

Lung Nodule Detection from CT scans using Gaussian Mixture Convolutional AutoEncoder and Convolutional Neural Network

Nuthanakanti Bhaskar¹, Dr.Ganashree T S²

¹Research Scholar, CSE Dept., VTU-RRC, Visvesvaraya Technological University, Belagavi, India.

¹Assistant Professor, CMR Technical Campus, Hyderabad, Telangana.

²Associate Professor, Department of Telecommunication Engineering, Dayanandasagar College of Engineering, Bangalore, India.

¹bhaskar4n@gmail.com, ²ganashreeec@gmail.com

Abstract: A CAD (computer-aided diagnosis) framework based on a Deep Convolutional Neural Network was built in this paper. Initially, we applied Gaussian Mixture Convolutional Auto Encoder (GMCAE) on DSB2017 dataset (Kaggle Data Science Bowl), extracted 3D features from Lung CT images, and reconstructed the 3D array lung area. The candidates further classified in the 3D deep CNN to get the final result. The proposed method archives 74% of accuracy with a validation loss of 0.57.

Keywords: Lung Nodule Detection, Gaussian Mixture Convolutional Auto Encoder (GMCAE), computer-aided detection (CAD), convolutional neural network (CNN)

1. Introduction:

According to the National Cancer Registry Programme Reports for India 2020, cancer affects 6,79,421 males (94.1 per 1,000) and 7,12,758 females (103.6 persons per 1,00,000). Lung cancer affects one in every 68 men, breast cancer affects one in every 29 women, and one in every nine Indians develops cancer throughout their lifetime (age from 0 to 74 years). Non-Communicable Diseases (NCDs) account for 71 percent of all deaths worldwide, with cancer accounting for 9% of those [1].

The American Cancer Society estimated 16,65,540 cancer investigations in 2014, with 5,85,720 deaths expected in the United States. According to this report's review of recent 5 years data (from 2006 to 2010), the incidence of cancer decreased marginally in men (0.6 percent per year) and remained constant in women, while the rate of deaths decreased in both men and women (i.e., men 1.8% / year and women 1.4% / year). However, over a 20-year period (1991-2010), the overall cancer death rate (instances per 100,000 community) decreased from 215.1 to 171.8. During this time, cancer deaths dropped by 20%, or approximately 13,40,400 people, with 9,52,700 men and 3,87,700 women, based on age, race, and sex. In 80, little has changed [2].

Many systems developed and research is continuing on lung cancer detection. However, some systems are unsatisfactory by means of cancer detection accuracy, and some more systems improved to achieve the highest accuracy in the classification of nodules in CT images. Digital Image Processing and Deep Learning methods implemented Lung cancer nodule recognition and classification in CT images. We reviewed the latest systems for pulmonary lung nodule recognition and classification on CT scans to choose the best systems and analysis performed on them and the new system proposed.

2. Literature Review:

Bishop et al. [14] introduced a new Mixture Density Networks, which can model general conditional probability densities. By contrast, the conventional network approach, involving the reduction of a sum-of-squares error function, only permits the determination of the conditional average of the target data, together with a single global variance parameter. In this work, they illustrated a simple 1-input 1-output mapping network and robot inverse kinematics problem.

Riquelme et al. [4] Proposed popular techniques such as U-Net, Faster R-CNN, Mask R-CNN, YOLO, VGG, ResNet, and that can use for 2D or adapted to 3D data processing. Using multiple slices as input in the 2D deep networks can also be considered. 2D networks are more efficient in terms of processing time and memory requirement, which makes them an interesting solution for processing large medical DICOM data.

Ning J et al. [5] designed a 3-section U-NET and a CAD system false positive reduction algorithm formed on a 3D residual network. The outcomes of this work show 3D residual network had recommended performance and a better CPM score (0.780). Based on the computational outcomes, 3D convolution neural networks have a powerful spatial

feature extraction capability for 3D data with spatial information (e.g., CT scans) than 2D networks, and obtained better performance and CPM scores.

