International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 02 | February - 2023 Impact Factor: 7.185 ISSN: 2582-3930

A Contemporary Technique for Lung Disease Prediction using Deep Learning

Janhvi Kawale[1] ***Computer Science & Engineering (Priyadarshini JL college of engineering)*** **Nagpur, India** **@gmail.com**	**Rashmi Tiwari[2]** ***Computer Science & Engineering (Priyadarshini JL college of engineering)*** **Nagpur, India** **@gmail.com**	**Angha Ambrkar[3]** ***Computer Science & Engineering (Priyadarshini JL college of engineering)*** **Nagpur, India** **@gmail.com**
Tejal Bhoyar[4] ***Computer Science & Engineering (Priyadarshini JL college of engineering)*** **Nagpur, India** **@gmail.com**	**Ankit Bhujade[5]** ***Computer Science & Engineering (Priyadarshini JL college of engineering)*** **Nagpur, India** **@gmail.com**	**ER. Mausami Dagwar6** ***Computer Science & Engineering (Priyadarshini JL college of engineering)*** **Nagpur, India** **@gmail.com**

***Abstract*—** The recent developments of deep learning support the identification and classification of lung diseases in medical images. Hence, numerous work on the detection of lung disease using deep learning can be found in the literature. This paper presents a survey of deep learning for lung disease detection in medical images. There has only been one survey paper published in the last five years regarding deep learning directed at lung diseases detection. However, their survey is lacking in the presentation of taxonomy and analysis of the trend of recent work. The objectives of this paper are to present a taxonomy of the state-of-the-art deep learning based lung disease detection systems, visualise the trends of recent work on the domain and identify the remaining issues and potential future directions in this domain. Ninety-eight articles published from 2016 to 2020 were considered in this survey. The taxonomy consists of seven attributes that are common in the surveyed articles: image types, features, data augmentation, types of deep learning algorithms, transfer learning, the ensemble of classifiers and types of lung diseases. The presented taxonomy could be used by other researchers to plan their research contributions and activities. The potential future direction suggested could further improve the efficiency and increase the number of deep learning aided lung disease detection applications.

Keywords: Deep Learning, lung disease detection.

I. Introduction

The technological development trend in software engineering has been improving, where the design of software began move from the desktop to the web. Nowadays, many IDE (Integrated Development Environment) application has been made, such as Eclipse, Visual Studio, etc. but IDEs which based on desktop still have significant disadvantages such as long time for configuration and installing the plug-in needed for IDE to run the project. This problem could be a huge waste of time when there are many devices that have to be configured Many software applications have been run in the cloud, and use a web browser as a user interface that allows ubiquitous access, instant collaboration, and avoid installation and configuration on desktop computers one of the technologies used for instant collaboration is single IDE (like pair programming) Pair programming is the practice of having two programmers' access and work on the same code in a single development environment in pair programming programmers have the abilities to create, edit and delete source code in real time. Therefore, it needs to make an application that can improve performance while writing program such as real-time collaboration, create, execute and display the result of the program using terminal. Online compiler that is development environment save lot of time as well as user can access it from any device without installing ide and other requirement.

Coide is designed for coding practice and improvement, Coide offers code challenges and courses aimed at helping you prepare for job interviews. Coide provides more than 50+ practice problems to its

programmer community. Individual challenges in topics like algorithms, database structures, and dynamic programming accompany entire study plans. The server based code editor is used to run multiple programming languages, many programmers who want to edit the source code urgently might not access convenient resource without installing any application on the computer or notebook in our project programmer can compile and run source code via web browser and the code will be done at the server side. Then the output from the compilation will be displayed at the browser of client side. server based code editor provides facility to run on the small resources such as pc's, tabs, androids, notebook's and laptops. programmers need to have atleast one computer in order to edit program source code. if programmer wants to run multiple languages then the programmer needs to install the software of that particular languages but by using server based code editor if programmer wants to run multiple programming languages then they can run all languages in one platform means server based code editor provider same platform for multiple programming languages that's why programmer doesn't need to install software of that particular programming languages which they to enter a program language source statements or to create a documents such as technical manuals.

