International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8, Issue-8

Integrated Clinician Decision Supporting System for Pneumonia and Lung Cancer Detection

Venkata Tulasiramu Ponnada, S.V.Naga Srinivasu

Abstract*: Our research work proposes a system "Integrated Clinician Decision Supporting System for Pneumonia and Lung Cancer Detection (ICDSSPLD)" that can detect Pneumonia and Lung cancer. ICDSSPLD uses CT scan images and leverages the convolution neural network (CNN). The integrated system works for both pneumonia and lung cancer and the end user for this system are clinician and or patient. Our research work objectives are as follows. Propose an integrated system for Pneumonia and Lung Cancer Detection; Implement best in class Edge AI system for Pneumonia detection; Implement best in class Edge AI system for Lung Cancer detection; Implement integrated edge AI system for Pneumonia and Lung Cancer Detection.The present work proposes an integrated system for Pneumonia and lung cancer detection i.e. the first research objective.

***Index Terms*: Pneumonia Detection, Lung Cancer Detection, Edge AI System, CNN, deep learning, neural networks**

I. INTRODUCTION

Any method or a system which is alleviating the pain, restoring the health and extending life is considered as a promoter for quality of life. The motto of ICDSSPLD falls under the same category. Matter of the fact that - lung cancer is a life-threatening disease and also widely known as critical cancer. 13% of all new cancer diagnosis was categorized as lung cancer, and 24% of cancer deaths are due to lung cancer, as per SEER Cancer Statistics review [1]. One out of 16 American people is diagnosed with lung cancer and a new diagnosis [1] nearly every two minutes. The other lung organ disease considered in our research is Pneumonia. Pneumonia [2] is equally dangerous and accounts for toddlers, accounting for 16% of all deaths of children fewer than five years. Deep learning techniques motivated us to leverage the best in class technique Convolution Neural Network (CNN) to come up with a unique and integrated system to detect lung cancer and pneumonia. In our research, we have used the data sets for pneumonia and lung cancer from the LIDC-IDRI [3] and Mendeley [4].

In this paper, we are proposing an integrated system for most common lung organ diseases - pneumonia and lung cancer. Also, we discuss the preliminary research results in this paper. Further, we will come up with a best in class edge AI system for Pneumonia and lung cancer separately and leverages the results and fine-tune the ICDSSPLD.

Here, we are proposing ICDSSPLD-CNN as an advanced CNN architecture to address the ICDSSPLD. The ICDSSPLD-CNN comprised of 3 sets of sandwiches. Each sandwich comprises two convolution layers and one max pooling layer. This multi-layer CNN architecture enhances the detection accuracy of the diseases. To come up with the ICDSSPLD-CNN, we have exercised 16 combinations of CNN layers. Results of all these combinations are analyzed to finalize the combination. We have used the LIDC-IDRI [3] and Mendeley [4] data sets for ICDSSPLD-CNN validation.

We summarize the CNN concepts in section 2. In section 3, we discuss our proposed deep learning architecture ICDSSPLD-CNN. Section 4 is about preliminary research results, and in section 5 conclude with future work.

II. CONVOLUTION NEURAL NETWORK (CNN)

The convolutional neural network (CNN) is a deep learning neural network dedicated to image classification [6][7]. CNN, also called as ConvNets has shown outstanding [13] performance in natural image classification.

To identify the normal versus abnormal medical image, the 2D and 3D structures of an organ under study is critical. CNN is the best in class [14][15] method to perform image analysis. CNN performs image classification, localization, detection, segmentation and registration for medical image analysis.

CNN preserves local image attributes during the image compression process, in this context also described dimension reduction. This unique characteristic keeps CNN as apart from the other deep learning algorithms. Another key CNN characteristic is the processing of 2D and 3D images with slight modifications. In the health care domain, the X-rays are 2D images, and CT/MRI scans are 3D images [11][12].

