iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** XI **Month of publication:** November 2020

DOI: https://doi.org/10.22214/ijraset.2020.32304

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue XI Nov 2020- Available at www.ijraset.com

A Review of Deep learning in Bioinformatics

Rachna Kumari[1], Amita Verma[2], Sunita[3]
[1, 2, 3]*Department of Computer Science & Engineering, Guru Jambheshwar University of Science & Technology, Hisar, India*

Abstract: Bioinformatics is a new discipline related to the solution of molecular level biological problems by different techniques like applied mathematics, computer science, statistics and artificial intelligence. Bioinformatics means extraction of knowledge or meaningful data from the large amount of biological database. This paper described deep learning technique use in bioinformatics.

I. INTRODUCTION

Bioinformatics means extraction of knowledge or meaningful data from the large amount of biological database. Bioinformatics is a new discipline related to the solution of molecular level biological problems by different techniques like applied mathematics, computer science, statistics and artificial intelligence. Bioinformatics is being used in analyzing genomes protein sequences, three-dimensional modelling of bio molecules and biologic systems etc. Different biological problems considered within the scope of bioinformatics fall into main tasks which are given below:

A. Alignment and comparison of DNA, RNA, and protein sequences.
B. Gene finding and promoter identification from DNA sequences.
C. Interpretation of gene expression and microarray data.
D. Gene regulatory network identification.
E. Construction of phylogenetic trees for studying evolutionary relationship.
F. Protein structure prediction and classification.
G. Molecular design and molecular docking.

First section of this paper gives introduction to Bioinformatics field. Then in second section of the paper described deep learning in Bioinformatics field.
Third section discussed literature review and fourth section described conclusion and future scope of paper. Then last section contain references.

II. DEEP LEARNING IN BIOINFORMATICS AND MEDICAL DIAGNOSIS

Now a day's deep learning is used in more and more domains and provides encouraging results in almost all the areas. The success of deep learning in the domain of image recognition, natural language processing and computer vision lead to its use in health and medical fields.
Automatic feature extraction capability of deep learning plays a vital role in most emerging fields such as bioinformatics and medical diagnosis for analyzing, visualization and manipulation of biological data. For example in medical imaging, deep learning techniques are used to extract the feature that is very difficult and complex by using descriptive means. Deep learning is successfully used for classification of tissue and cell classification and also widely used for diagnosing diseases such as cancer, tumour and diabetes etc. [1].

III. LITERATURE REVIEW

Riya Roy and Swapna Sasi [2] proposed frameworks for classifying of each kind of white blood cells. They used multi class support vector machine classification and convolutional neural networks. Deep learning method identified the Neutrophils, Lymphocytes, Monocytes, Eosinophil and Basophils variation for diagnosis diseases. SVM is used to discover the level of malignant cell in blood and leukocytes for different age domain.
R. Shimizu *et al.* [3] developed a model of deep learning for diagnosing lung cancer. Model developed achieved accuracy around 90% in diagnosing whether a patient have lung cancer or not. This model was very flexible as input was generated from human urine by Gas Chromatography Mass Spectrometer and collection of urine is very easy and these are not harmful to human. For building model Stacked Autoencoder architecture of deep learning was used. They used this model for pre-diagnosis test it can further enhanced for complete diagnosis test and can also use it in diagnosing other diseases.

©IJRASET: All Rights are Reserved

Hue Cheng *et al.* [4] proposed a new deep learning model particle and cell identification for Live-cell time-lapse images. They used RCNN architecture and tenserflow in Linux for implementing this model. Results of this model provide more accuracy then particle Cell relation mining method. Advantage of this method is that it can identify cell particle in much less time than particle cell mining method. Disadvantage is that it cannot detect some particles that generally have small or their brightness is too dark.