II. LITERATURE REVIEW

A Review on Server Based Code Editor In this paper the server based code editor is created to run multiple programming as well as web based languages, and provides the features create file, compile, and run file were created to make this editor complete. In addition, the advantage of sever based code editor the programmers in the business or organization can urgently write or modify program source code without any specific physical computers or without installing the software of programmes. web based languages HTML, CSS, JavaScript, and programming language Java, python, c, c++, c#, can be written within this editor. Sr. No. Paper Name and Author Study Title Research method used Key feature of the study 1. Boubahi and Alrazgan 2015 Investigation IT faculty resistance to learning and examination system adoption using latent variables in an acceptance technology model Survey Use of learning management system at the user end. 2. Kuijk et al. 2015 Usability in product development practice; an exploratory case study comparing four markets Interview Attitude towards usability 3. Medina and Gamboa 2015 Usability evaluation by experts of a learning Management System Questionnaire (a list of questions that are answered by many people.) Evaluates the usability of locall

IV. METHODOLOGY

1 Data Augmentation : Data augmentation is a strategy that utilized in AI to improve model precision, speculation, and to control overfitting [12]. The image augmentation procedures used in this study are acknowledged by the Keras library in Python [13]. In image augmentation, rotation, width and height shift, shear, zoom, horizontal flip and fill operations are implemented. The rotation range fix to 40 and width and height shift range, the sheer range and zoom range fix to 0.2 and horizontal flip fix to true. The augmentation was only applied to training dataset and on the other hand, the test dataset used without augmentation.

2 Convolutional Neural Network (CNN) In this study, AlexNet was used from deep learning models. AlexNet was fundamentally planned by Alex Krizhevsky. It was disclosed with Ilya Sutskever and Krizhevsky's doctoral counsellor Geoffrey Hinton, and is a Convolutional Neural Network or CNN. In the wake of contending in ImageNet Large Scale Visual Recognition Challenge, AlexNet shot to popularity. It accomplished a top-5 mistake of 15.3%. It was trained by 1.2 million images [14]. From three deep learning methods, AlexNet performed better to detect lung cancer [9]. The AlexNet architecture formed by five convolutional layers, some of these are ended by maximum pooling layers and afterwards three entirely joint layers lastly a 1000-way softmax classifier.

3 Optimization The primary reason for optimization techniques is to refresh the loads at each level until the best training in CNN design is figured it out. Every technique plays out the update measure with its own calculation. In this study, we used stochastic gradient descent (SGD) optimization technique in Keras [17]. Stochastic gradient descent (SGD) interestingly plays out a boundary update for each training instance x and label y [15][16].

4 Machine Learning Model Machine learning model is employed to separate the classes. In this experiment, we used the KNearest Neighbor algorithm (K-NN) to classify lung cancer. When it comes to classification, kNN machine learning algorithm is going to set which objects goes under which class by investigating the features of the sample. According to the maximum votes of the neighbors, the sample is sent to the appropriate class. The kNN performed outstanding from various machine learning algorithms [9]. In this experiment,

we observed the value of k in the range of 1 to 7 discussed in figure 4. And finally, the value of k is set to 3.

5 Proposed Methodology The dataset applied in this study is a public dataset from SPIE-AAPM. Firstly, the data augmentation technique applied to training dataset and the test data set were kept as it was. Then the AlexNet deep learning architecture applied to the augmented dataset. Parameters for all the layers of AlexNet kept as mentioned in table1. And then stochastic gradient descent (SGD) optimization technique applied for feature optimization. After that relative features were turned out from the last fully-connected layer of AlexNet. Lastly, extracted features from AlexNet are applied to kNN with cross validation for separating cancerous and non-cancerous images

V. CONCLUSION

As time goes on, more works on lung disease detection using deep learning have been published. However, there was a lack of systematic survey available on the current state of research and application. This paper is thus produced to offer an extensive survey of lung disease detection using deep learning, specifically on tuberculosis, pneumonia, lung cancer and COVID-19, published from 2016 to September 2020. In total, 98 articles on this topic were considered in producing this survey.