To assess true nodules, Chi J et al. [6] proposed a U-Net-like network that revamped Multi-Scale Pooling and Multi-Resolution Convolution Connections. Furthermore, the three sub-networks were combined and configured using a fused loss method that included MSE, perceptual, and dice losses. On two datasets, experimental results show that the proposed approach outperforms state-of-the-art methods in terms of pulmonary nodule detection (i.e., LUNA16 and TianChi competition dataset).

Xiao Z et al. [7] suggested a 3D-Res2UNet neural network and it enhances the training rate of the model while building the segmentation method finished. After completion of testing, it provided better outcomes than the other techniques by means of dice coefficient and recall rate.

Nasrullah N et al. [8] introduced a multi-strategy-based method for nodule detection and classification in the early stages of false-positive trimming. A 3D CT scan of the lungs is used to screen for the existence of malignant nodules. The CT lung scan was first subjected to a 3D Faster R-CNN with CMixNet and a U-Net-like encoder-decoder to detect nodules. To determine whether the nodules were regular or abnormal, 3D CMixNet and GBM were used. Finally, nodules were identified using deep learning techniques, which took into account many factors such as family background, age, smoking history, clinical biomarkers, height, and nodule location.

Alakwaa et al. [9] developed a CNN architecture (Convolutional Neural Network) to detect nodules in patients of lung cancer and detect the interest points using U-Net architecture. The deep 3D CNN models achieve performance AUC of 0.83, considering less labelled data than most state-of-the-art CAD systems.

Jalali Y et al. [10] implemented a segmentation framework utilizing a BCDU-Net with an encoder of pre-trained ResNet-34 and it named Res BCDU-Net. This architecture is a hybrid of the ResNet and BCDU-Net networks, with additional channels added. There were a few false positives, but the dice similarity scores were higher.

Xavier Rafael-Palou et al. [16] implemented computerized re-identification of pulmonary nodules and follow-up studies, using siamese neural networks (SNNs) to rank closeness between nodules without image registration. Non-identical designs of the conventional SNN inspected for transfer learning applications, loss functions, and a mixture of some attribute maps of distinct network levels. During off-line training of SSN achieved 92% accuracy with 89% cross-validation which is the same as state-of-the-art registration mechanisms. They implanted the SNN into a 2-stage nodule progress identification channel. Finally, the results found faster with an accuracy of 88% and 92% sensitivity, predicted and ground truth was not notable.

With the support of multi-level cross ResNet, Lyu J et al. [11] implemented a novel framework for deep learning computational. They looked at both ternary and binary classification, which involves categories like benign, unpredictable, and malignant. The results show that multi-level and cross residual structures help in the development of multi-scale attributes and their fusion to boost efficiency in binary and ternary classifications.

Perez G et al. [12] proposes an approach that is able to identify lung nodules and predict cancer effectively. They designed a candidate proposal method with almost perfect recall. They also trained a 3D convolutional neural network that successfully distinguished nodules from non-nodules and improved the exactness (compared to the HOG + SVM baseline), achieving human-like performance in this difficult task.

Xiao Z et al. [13] 3D CNN (Multi-Scale Heterogeneous) has been proposed for the reduction of false positives in pulmonary nodules. In this network, three methods used: (1) 3D multi-scale gradual integration (2) heterogeneous feature extraction (3) intellectual weight fusion. They tested the proposed algorithm by applying the method to the LUNA16 dataset and estimated the effects of multi-scale heterogeneous 3D convolutional neural networks with different structures.

On publicly accessible datasets, CNN obtained stronger results in recent research. In recent years, CNNs have shown a good capacity to learn valuable feature representation via representation. However, in this paper, we used the Gaussian

Mixture Convolutional Auto Encoder (GMCAE) and CNN to build a CAD scheme.



Fig.1 Proposed Method

3. Methodology:

3.1. Data:

We used primarily DSB-2017dataset(Kaggles Data Science Bowl) [3]. It contains 1397 patients (1 sample = 1 patient) labelled data, which we divided 1117 samples (291 cancerous) (approximately ~80%) for training, 280 samples (71 cancerous) (approximately ~20%) for validation, and 198 samples for test. The dataset contains CT images for each patient and labelled 0 for Non-Cancer and 1 for Cancer. In this dataset nodules not labelled and slices supplied in DICOM format. In this, a different number of images (512 * 512 pixels) are available in each patient CT scan data (approx. 100-400, everyscan is an axial slice). By merging the slices, we constructed a 3D array for each patient which constitutes a sample and associated a binary label (0 for Non-Cancer and 1 for Cancer).