Supervised machine learning algorithms and unsupervised machine algorithm are suitable methods for medical image analysis. In general, supervised machine algorithms need significant amounts of training data. CNN is classified under supervised machine learning algorithms. The examples for unsupervised algorithms are Generative Adversarial Networks (GANs), Restricted Boltzmann Machines (RBMs) and Deep Belief Networks (DBNs.

The input for CNN is n*n pixel image segment. But the classical methods take feature vectors as input. In CNN, the image classification happens based on the appearance of the image segment. CNN considers the highest score of the class as CNN output.

The current machine learning disease detection/prediction solutions are specific to the disease. But the present research work came up with an integrated solution for lung diseases pneumonia and lung cancer.

In our research work, the input for CNN is pneumonia and lung cancer CT scan

Revised Manuscript Received on June 12, 2019

Venkata Tulasiramu Ponnada, Research Scholar, Acharya Nagarjuna University, Nagarjuna Nagar, Guntur, Andhra Pradesh 522510, India.

Dr.S.V.Naga Srinivasu, Professor, Computer science and engineering, Narasaraopeta Engineering College, Narasaraopet, Andhra Pradesh 522601, India.

image segments. The outputs are four scores of pneumonia and non-pneumonia, lung cancerous and non-lung cancerous [8][9][10].

The CNN architecture looks like the visual cortex of the brain, which processes visual information. The input image segment is passing through by consecutive layers to derive the feature vector [18].

III. ICDSSPLD

Our proposed system ICDSSPLD comprised of ICDSSPLD engine, Input image system, input enhancement system and User alarm system. Fig. 1 describes the ICDSSPLD system Black box view. The input patient image is captured using the Input Image System. The captured input image acts as the input for input image enhancement system. The input image enhancement system leverages image enhancement techniques. The enhanced input image is the input for ICDSSPLD engine. The ICDSSPLD engine acts as an intermediate layer between the input image system and the ML system. Fig. 2 describes the interaction between different components of ICDSSPLD.

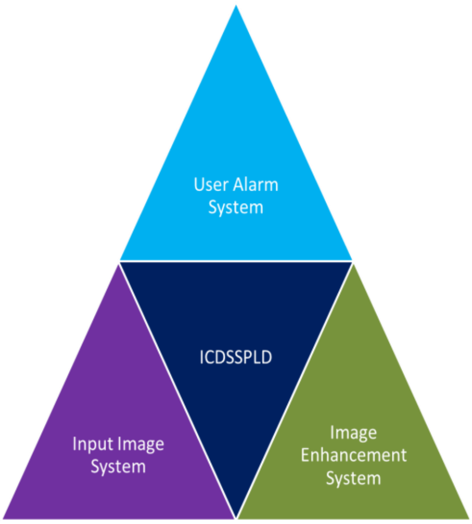

Fig.1 ICDSSPLD System Black Box View

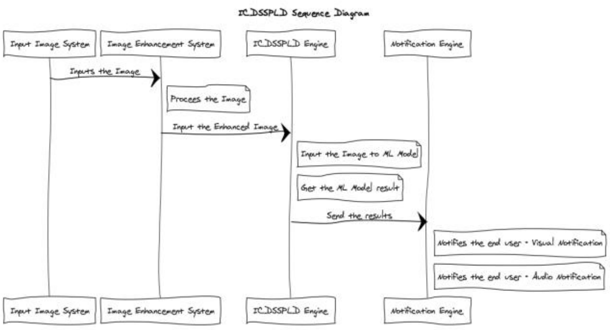

Fig.2 ICDSSPLD Sequence Diagram

The Pneumonia and lung cancer data set has been taken from LIDC-IDRI [3] and Mendeley [4] data sets. The proposed ICDSSPLD-CNN is used to train the models for pneumonia and lung cancer detection. Fig. 3 and Fig. 4 describe the model generation process for pneumonia and lung cancer models.