To summarise and provide an organisation of the key concepts and focus of the existing work on lung disease detection using deep learning, a taxonomy of state-of-the-art deep learning aided lung disease detection was constructed based on the survey on the works considered. Analyses of the trend on recent works on this topic, based on the identified attributes from the taxonomy, are also presented. From the analyses of the distribution of works, the usage of both CNN and transfer learning is high. Concerning the trend of the surveyed work, all the identified attributes in the taxonomy observed, on average, a linear increase over the years, with an exception to the ensemble attribute. The remaining issues and future direction of lung disease detection using deep learning were subsequently established and described. Four issues of lung disease detection using deep learning were identified: data imbalance, handling of huge image size, limited available datasets and high correlation of errors when using ensemble techniques. Four potential works for lung disease detection using deep learning are suggested to resolve the identified issues: making datasets available to the public, usage of cloud computing, usage of more features and usage of the ensemble.

To conclude, investigating how deep learning was employed in lung disease detection is highly significant to ensure future research will concentrate on the right track, thereby improving the performance of disease detection systems. The presented taxonomy could be used by other researchers to plan their research contributions and activities. The potential future direction suggested could further improve the efficiency and increase the number of deep learning aided lung disease detection applications.

VI. REFERENCES

1. Waheed, A.; Goyal, M.; Gupta, D.; Khanna, A.; Al-Turjman, F.; Pinheiro, P.R. CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection. *IEEE Access* **2020**, *8*, 91916–91923. [Google Scholar] [CrossRef]

1. Muhammad Azizi Mohd Ariffin, “Data Leakage Detection in Cloud Computing Platform”, August 2019, International Journal of Advanced Trends in Computer Science and Engineering 8(1.3):400-408 DOI:10.30534/ijatcse/2019/7081.32019

2. M.Sai Charan Reddy, T. Venkata Satya Yaswanth, T. Gopal, L. Raji, Dr. K.Vijaya, “DATA LEAKAGE DETECTION USING CLOUD COMPUTING”, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, Volume: 06 Issue: 03, Mar 2018

3. N. Kumar, V. Katta, H. Mishra and H. Garg, "Detection of Data Leakage in Cloud Computing Environment," 2021 International Conference on Computational Intelligence and Communication Networks, 2021, pp. 803-807, doi: 10.1109/CICN.2014.172.

4. R. Naik and M. N. Gaonkar, "Data Leakage Detection in cloud using Watermarking Technique," 2019 International Conference on Computer Communication and Informatics

(ICCCI), 2019, pp. 1-6, doi: 10.1109/ICCCI.2019.8821894.

5. M. Xu, J. Zhao, Q. Ye, W. Ding, X. Zhu and Q. Zhang, "Cloud Computing Methods for Leaking Detection of Halogen Pipeline Based on Distributed Terminals," 2020 International Conference on Robots & Intelligent System (ICRIS), 2020, pp. 561-563, doi: 10.1109/ICRIS52159.2020.00143.

6. Lord, N. (2019). Data Protection: Data In transit vs. Data At Rest. Retrieved 24 July 20,from https://digitalguardian.com/blog/data-protection-data-in-transit-vs-data-at-rest

7. El Sharif Karrar, D., & Idris Fadl, M. (2018). Security Protocol for Data Transmission in Cloud Computing. International Journal Of Advanced Trends In Computer Science And Engineering, 7(1), 1-5. doi: 10.30534/ijatcse/2018/01712018

8. Yasmin R., Memarian, M.R., Hosseinzadeh, S., Conti, M., & Leppänen, V. (2017). Investigating the Possibility of Data Leakage in Time of Live VM Migration. In: Dehghantanha A., Conti M., Dargahi T. (eds) Cyber Threat Intelligence. Advances in Information Security, vol 70. Springer, Cham.

9. Herrera Montano, I., García Aranda, J.J., Ramos Diaz, J. *et al.* Survey of Techniques on Data Leakage Protection and Methods to address the Insider threat. *Cluster Comput* **25**, 4289–4302 (2022). https://doi.org/10.1007/s10586-022-03668-2