In the histograms below, Fig.2 indicates the distribution of the number of fragments and spatial resolution.

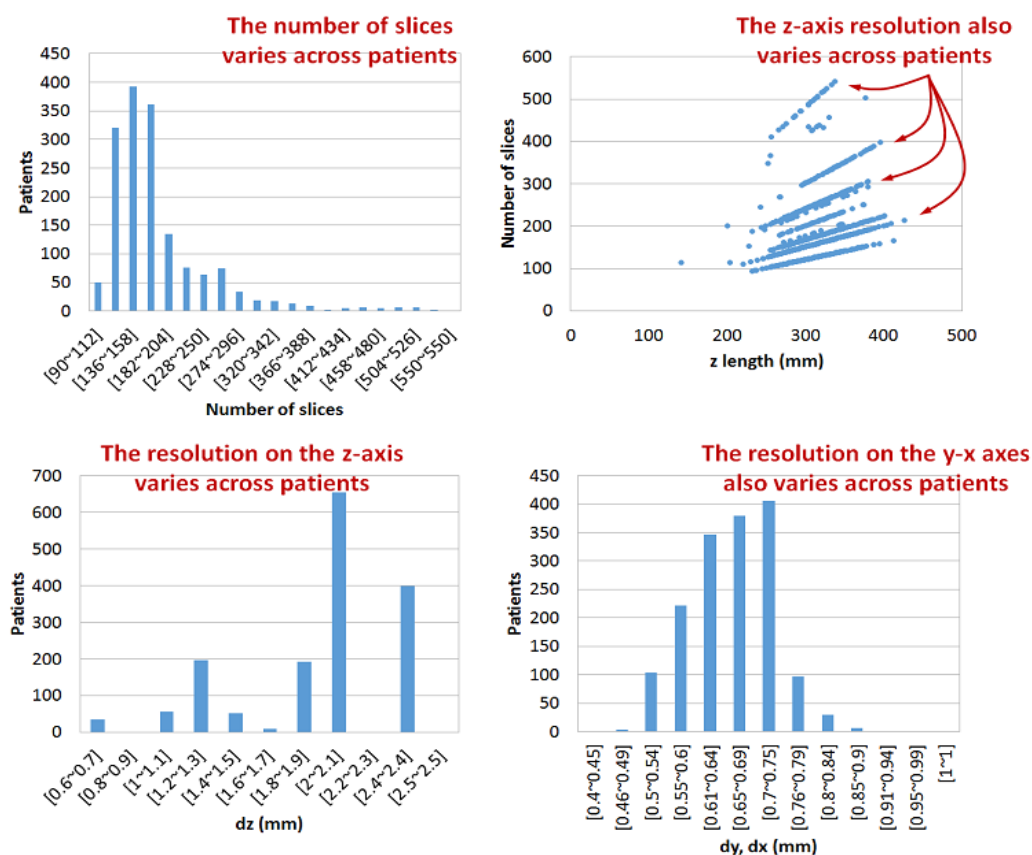


Fig.2 Slice and Resolution Distribution

3.2. Preprocessing:

The CT image slices are preprocessed and consolidated into a single 3D array for each case using the methods shown in Fig. 3.

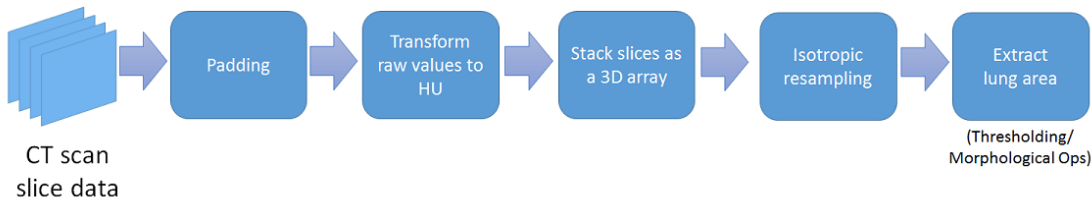


Fig.3 Preprocessing

In this process we used 6 phases: 1. Read sequence of DICOM files of a patient. 2. Applied Padding on images. 3. Converting DICOM raw pixel array into HU units image array. 4. Stack DICOM slice sequence into 3D array. 5. Resampling (Resamples an array to specified pixel spacing (resolution)). 6. Extract lung array.