P. Kiran Sree and P. S. V Srinivasa Rao [5] developed a novel unsupervised classifier using deep learning for finding gene in a given DNA sequence. Model is built on Fickett and Taung data sets and also uses hybrid cellular automata. This model provides accuracy more than 97.45%. Time taken by this model is much less than traditional techniques. Further this model can also be applied in electric and electronic and other field also.

O. Yadav *et al.* [6] applied Deep Learning technique to classify X-ray images of potential Tuberculosis patients. Different learning rate enhancement techniques were used to find better results. They used FastAi as the main tool and PyTorch library for implementation. They achieved an accuracy of 94.89% on the augmented images classification. Limitation of this model is that this is trained only for china data sets. In future this model can be applied to other datasets.

A. Akay and H. Hess [7] provided an overview of the deep learning, its applications and needs to understand swarm behaviour of molecular shuttles. Deep learning applications in fields such as Diagnosis of Diabetic Retinopathy, Microrobot Localization, Computational Biology, and Bioinformatics were discussed. Various uses of deep learning in current trends like drug design, Multi-platform Data, Intelligent Pre-Diagnosis are explained. Use of RCNN in analyzing the behaviour of molecular shuttles is provided. We can apply deep learning techniques in fields like displacing the Workforce through Automation, complex datasets.

Z. Luo *et al.* [8] developed NgramCNN for classification of graphs based on common substructure patterns and their latent relationships in the collection of graphs. NgramCNN was a noval convolutional neural network that consists of three features. In First phase, it transforms raw graph object into a sequence of n-gram blocks connected through overlapping regions by using the concept of n-gram block. Second, it extracts local patterns and connectivity features hidden by using convolution operation of n-gram normalization and diagonal. And in third phase but not the least, it iteratively learns to extract global patterns based on the local patterns. Ngram is implemented through Tenserflow.

G. Yolcu *et al.* [9] presented a deep learning approach for recognition of facial expression. The aim of study is to design a deep learning architecture that can detect facial expression to assist diagnosis and monitoring of neurological disorders. The proposed system achieved accuracy of 93.43% for facial expression recognition. This model consist a cascade of two CNN architecture. The first part of the architecture was trained to segment facial components and second part was trained to perform classification of various facial expressions.

N. N. Leite *et al.* [10] presented deep convolutional autoencoder for reducing the effort of projecting denoising filters. This method filtered EEG data without extended knowledge. Two types of noise, originated from eye blink and from jaw clenching were used for experiment. Result show that deep convolutional approach provides better accuracy than traditional filtering method. This method is only trained for initial stages of testing we can train it for further stages of testing. RNN technique can also be implemented for filtering noise in EEG. M. Motlagh *et al.*[11] applied various Inception and ResNet deep learning approaches for diagnosing breast cancer. They developed a fine-tuned pre-trained deep neural networks that can detect four cancer types including breast, bladder, lung and lymphoma with high accuracy using ResNet pretrained model. Accuracy can further extended by applying different deep learning semantic segmentation algorithms. The framework principals can be extended in the field of computer-assisted diagnosis using medical image and pathological analysis.

A. Khan and K. Wang [12] proposed a deep learning based scoring system for analyzing gene sequencing data on personal genomes. Model developed, integrate contributions from known brain expression quantitative trait locus, coding, non-coding, structural variants, and enhancer/promoter peaks from PsychENCODE. The technique was found helpful in identifying variants with major effect sizes for mental disorders. This technique can be further used for detecting susceptibility mutations for mental diseases specifically.

Y. Ma and C. Yan [13] applied concurrent neural network (CNN) for identifying RNA-binding residues on the proteins. Dataset for performing experiment were taken from SWISS database. Using CNN accuracy up to 97.2% was achieved. Results obtained were also compared with support vector machine and Random Forest and it was found that CNN provides better accuracy than them. Further it can also apply to other protien-protien network.

Pablo Guillen and Jerry Ebalunode [14] applied a deep learning algorithm based on multilayer perception for classification of cancer. Library H20 was used for implementation purpose. Datasets are taken from 11_Tumors database for training and testing purpose. This approach can efficiently classify kidney cancer, lung cancer and liver cancer. Multilayerpreceptron can also be in field of tumours cell, leading to the discovery of selective drugs.

©IJRASET: All Rights are Reserved

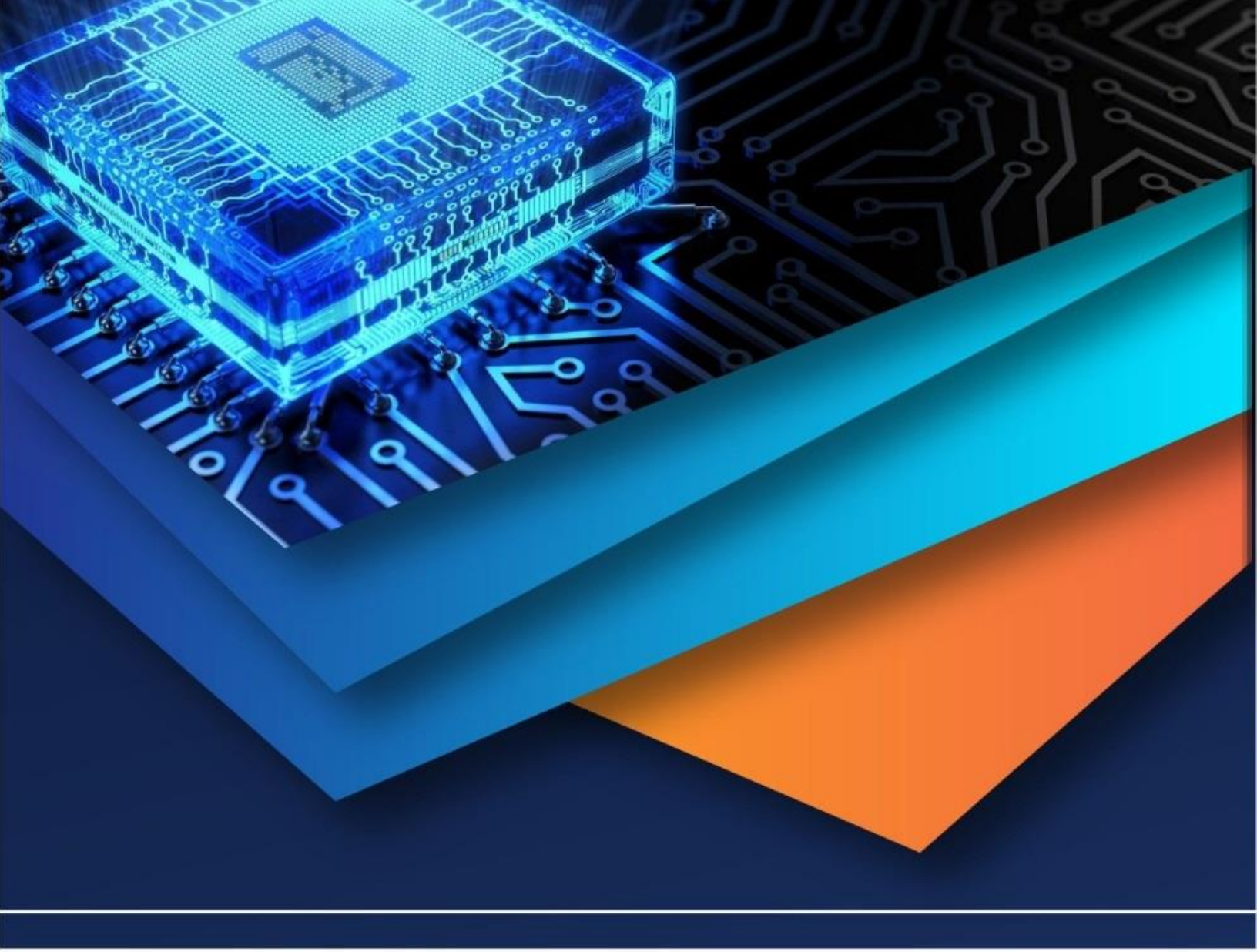
IV. CONCLUSION AND FUTURE SCOP

Deep learning is an attractive technique for analyzing biological data and images of bioinformatics fields. It has shown a great potential in promoting medical industry. From the literature review, it is observed that there is huge scope of work in the field of bioinformatics with the help of deep learning techniques.

REFERENCES

[1] D. Ravi, F. Deligianni, M. Berthelot and J. Perez, "Deep Learning for Health Informatics", IEEE Journal of Biomedical and Health Informatics, Vol. 21, pp. 1-18, 2017.
[2] Riya Roy and Swapna Sasi, "Classification of WBC Using Deep Learning for Diagnosing Diseases", International Conference on Inventive Communication and Computational Technologies, IEEE, pp. 1634-1638, 2018.
[3] R. Shimizu, S. Yanagawa, Y. Monde and H. Yamagishi, "Deep Learning Application Trial to Lung Cancer Diagnosis for Medical Sensor Systems", IEEE, pp. 191-192, 2016.
[4] Hui Cheng, Cheng Wu and Wei Chen, "Using Deep Learning to Identify Cell and Particle in Live-Cell Time-lapse Images", IEEE International Conference on Bioinformatics and Biomedicine, pp. 1327-1331, 2018.
[5] P. Kiran Sree and P. S. V Srinivasa Rao, "CDLGP: A Novel Unsupervised Classifier using Deep Learning for Gene Prediction", IEEE International Conference on Power, Control, Signals and Instrumentation Engineering, pp. 2811-2813, 2017.
[6] O. Yadav, K. Passi and C. Jain, "Using Deep Learning to Classify X-ray Images of Potential Tuberculosis Patients", IEEE International Conference on Bioinformatics and Biomedicine, pp. 2368-2375, 2018.
[7] A.Akay and H. Hess, "Deep Learning: Current and Emerging Applications in Medicine and Technology", IEEE, pp. 1-15, 2018.
[8] Z. Luo, L. Liu, J. Yin, Y. Li and Z. Wu, "Deep Learning of Graphs with Ngram Convolutional Neural Networks",IEEE 34th International Conference on Data Engineering, pp. 1791-1792, 2018.
[9] G. Yolcu, K. Palaniappan, T. E. Lever and F.Bunyak, "Deep Learning-based Facial Expression Recognition for Monitoring Neurological Disorders", IEEE International Conference on Bioinformatics and Biomedicine, pp. 1652-1657, 2017.
[10] N. N. Leite, E. C. Gurjao and L. R. Veloso, "Deep Convolutional Autoencoder for EEG Noise Filtering", IEEE International Conference on Bioinformatics and Biomedicine, pp. 2605-2612, 2018.
[11] M. Motlagh, M. Jannesari, M. Totonchi and I. Hajirasouliha, "Breast Cancer Histopathological Image Classification: A Deep Learning Approach", IEEE, pp. 1-8, 2018.
[12] A. Khan and K. Wang, "A deep learning based scoring system for prioritizing susceptibility variants for mental disorders", IEEE International Conference on Bioinformatics and Biomedicine, pp. 1698-1704, 2017.
[13] Y. Ma and C. Yan, "A Concurrent Neural Network (CNN) Method for RNA-binding Site Prediction", IEEE 8th Joint International Information Technology and Artificial Intelligence Conference, pp. 567-570, 2018.
[14] Pablo Guilin and Jerry Ebalunode, "Cancer Classification Based on Microarray Gene Expression Data Using Deep Learning", International Conference on Computational Science and Computational Intelligence IEEE, pp. 1403-1405, 2016.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)