Fig.3 Pneumonia Data model generation

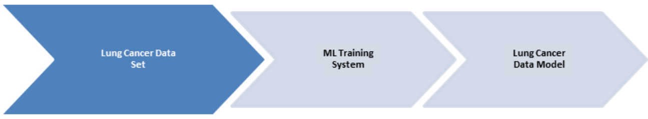

Fig.4 Lung Cancer Data model generation

ICDSSPLD engine routes the given enhanced image to the associated model, i.e. Pneumonia model and Lung cancer model respectively. ICDSSPLD engine sends the ICDSSPLD-CNN results to user alarm engine. Fig. 5 describes the ICDSSPLD system diagram.

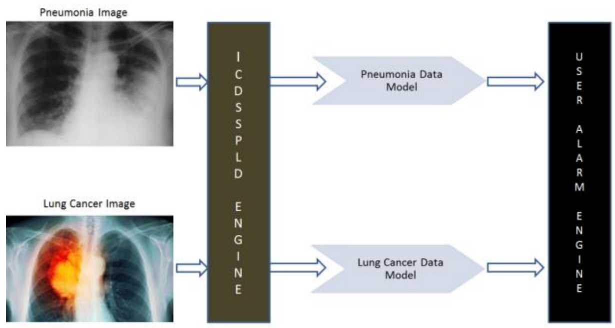

Fig.5 ICDSSPLD System Diagram

User alarm engine sends the notifications to the end user (patient/clinician) in three forms, i.e. visual notification, audio notification and email notification. Fig. 6 describes the user alarm system and different forms of user notification.

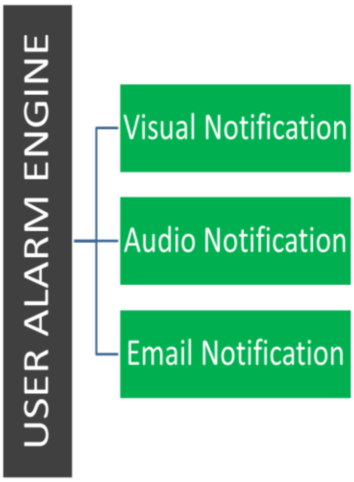

Fig.6 User Alarm Engine

Fig. 7, Fig. 8, Fig. 9 and Fig. 10 describes the ICDSSPLD-CNN. The NN-SVG architecture [5] tools are used to present the ICDSSPLD-CNN. Three sets of convolution layer-convolution layer-Max-Pooling layers combination is used to enhance the detection accuracy.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

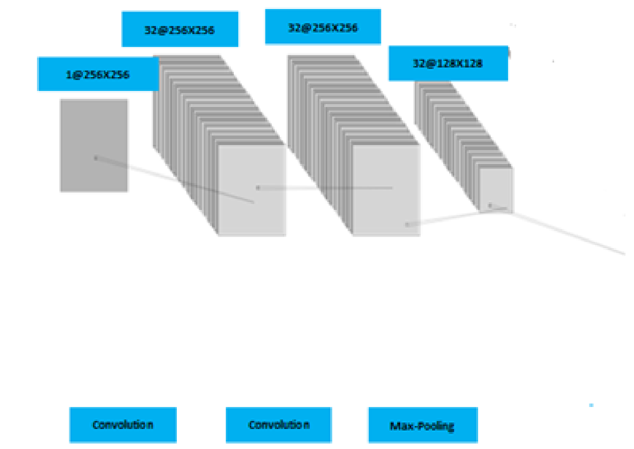

Fig.7 ICDSSPLD-CNN Set-I Layers

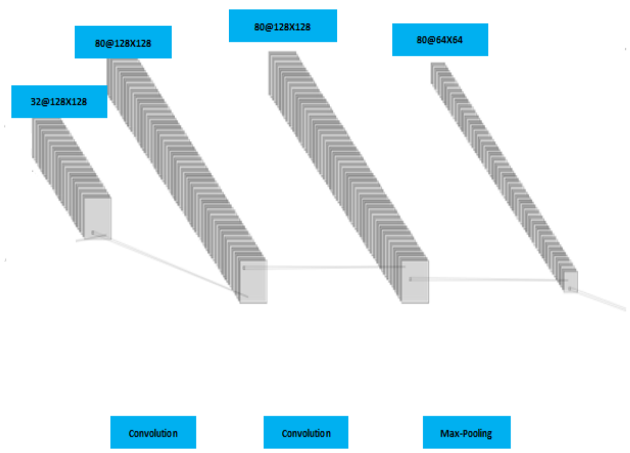

Fig.8 ICDSSPLD-CNN Set-II Layers

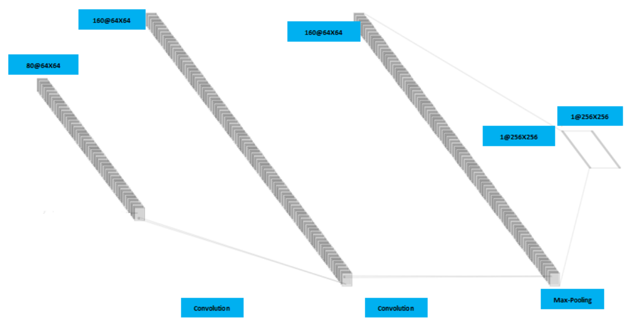

Fig.9 ICDSSPLD-CNN Set-III Layers

The ICDSSPLD-CNN used as a linear classifier. ICDSSPLD-CNN uses ReLU activation, Adam Optimizer and weighted soft max cross entropy loss. Smaller regions of actual input image feed are used to enhance the detection results,

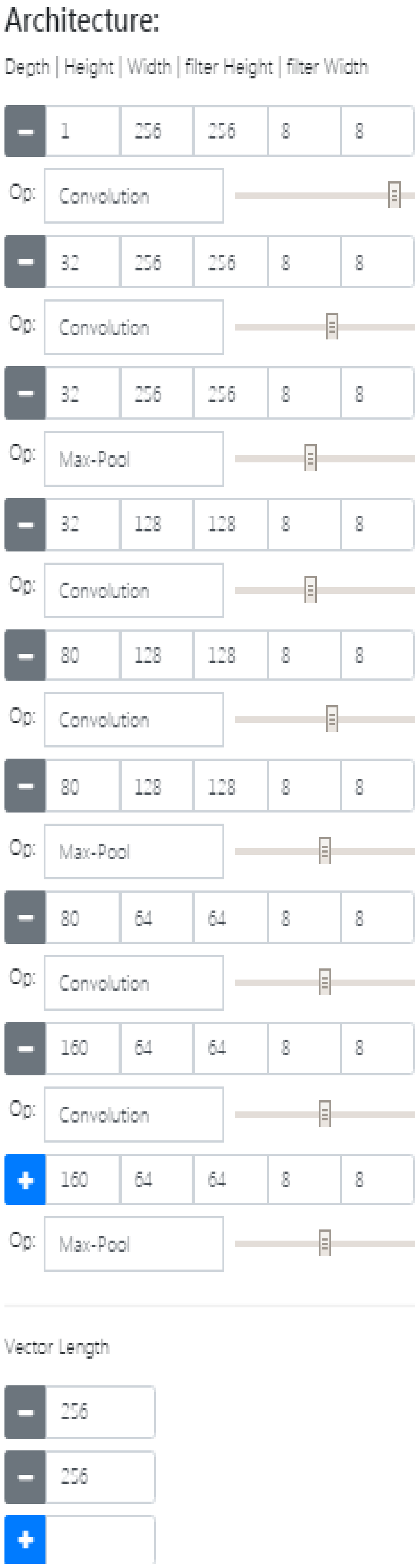

Fig.10 ICDSSPLD-CNN architecture schematics

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The first set of ICDSSPLD-CNN layers are described as below.

i. The input image is 256x256x1

ii. The first convolution layer (Params: 3x3x32 /Activation: ReLU) output is 256x256x32

iii. The second convolution layer (Params: 3x3x32 /Activation: ReLU) output is 256x256x32

iv. Max Pool layer(Params: 2x2, stride 2) output is 128x128x32

v. The output of one layer acts as input for the next CNN layer

vi. The output of the first set of ICDSSPLD-CNN layers acts as input for the next CNN layer

The Second set of ICDSSPLD-CNN layers are described as below.

i. The first convolution layer (Params: 3x3x80 /Activation: ReLU) output is 128x128x80

ii. The second convolution layer (Params: 3x3x80 /Activation: ReLU) output is 128x128x80

iii. Max Pool layer (Params: 2x2, stride 2) output is 64x64x80

iv. The output of one layer acts as input for the next CNN layer

v. The output of the second set of ICDSSPLD-CNN layers acts as input for the next CNN layer

The Third set of ICDSSPLD-CNN layers are described as below.

i. The first convolution layer (Params: 3x3x160 /Activation: ReLU) output is 64x64x160

ii. The second convolution layer (Params: 3x3x80 /Activation: ReLU) output is 64x64x160

iii. Max Pool layer(Params: 2x2, stride 2) output is 256x256x1

IV. RESULTS

The ICDSSPLD-CNN is implanted using TesnorFlow. The preliminary research results are described in Fig.11 and Fig.12.

The pneumonia test results are as given below.

- Sample Data set: 624/624
- Processing time for each step is 487s 780ms
- Loss on test set: 0.9056972557536821
- Accuracy on test set: 0.8269230769230769
- Recall rate of the model is 0.98
- The precision of the model is 0.79

The Lung cancer test results are as given below.

- Sample Data set: 880/880
- Processing time for each step is 520s 820ms
- Loss on test set: 0.79956972557536821
- Accuracy on test set: 0.85230769230769
- Recall rate of the model is 0.97
- The precision of the model is 0.78

TesnsorfFlow Results- Pneumonia

624/624 [==============================] - 487s 780ms/step

Loss on test set: 0.9056972557536821

Accuracy on test set: 0.8269230769230769

recall of the model is 0.98

Precision of the model is 0.79

Fig.11 ICDSSPLD-Pneumonia preliminary results

TesnsorfFlow Results-Lung Cancer

880/880 [==============================] - 520s 820ms/step

Loss on test set: 0.79956972557536821

Accuracy on test set: 0.85230769230769

recall of the model is 0.97

Precision of the model is 0.78

Fig.12 ICDSSPLD-Lung cancer preliminary results

V. CONCLUSION AND FUTURE WORK

Pneumonia and lung cancer are critical lung organ diseases. Early detection of these terminal diseases will help in preventing deaths. We developed an integrated system (ICDSSPLD), which detects pneumonia and lung cancer in early stages. The first research objective is addressed in this paper. The proposed ICDSSPLD-CNN contains the three sets of CNN layers. Each set contains two convolution layers and one max pooling layer. ICDSSPLD-CNN architecture enhanced the detection accuracy.

The preliminary results of proposed system are satisfactory. We will extend our work to achieve the rest of the research objectives.

REFERENCES

1. Howlader N, Noone AM, Krapcho M, Miller D, Bresi A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2016, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2016, based on November 2018 SEER data submission, posted to the SEER website, April 2019.
2. https://www.thoracic.org/patients/patient-resources/resources/top-pneumonia-facts.pdf
3. https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI
4. https://data.mendeley.com/datasets
5. http://alexlenail.me/NN-SVG/LeNet.html
6. Hongkai Wang, Zongwei Zhou, Yingci Li, Zhonghua Chen, Peiou Lu, Wenzhi Wang, Wanyu Liu and Lijuan Yu3. Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18F-FDG PET/CT images.Wang et al. EJNMMI Research (2017) 7:11 DOI 10.1186/s13550-017-0260-9
7. Toloza, Eric M., Linda Harpole, and Douglas C. McCrory. "Noninvasive staging of non-small cell lung cancer: a review of the current evidence." *Chest* 123.1 (2003): 137S-146S.
8. Silvestri, G. A., Gonzalez, A. V., Jantz, M. A., Margolis, M. L., Gould, M. K., Tanoue, L. T., ... & Detterbeck, F. C. (2013). Methods for staging non-small cell lung cancer: diagnosis and management of lung cancer: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*, *143*(5), e211S-e250S.
9. Kumar, A., Dutta, R., Kannan, U., Kumar, R., Khilnani, G. C., & Gupta, S. D. (2011). Evaluation of mediastinal lymph nodes using 18F-FDG PET-CT scan and its histopathologic correlation. *Annals of thoracic medicine*, *6*(1), 11.
10. Kim SK, Allen-Auerbach M, Goldin J, Fueger BJ, Dahlbom M, Brown M, et al. Accuracy of PET/CT in characterization of solitary pulmonary lesions. J Nucl Med. 2007;48:214–20.
11. Li, X., Zhang, H., Xing, L., Zhang, L., Xu, W., Yang, G. R., & Yu, J. (2011). Mediastinal lymph nodes staging by 18F-FDG PET/CT for early-stage non-small cell lung cancer: A multicenter study. *Journal of Clinical Oncology*, *29*(15_suppl), e17523-e17523.
12. Schmidt-Hansen M, Baldwin DR, Hasler E, Zamora J, Abraira V, Roque IFM. PET-CT for assessing mediastinal lymph node involvement in patients with suspected resectable non-small cell lung cancer. Cochrane Database Syst Rev. 2014;11:CD009519.
13. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436–44.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

14. JUSTIN KER1, LIPO WANG, JAI RAO AND TCHOYOSON LIM. Deep Learning Applications in Medical Image Analysis,Digital Object Identifier 10.1109/ACCESS.2017.2788044
15. https://www.cdc.gov/nchs/fastats/pneumonia.htm
16. Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., & Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLoS medicine*, *15*(11), e1002683.
17. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
18. https://www.kaggle.com/nih-chest-xrays/datasets
19. Yan, C., Yao, J., Li, R., Xu, Z., & Huang, J. (2018, August). Weakly supervised deep learning for thoracic disease classification and localization on chest x-rays. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics* (pp. 103-110). ACM.
20. http://openaccess.thecvf.com/content cvpr 2017/papers/Wang ChestX-ray8 Hospital-Scale ChestCVPR 2017 paper.pdf
21. Diederik P. Kingma, Jimmy Lei Ba. Adam: A Method for Stochastic Optimization https://arxiv.org/pdf/1412.6980.pdf
22. Laurens van der Maaten, Geoffrey Hinton Visualizing Data using t-SNE 2008

AUTHORS PROFILE

Venkata Tulasiramu Ponnada, pursuing PhD in Computer science and engineering from Acharya Nagarjuna University, Nagarjuna Nagar, Guntur, Andhra Pradesh 522510, India. His area of interest includes machine learning, deep learning, artificial intelligence and safety critical systems, software engineering, software testing, software project management and System engineering.

Dr. S.V.Naga Srinivasu, Professor, Computer science and engineering, Narasaraopeta Engineering College, Narasaraopet, Andhra Pradesh 522601, India. His area of interest includes Software Engineering, Software Testing, Machine Learning, Deep learning, Mobile Networking, Operating systems, Data Mining, Image Processing and Black chain technology. He has published more than 58 articles in reputed journals which includes Scopus and other indexed bodies. He is a member of IEEE, MIET.