Below Fig.4 is shown an array obtained after preprocessing the CT scan slices:

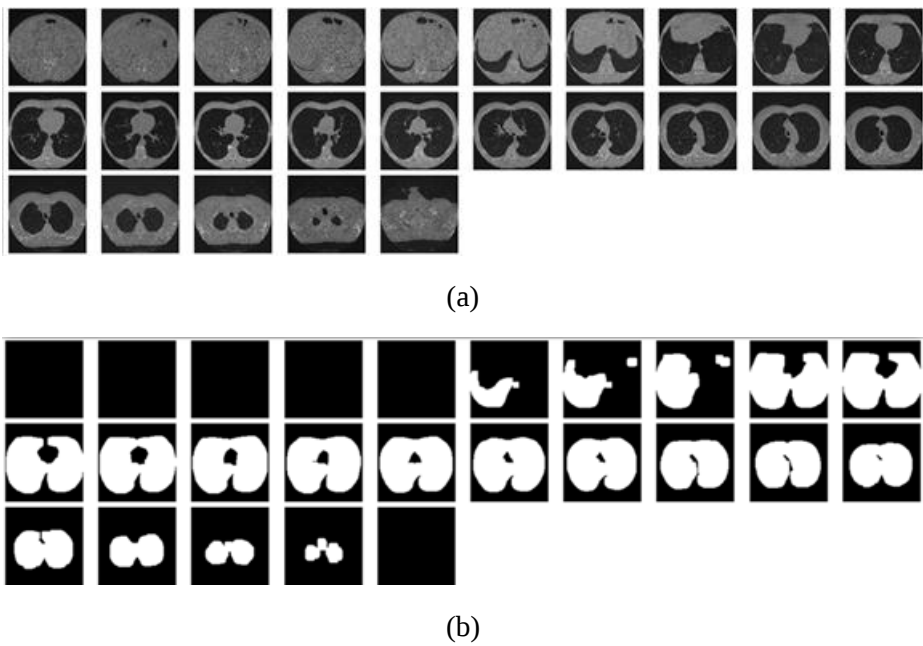


Fig.4 (a) Lung Slices (b) Masks

The size of the lung array along the z-y-x axes vary across patients and thus the model should be able to handle arrays of varying size.

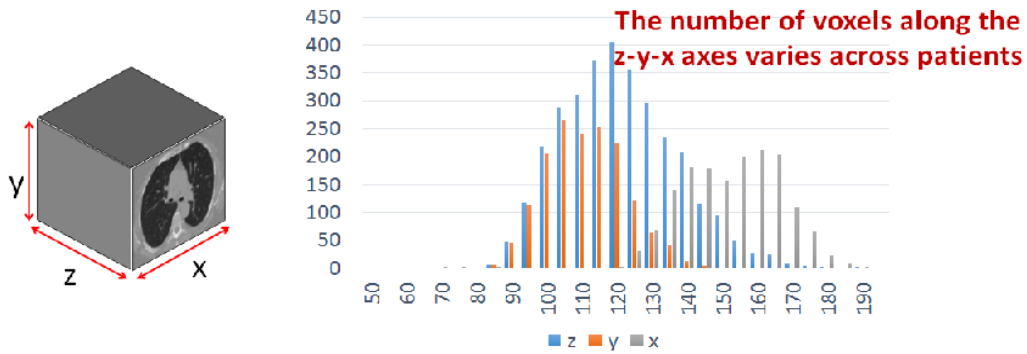


Fig.5 Distribution of Lung Array Size

4. Methods:

The whole lung nodule identification system mainly divided into 2 stages. In stage-1, Gaussian Mixture Convolutional AutoEncoder (GMCAE) used to extract required features from a 3D CT lung array. The candidates further classified using CNN classifier in stage-2 to get the final result.

4.1. Gaussian Mixture Convolutional AutoEncoder (GMCAE):

The purpose of this network is to learn features from the 3D CT lung arrays that could be transferred to the second network for classification. This is done through unsupervised learning to use an autoencoder with a reconstruction task.

As a reconstruction objective for the autoencoder, one could attempt to minimize an MSE objective, but this would fail because the CT scan voxels have a multimodal distribution and an MSE objective would tend to predict the average of the distribution and thus likely yield meaningless predictions.

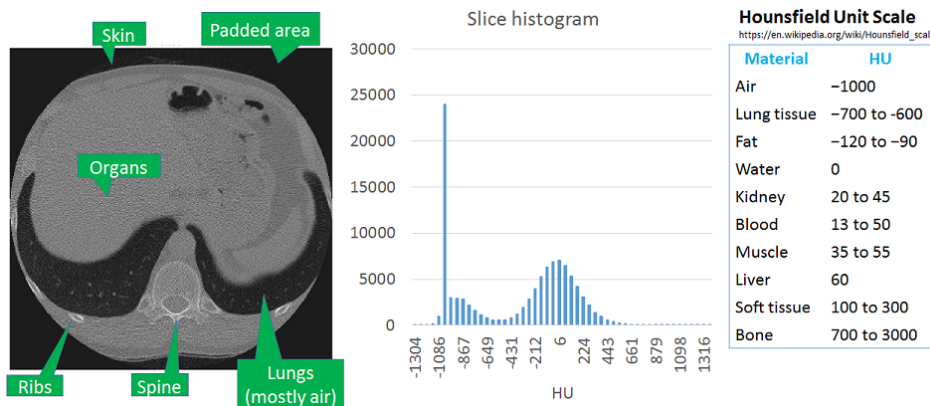


Fig.6 CT Scan Voxels Multimodal Distribution

This is because enhancing the log-likelihood by assuming a (uni-modal) Gaussian distribution for bound up with the probability of the output provided the data is equivalent to an MSE objective.

Thus, the conditional probability is instead formulated as a mixture of Gaussians as:

$$p(t|X) = \sum_{k=1}^m \alpha_i(X) \phi_k(t|X) \quad (1)$$

$$\phi_k(t|X) = \frac{1}{(2\pi\sigma_k(X)^2)^{c/2}} \exp\left(-\frac{\|t - \mu_k(X)\|^2}{2\sigma_k(X)^2}\right) \quad (2)$$

Where m is the number of gaussians in the mixture, and c is the number of output dimensions (number of voxels in the reconstruction).

The GMCAE is trained to produce outputs that determine the parameters α (priors), σ^2 (variances) and μ (means) of the mixture of Gaussians. α , σ and μ are functions of X and the network parameters θ .

Since we are doing reconstruction, $t=x$ in this case. Specifically, the network is trained to diminish by reason of loss function:

$$J(\theta; X, t) = - \sum_{n=1}^N \log \sum_{k=1}^M \exp \left\{ \log(\alpha_k(X_n)) - \frac{c}{2} \log(2\pi) - \frac{c}{2} \log(\sigma_k(X_n)^2) - \frac{\|t_n - \mu(X_n)\|^2}{2\sigma_k(X_n)^2} \right\} \quad (3)$$

In this, the priors and normalizing constants of the Gaussians are moved inside the exponential function, allowing to exhibit the loss as a logsumexp and improves the numerical stability.

4.2. CNN Classifier:

The aim of this network is to identify patients based on GMCAE features. The network is trained to decrease the Loss Function, as well as the output is a single sigmoid unit.

A Spatial Pyramid Pooling layer is used to communicate between the convolutional and fully-connected layers since the model should be able to accommodate arrays of varying sizes[15].

5. Implementation:

Gaussian Mixture Convolutional AutoEncoder (GMCAE): A convolutional autoencoder cast as Mixture Density Network [14]. This network is used to learn high-level features of patches of lung scans (3D arrays of CT scans in Hounsfield Units), using unsupervised learning and maximum likelihood on a mixture of Gaussians.

CNN classifier performs binary classification upon the characteristics extracted by the encoding layers of the GMCAE.

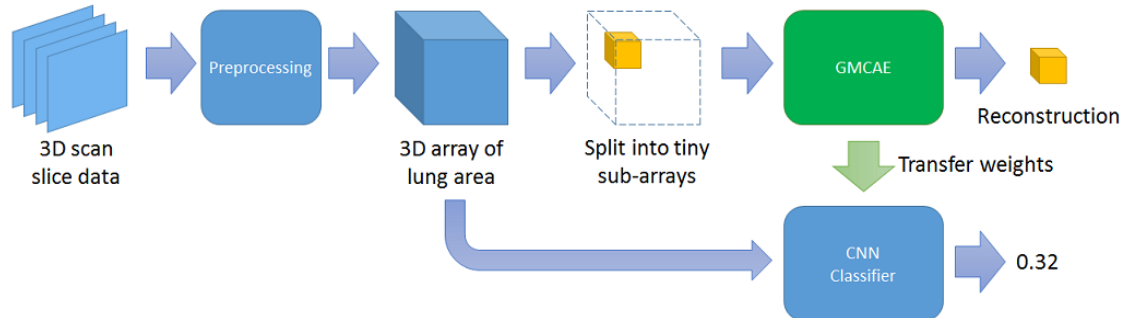


Fig.7 Model Overview

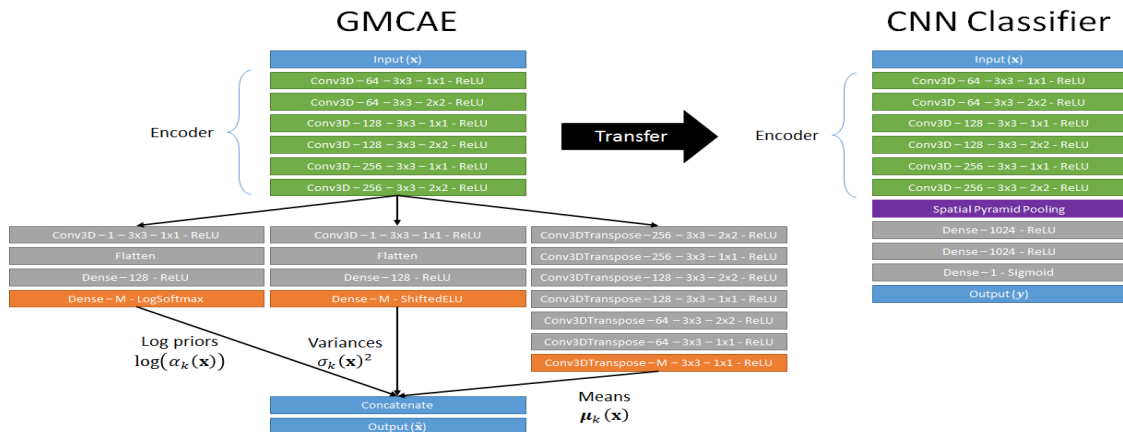


Fig.8 Network

5.1. GMCAE Setup:

Inputting a 3D sub-array to GMCAE which corresponding to a cube patch of fixed size, which is big enough to contain a lung nodule. The sub-arrays set to 32x32x32 corresponding to a cube of 3.2 mm. And data augmentation performed by random rotations or mirroring of the sub-arrays (A cube has 48 symmetries this allows a 48-fold augmentation).

5.2. CNN Classifier Setup:

Inputting a Full 3D array of lung area to CNN Classifier and Data augmentation is performed by random rotations or mirroring of the sub-arrays.

6. Results and Discussion:

The reconstruction computed with the GMCAE using the combination of 4 Gaussians shown below Fig.9. The 24 slice patches on the left are the original, and those on the right are the reconstructions produced by the model.

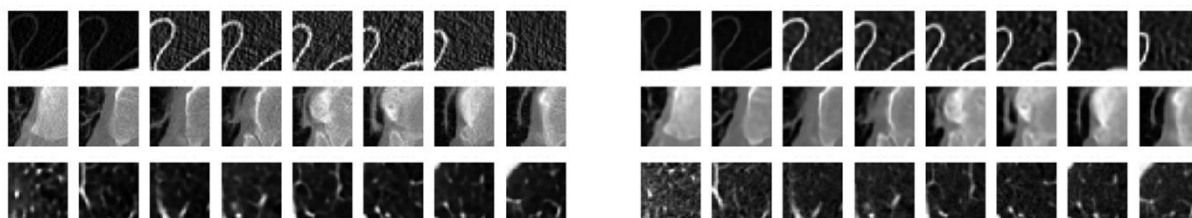


Fig.9 Reconstruction computed with the GMCAE using the combination of 4 Gaussians

The train and validation loss for $m = 2, 4$ is shown below Fig. 10.

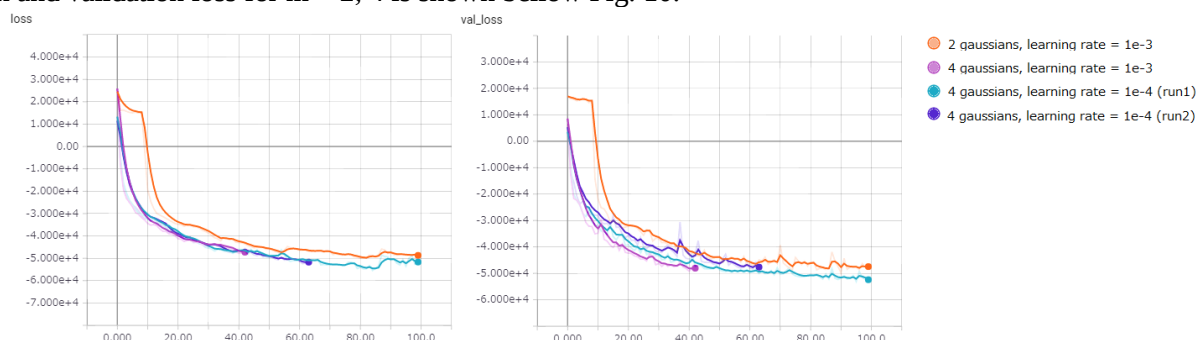


Fig.10 Train and Validation Plots

The log likelihood can take negative values because point estimates of the density can take values greater than 1 if the variances are made small enough.

7. Conclusion:

A 3D deep CNN-based CAD device for nodule detection in lung CT images has been proposed. High-level features were extracted from the DSB-2017 dataset (Kaggle's Data Science Bowl) and reconstructed as a 3D Array using a GMCAE. Finally, CNN investigation findings confirmed the effectiveness of the suggested mechanism, classifying nodules as normal or abnormal with a validity loss of 0.57 and a 74 percent accuracy.

For potential projects, to begin, control gradient explosion: gradients with low learning rates and/or gradient standard

clipping. And then directly parametrize to the opposite variance. Often, the variances are set to a constant value (determined empirically). Second, the loss function's lower bound: Since the Gaussians in the mixture are densities, point estimates of the probability can be negative if the variances are small enough. Because of the variable variances and priors, determining the lower bound of the loss function is difficult, as is determining how much the model underfits the data.

References:

- [1] Mathur P, Sathishkumar K, Chaturvedi M, Das P, Sudarshan KL, Santhappan S, Nallasamy V, John A, Narasimhan S, Roselind FS, ICMR-NCDIR-NCRP Investigator Group, “Cancer Statistics, 2020: Report From National Cancer Registry Programme, India”, *JCO Glob Oncol*, 2020. [\[CrossRef\]](#)[\[PubMed\]](#)
- [2] Siegel R, Ma J, Zou Z, Jemal A, “Cancer Statistics, 2014”, *CA Cancer J Clin*, Epub 2014. [\[CrossRef\]](#)
- [3] Kaggle, “Data science bowl 2017”, <https://www.kaggle.com/c/datascience-bowl-2017/data>, 2017. [\[CrossRef\]](#)[\[PubMed\]](#)
- [4] Riquelme, Diego & Akhloufi, Moulay, “Deep Learning for Lung Cancer Nodules Detection and Classification in CT Scans”, *AI* 2020. [\[CrossRef\]](#)
- [5] Ning J, Zhao H, Lan L, Sun P, Feng Y, “A Computer-Aided Detection System for the Detection of Lung Nodules Based on 3D-ResNet”, *Appl. Sci.* 2019. [\[CrossRef\]](#)
- [6] Chi J, Zhang S, Yu X, