data-driven features and disease characteristics. For example, Che et al. [14] have highlighted that although DNNs outperform conventional machine learning approaches in their ability to predict and classify clinical events, they suffer from the issue of model interpretability, which is important for clinical adaptation. They pointed out that interpreting individual units can be misleading and the behavior of DNNs are more complex than originally thought. They suggested that once a DNN is trained with big data, a simpler model can be used to distil knowledge and mimic the prediction performance of the DNN.

Deep learning has paved the way for personalized health care by offering an unprecedented power and efficiency in mining large multimodal unstructured information stored in hospitals, cloud providers and research organization. Although, it has the potential to outperform traditional machine learning approaches, appropriate initialization and tuning is important to avoid overfitting. Noisy and sparse datasets result in considerable fall of performance indicating that there are several challenges to be addressed. Furthermore, adopting these systems into clinical practice requires the ability to track and interpret the extracted features and patterns.

2.2 Sensing

Pervasive sensors, such as wearable, implantable, and ambient sensors have [15] allowed continuous monitoring of health and wellbeing. An accurate estimation of food intake and energy expenditure throughout the day can help tackle obesity and improve personal wellbeing. For elderly patients with chronic diseases, wearable and ambient sensors can be utilized to improve quality of care by enabling patients to continue living independently in their own homes. The care of patients with disabilities and patients undergoing rehabilitation can also be improved through the use of wearable and implantable assistive devices and human activity recognition. For patients in critical care, continuous monitoring of vital signs, such as blood pressure, respiration rate, and body temperature, are important for improving treatment outcomes by closely analyzing the patient's condition [16].

Wulsin et al. [17] proposed a DBN approach to detect anomalies in EEG waveforms. EEG is used to record electrical activity of the brain. Interpreting the waveforms from brain activity is challenging due to the high dimensionality of the input signal and the limited understanding of the brain operations. Jia et al. [18] have used a deep learning method based on RBM to recognize affective state of EEG. Although the sample sets are small and noisy, the proposed method achieves greater accuracy. A DBN was also used for detecting arrhythmias from electrocardiography (ECG) signals. The main purpose of the system is identifying arrhythmias which are a complex pattern recognition problem. Yan et al. have attained classification accuracies of 98% using a two-lead ECG dataset. For low power wearable and implantable EEG sensors, where energy consumption and efficiency are major concerns, Wang et al. [19] have designed a DBN to compress the signal. This results in more than 50% of energy savings while retaining accuracy for neural decoding. The introduction of deep learning has increased the utility of pervasive sensing across a range of health applications by improving the accuracy of sensors that measure food calorie intake, energy expenditure, activity recognition, sign language interpretation, and detection of anomalous events in vital signs. Many applications use deep learning to achieve greater efficiency and performance for real-time processing on low-power devices; however, a greater focus should be placed upon implementations on neuromorphic hardware platforms designed for low-power parallel processing. The most significant improvements in performance have been achieved where the data has high dimensionality as seen in the EEG datasets. Most current research has focused on the recognition of activities of daily living and brain activity. Many opportunities for other applications and diseases remain, and many currently studies still rely upon relatively small datasets that may not fully capture the variability of the real world.

2.3 Bioinformatics

Bioinformatics aim to investigate and understand biological processes at a molecular level. Advances in

biotechnology have helped reduce the cost of genome sequencing and steered the focus on prognostic, diagnostic and treatment of diseases by analyzing genes and proteins. This can be illustrated by the fact that sequencing the first human genome cost billions of dollars, whereas today it is affordable [20]. Further motivated by P4 (predictive, personalized, preventive, participatory) medicine [21], bioinformatics aim to predict and prevent diseases by involving patients in the development of more efficient and personalized treatments. The application of machine learning in bioinformatics can be divided into three areas: prediction of biological processes, prevention of diseases and personalized treatment. These areas are found in genomics, pharmacogenomics and epigenomics. Genomics explores the function and information structures encoded in the DNA sequences of a living cell [22]: it analyzes genes responsible for the creation of protein sequences and the expression of phenotypes. A goal of genomics is to identify gene alleles and environmental factors that contribute to diseases such as cancer. Pharmacogenomics evaluates variations in an individual's drug response to treatment brought about by differences in genes. It aims to design more efficient drugs for personalized treatment. Finally, epigenomics aims to investigate protein interactions and understand higher level processes, such as transcriptome (mRNA count), proteome, and metabolome, which lead to modification in the gene's expression. Understanding how environmental factors affect protein formation and their interactions is a goal of epigenomics.

The ability of deep learning to abstract large, complex, and unstructured data offers a powerful way of analyzing heterogeneous data such as gene alleles, proteins occurrences, and environmental factors [23]. In deep learning approaches, feature extraction and model fitting takes place in a unified step. Multilayer feature representation can capture nonlinear dependencies at multiple scales of transcriptional and epigenetic interactions and can model molecular structure and properties in a data driven way. These nonlinear features are invariant to small input changes which results in eliminating noise and increasing the robustness of the technique.

Several works have demonstrated that deep learning features outperformed methods relying on visual descriptors in the recognition and classification of cancer cells. For example, Ibrahim et al. [24] have proposed a DBN with an active learning approach to find features in genes and microRNA that resulted in the best classification performance of various cancer diseases such as hepatocellular carcinoma, lung cancer and breast cancer. For breast cancer genetic detection, Khademi et al. [25] have overcame missing attributes and noise by combining a DBN and Bayesian network to extract features from microarray data. Deep learning approaches have also outperformed SVM in predicting splicing code and understanding how gene expression changes by genetic variants [26]. Angermueller et al. [27] have used DNN to predict DNA methylation states from DNA sequences and incomplete methylation profiles. After applying to 32 embryonic mice stem cells, the baseline model was compared with the results. This method can be used for genome-wide downstream analyses.

Deep learning not only outperforms conventional approaches but also opens the door to more efficient methods to be developed. Kearnes et al. [28] have described how deep learning based on graph convolutions can encode molecular structural features, physical properties, and activities in other assays. This allows a rich representation of possible interactions beyond the molecular structural information encoded in standard databases. Similarly, multitask DNNs provides an intuitive model of correlation between molecule compounds and targets because information can be shared among different nodes. This increases robustness, reduces chances to miss information, and usually outperforms other methods that process large datasets [29].

2.4 Medical Imaging

Automatic medical imaging analysis is crucial in modern medicine. Diagnosis based on the interpretation of images can be highly subjective. Computer-aided diagnosis can provide an objective assessment of the underlying disease processes. Modeling of disease progression, common in several neurological conditions, such as Alzheimer's, multiple sclerosis, and stroke, requires analysis of brain scans based on multimodal data and detailed maps of brain regions. In recent years, CNNs have been adapted rapidly by the medical imaging research community because of their outstanding performance demonstrated in computer. The fact that CNNs in medical imaging have yielded promising results have also been highlighted in a recent survey of CNN approaches in brain pathology segmentation [30] and in an editorial of deep learning techniques in computer aided detection, segmentation, and shape analysis [31].

Among the biggest challenges in CAD are the differences in shape and intensity of tumors/lesions and the variations in imaging protocol even within the same imaging modality. In several cases, the intensity range of pathological tissue may overlap with that of healthy samples. Furthermore, Rician noise, non isotropic resolution, and bias field effects in magnetic resonance images (MRI) cannot be handled automatically using simpler machine learning approaches. To deal with this data complexity, hand-designed features are extracted and conventional machine learning approaches are trained to classify them in a completely separate step.

Deep learning provides the possibility to automate and merge the extraction of relevant features with the classification procedure. CNNs inherently learn a hierarchy of increasingly more complex features and, thus, they can operate directly on a patch of images centered on the abnormal tissue. Example applications of CNNs in medical imaging include the classification of interstitial lung diseases based on computed tomography (CT) images, the classification of tuberculosis manifestation based on X-ray images, the classification of neural progenitor cells from somatic cell source , the detection of haemorrhages in color fundus images and the organ or body-part-specific anatomical classification of CT images. A multistage deep learning framework based on CNNs extracts both the patches with the most as well as least discriminative local patches in the pretraining stage. Subsequently, a boosting stage exploits this local information to improve performance. The authors point out that training based on discriminative local appearances are more accurate compared to the usage of global

image context. Although CNNs have dominated medical image analysis applications, other deep learning approaches/architectures have also been applied successfully. In a recent paper, a stacked denoising autoencoder was proposed for the diagnosis of benign malignant breast lesions in ultrasound images and pulmonary nodules in CT scans [32]. The method outperforms classical CAD approaches, largely due to the automatic feature extraction and noise tolerance. Furthermore, it eliminates the image segmentation process to obtain a lesion boundary. Shan et al. [33] presented a stacked sparse autoencoder for microaneurysms detection in fundus images as an instance of a diabetic retinopathy strategy. The proposed method learns high-level distinguishing features based only on pixel intensities. Various autoencoder-based learning approaches have also been applied to the automatic extraction of biomarkers from brain images and the diagnosis of neurological diseases. These methods often use available public domain brain image databases such as the Alzheimer's disease neuroimaging initiative database. For example, a deep Autoencoder combined with a softmax output layer for regression is proposed for the diagnosis of Alzheimer's disease.

Low level image processing, such as image segmentation and registration can also benefit from deep learning models. Brosch et al. [34] described a manifold learning approach of 3-D brain images based on DBN. It is different than other methods because it does not require a locally linear manifold space. Mansoor et al. [35] developed a fully automated shape model segmentation mechanism for the analysis of cranial nerve systems. The deep learning approach outperforms conventional methods particularly in regions with low contrast, such as optic tracts and areas with pathology. A pipeline is proposed for object detection and segmentation in the context of automatically processing volumetric images. A novel framework called marginal space deep learning implements an object parameterization in hierarchical marginal spaces combined with automatic feature detection based on deep learning. A DNN architecture called input–output deep architecture is described to solve the image labelling problem. A single NN forward step is used

to assign a label to each pixel. This method avoids the handcrafted subjective design of a model with a deep learning mechanism, which automatically extracts the dependencies between labels. Deep learning is also used for processing hyperspectral images. Spectral and spatial learned features are combined together in a hierarchical model to characterize tissues or materials.

In general, deep learning in medical imaging provides automatic discovery of object features and automatic exploration of feature hierarch and interaction. In this way, a relatively simple training process and a systematic performance tuning can be used, making deep learning approaches improve over the state-of-the art. However, in medical imaging analysis, their potentials have not been unfolded fully. To be successful in disease detection and classification approaches, deep learning requires the availability of large labeled datasets. Annotating imaging datasets is an extremely time-consuming and costly process that is normally undertaken by medical doctors. Currently, there is a lot of debate on whether to increase the number of annotated datasets with the help of non-experts (crowd- sourcing) and how to standardize the available images to allow objective assessment of the deep learning approaches.

2.5 Public Health

Public health aims to prevent disease, prolong life, and promote healthcare by analyzing the spread of disease and social behaviors in relation to environmental factors. Public health studies consider small localized populations to large populations that encompass several continents such as in the case of epidemics and pandemics. Applications involve epidemic surveillance, modeling lifestyle diseases, such as obesity, with relation to geographical areas, monitoring and predicting air quality, drug safety surveillance, etc. The conventional predictive models scale exponentially with the size of the data and use complex models derived from physics, chemistry, and biology. Nevertheless, existing computational methods are able to accurately model several phenomena, including the progression of diseases or the spread of air pollution. However, they have limited abilities in incorporating real time information, which could be crucial in controlling an epidemic or the adverse effects of a newly approved medicine. In contrast, deep learning approaches have a powerful generalization ability. They are data-driven methods that automatically build a hierarchical model and encode the information within their structure. Most deep learning algorithm designs are based on online machine learning and, thus, optimization of the cost function takes place sequentially as new training datasets become available. One of the simplest online optimization algorithms applied in DNNs is stochastic gradient descent. For these reasons, deep learning, along with recommendation systems and network analysis, are suggested as the key analysis methods for public health studies [36].

Another interesting application is tracking outbreaks with social media for epidemiology and lifestyle diseases. Social media can provide rich information about the progression of diseases, such as Influenza and Ebola, in real time. Zhao et al. [37] used the microblogging social media service, Twitter, to continuously track health states from the public. DNN is used to mine epidemic features that are then combined into a simulated environment to model the progression of disease. Text from Twitter messages can also be used to gain insight into antibiotics and infectious intestinal diseases. DBN is used to categorize antibiotic- related Twitter posts into nine classes (side effects, wanting/needing, advertisement, advice/information, animals, general use, resistance, misuse, and other). To obtain the classifier, Twitter messages were randomly selected for manual labeling and categorization. They used a training set of 412 manually labeled and 150 000 unlabeled examples. A deep learning approach based on RBMs was pretrained in a layer-by-layer procedure. Fine-tuning was based on standard back propagation and the labeled data. Deep learning is used to create a topical vocabulary of keywords related to three types of infectious intestinal disease—campylobacter, norovirus, and food poisoning. When compared to officially documented cases, their results show that social media can be a good predictor of

intestinal diseases.

For tracking certain stigmatized behaviors, social media can also provide information that is often undocumented; Garimella et al. [38] used geographically-tagged images from Instagram to track lifestyle diseases, such as obesity, drinking, and smoking, and compare the self-categorization of images from the user against annotations obtained using a deep learning algorithm.

Furthermore, mining food and drug records to identify adverse events could provide vital large scale alert mechanisms. We have presented a few examples that use deep learning for early identification and modeling the spread of epidemics and public health risks. However, strict regulation that protects data privacy limits the access and aggregation of the relevant information. For example, Twitter messages or Facebook posts could be used to identify new mothers at risk from postpartum depression. Although, this is positive, there is controversy associated of whether this information should become available, since it stigmatizes specific individuals. Therefore, it has become evident that we need to strike a balance between ensuring individuals can control access to their private medical information and providing pathways on how to make information available for public health studies [39]. The complexity and limited interpretability of deep learning models constitute an obstacle in allowing an informed decision about the precise operation of a DNN, which may limit its application in sensitive data.

3. DEEP LEARNING IN HEALTHCARE: LIMITATIONS AND CHALLENGES

Although for different artificial intelligence tasks, deep learning techniques can deliver substantial improvements in comparison to traditional machine learning approaches, many researchers and scientists remain sceptical of their use where medical applications are involved. These scepticisms arise since deep learning theories have not yet provided complete solutions and many questions remain unanswered. The following four aspects summarize some of the potential issues associated with deep learning:

- ➢ Despite some recent work on visualizing high level features by using the weight filters in a CNN, the entire deep learning model is often not interpretable. Consequently, most researchers use deep learning approaches as a black box without the possibility to explain why it provides good results or without the ability to apply modifications in the case of misclassification issues.
- ➢ To train a reliable and effective model, large sets of training data are required for the expression of new concepts. Although an explosion of available healthcare data with many organizations is often limited. Therefore, not all applications, particularly rare diseases or events are well suited to deep learning. A common problem that can arise during the training of a DNN (especially in the case of small datasets) is overfitting, which may occur when the number of parameters in the network is proportional to the total number of samples in the training set. Therefore, although the error on the training set is driven to a very small value, the errors for new data will be high. To avoid the overfitting problem and improve generalization, regularization methods, such as the dropout [40], are usually exploited during training.
- ➢ Another important aspect to take into account when deep learning tools are employed, is that for many applications the raw data cannot be directly used as input for the DNN. Thus, preprocessing, normalization or change of input domain is often required before the training. Finding the correct preprocessing of the data and the optimal set of hyperparameters can be challenging, since it makes the training process even longer, requiring significant training resources and human expertise, without which is not possible to obtain an effective classification model.
- ➢ The last aspect is that many DNNs can be easily fooled. For example, it is possible to add small changes to the input samples (such as imperceptible noise in an image) to cause samples to be misclassified. However, it is important to note that almost all machine learning algorithms are susceptible to such issues. Values of particular features can be deliberately set very high or very low to induce misclassification in logistic regression. Similarly, for decision tress, a single binary feature can be used to direct a sample along the wrong partition by simply switching it at the final layer. Hence in general, any machine learning models are susceptible to such manipulations. On the other hand, the work in [41] discusses the opposite problem.

To conclude, the healthcare informatics today is a human machine collaboration that may ultimately become a symbiosis in the future. As more data becomes available, deep learning systems can evolve and deliver where human interpretation is difficult. This can make diagnoses of diseases faster and smarter and reduce uncertainty in the decision making process. Finally, the last boundary of deep learning could be the feasibility of integrating data across disciplines of health informatics to support the future of precision medicine.

CONCLUSION

Deep learning has gained a central position in recent years in machine learning and pattern recognition. It has enabled the development of more data driven solutions in health informatics by allowing automatic generation of features that reduce the amount of human intervention in this process. This is advantageous for many problems in health informatics and has eventually supported a great leap forward for unstructured data such as those arising from medical imaging, medical informatics, and bioinformatics. Until now, most applications of deep learning to health informatics have involved processing health data as an unstructured source. A significant amount of information is equally encoded in structured data such as EHRs, which provide a detailed picture of the patient's history, pathology, treatment, diagnosis, outcome, and the like. In the case of medical imaging, the cytological notes of a tumor diagnosis may include compelling information like its stage and spread. This

information is beneficial to acquire a holistic view of a patient condition or disease and then be able to improve the quality of the obtained inference. In fact, robust inference through deep learning combined with artificial intelligence could ameliorate the reliability of clinical decision support systems.

Therefore, methodological aspects of NNs need to be revisited in this regard. Another concern is that deep learning predominantly depends on large amounts of training data. Many researchers have been encouraged to apply deep learning to any data-mining and pattern recognition problem related to health informatics in light of the wide availability of free packages to support this research. In practice, it is still questionable whether the large amount of training data and computational resources needed to run deep learning at full performance is worthwhile, considering other fast learning algorithms that may produce close performance with fewer resources, less parameterization, tuning, and higher interpretability.

Therefore, deep learning has provided a positive revival of NNs and connectionism from the genuine integration of the latest advances in parallel processing enabled by coprocessors. Nevertheless, a sustained concentration of health informatics research exclusively around deep learning could slow down the development of new machine learning algorithms with a more conscious use of computational resources and interpretability.

REFERENCES

[1] H. R. Roth et al., ‒Improving computer-aided detection using convolutional neural networks and random view aggregation,‖ IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1170–1181, May 2016.

[2] R. Fakoor, F. Ladhak, A. Nazi, and M. Huber, ‒Using deep learning to enhance cancer diagnosis and classification,‖ in Proc. Int. Conf. Mach. Learn., 2013, pp. 1–7.

[3] B. Alipanahi, A. Delong, M. T. Weirauch, and B. J. Frey, ‒Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning,‖ Nature Biotechnol., vol. 33, pp. 831–838, 2015.

[4] Y. LeCun, Y. Bengio, and G. Hinton, ‒Deep learning,‖ Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[5] J. L. McClelland et al., Parallel distributed processing. Cambridge, MA, USA: MIT Press, vol. 2, 1987.

[6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, ‒Learning representations by back-propagating errors,‖ in Neurocomputing: Foundations of Research, J. A. Anderson and E. Rosenfeld, Eds. Cambridge, MA, USA: MIT Press, 1988, pp. 696–699. [Online].

Available: http://dl.acm.org/ citation.cfm?id=65669.104451

[7] H. Shin, L. Lu, L. Kim, A. Seff, J. Yao, and R. M. Summers,

‒Interleaved text/image deep mining on a large-scale radiology database for automated image interpretation,‖ CoRR, vol. abs/1505.00670, 2015. Also available: http://arxiv.org/abs/1505.00670

[8] Z. Liang, G. Zhang, J. X. Huang, and Q. V. Hu, ‒Deep learning for healthcare decision making with emrs,‖ in Proc. Int. Conf. Bioinformat. Biomed., Nov 2014, pp. 556–559.

[9] E. Putin et al., ‒Deep biomarkers of human aging: Application of deep neural networks to biomarker development,‖ Aging, vol. 8, no. 5, pp. 1–021, 2016.

[10] L. Nie, M. Wang, L. Zhang, S. Yan, B. Zhang, and T. S. Chua,

‒Disease inference from health-related questions via sparse deep learning,‖ IEEE Trans. Knowl. Data Eng, vol. 27, no. 8, pp. 2107– 2119, Aug. 2015.

[11] R. Miotto, L. Li, B. A. Kidd, and J. T. Dudley, ‒Deep patient: An unsupervised representation to predict the future of patients from the electronic health records,‖ Sci. Rep., vol. 6, pp. 1–10, 2016.

[12] J. Futoma, J. Morris, and J. Lucas, ‒A comparison of models for predicting early hospital readmissions,‖ J. Biomed. Informat., vol. 56, pp. 229–238, 2015.

[13] S. Mehrabi et al., ‒Temporal pattern and association discovery of diagnosis codes using deep learning,‖ in Proc. Int. Conf. Healthcare Informat., Oct. 2015, pp. 408–416.

[14] Z. Che, S. Purushotham, R. Khemani, and Y. Liu, ‒Distilling knowledge from deep networks with applications to healthcare domain,‖ ArXiv e-prints, Dec. 2015.

[15] G.-Z. Yang, Body Sensor Networks, 2nd ed. New York, NY, USA: Springer, 2014.

[16] A. E. W. Johnson, M. M. Ghassemi, S. Nemati, K. E. Niehaus, D. A. Clifton, and G. D. Clifford, ‒Machine learning and decision support in critical care,‖ Proc. IEEE, vol. 104, no. 2, pp. 444–466, Feb. 2016.

[17] D. Wulsin, J. Blanco, R. Mani, and B. Litt, ‒Semi-supervised anomaly detection for eeg waveforms using deep belief nets,‖ in Proc. 9th Int. Conf. Mach. Learn. Appl., Dec. 2010, pp. 436–441.

[18] X. Jia, K. Li, X. Li, and A. Zhang, ‒A novel semi-supervised deep learning framework for affective state recognition on eeg signals,‖ in Proc. Int. Conf. Bioinformat. Bioeng., 2014, pp. 30–37.

[19] A. Wang, C. Song, X. Xu, F. Lin, Z. Jin, and W. Xu, ‒Selective and compressive sensing for energy-efficient implantable neural decoding,‖ in Proc. Biomed. Circuits Syst. Conf., Oct. 2015, pp. 1–4.

[20] L. A. Pastur-Romay, F. Cedr´ n, A. Pazos, and A. B. Porto-Pazos,

‒Deep oartificial neural networks and neuromorphic chips for big data analysis: Pharmaceutical and bioinformatics applications,‖ Int. J. Molecular Sci., vol. 17, no. 8, 2016, Art. no. 1313.

[21] L. Hood and S. H. Friend, ‒Predictive, personalized, preventive, participatory (p4) cancer medicine,‖ Nature Rev. Clin. Oncol., vol. 8, no. 3, pp. 184–187, 2011.

[22] M. K. Leung, A. Delong, B. Alipanahi, and B. J. Frey, ‒Machine learning in genomic medicine: A review of computational problems and data sets,‖ Proc. IEEE, vol. 104, no. 1, pp. 176–197, Jan. 2016.

[23] V. Marx, ‒Biology: The big challenges of big data,‖ Nature, vol. 498, no. 7453, pp. 255–260, 2013.

[24] R. Ibrahim, N. A. Yousri, M. A. Ismail, and N. M. El-Makky, ‒Multi- level gene/mirna feature selection using deep belief nets and active learning,‖ in Proc. Eng. Med. Biol. Soc., 2014, pp. 3957–3960.

[25] M. Khademi and N. S. Nedialkov, ‒Probabilistic graphical models and deep belief networks for prognosis of breast cancer,‖ in Proc. IEEE 14th Int. Conf. Mach. Learn. Appl., 2015, pp. 727–732.

[26] D. Quang, Y. Chen, and X. Xie, ‒Dann: A deep learning approach for annotating the pathogenicity of genetic variants,‖ Bioinformatics, vol. 31, p. 761–763, 2014.

[27] C. Angermueller, H. Lee, W. Reik, and O. Stegle, ‒Accurate prediction of single-cell dna methylation states using deep learning,‖ bioRxiv, 2016, Art. no. 055715.

[28] S. Kearnes, K. McCloskey, M. Berndl, V. Pande, and P. Riley, ‒Molecular graph convolutions: moving beyond fingerprints,‖ J. Comput. Aided Mol. Des., vol. 30, no. 8, pp. 595–608, 2016.

[29] B. Ramsundar, S. Kearnes, P. Riley, D. Webster, D. Konerding, and V. Pande, ‒Massively multitask networks for drug discovery,‖ ArXiv e-prints, Feb. 2015.

[30] M. Havaei, N. Guizard, H. Larochelle, and P. Jodoin, ‒Deep learning trends for focal brain pathology segmentation in MRI,‖ CoRR, vol. abs/1607.05258, 2016. [Online]. Available: http://arxiv.org/abs/1607.05258

[31] H. Greenspan, B. van Ginneken, and R. M. Summers, ‒Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique,‖ IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1153–1159, May 2016.

[32] J.-Z. Cheng et al., ‒Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in us images and pulmonary nodules in ct scans,‖ Sci. Rep., vol. 6, 2016, Art. no. 24454.

[33] J. Shan and L. Li, ‒A deep learning method for microaneurysm detection in fundus images,‖ in Proc. IEEE Connected Health, Appl., Syst. Eng. Technol., 2016, pp. 357–358.

[34] T. Brosch et al., ‒Manifold learning of brain mris by deep learning,‖ in Proc. MICCAI, 2013, pp. 633–640.

[35] A. Mansoor et al., ‒Deep learning guided partitioned shape model for anterior visual pathway segmentation,‖ IEEE Trans. Med. Imag., vol. 35, no. 8, pp. 1856–1865, Aug. 2016.

[36] T. Huang, L. Lan, X. Fang, P. An, J. Min, and F. Wang, ‒Promises and challenges of big data computing in health sciences,‖ Big Data Res., vol. 2, no. 1, pp. 2–11, 2015.

[37] L. Zhao, J. Chen, F. Chen, W. Wang, C.-T. Lu, and N. Ramakrishnan, ‒Simnest: Social media nested epidemic simulation via online semi- supervised deep learning,‖ in Proc. IEEE Int. Conf. Data Mining, 2015, pp. 639–648.

[38] V. R. K. Garimella, A. Alfayad, and I. Weber, ‒Social media image analysis for public health,‖ in Proc. CHIConf. Human Factors Comput. Syst., 2016, pp. 5543–5547.

[39] E. Horvitz and D. Mulligan, ‒Data, privacy, and the greater good,‖ Science, vol. 349, no. 6245, pp. 253–255, 2015.

[40] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, ‒Dropout: a simple way to prevent neural networks from overfitting.‖ J. Mach. Learn. Res., vol. 15, no. 1, pp. 1929–1958, 2014.

[41] C. Szegedy et al., ‒Intriguing properties of neural networks.‖ CoRR, vol. abs/1312.6199, 2013.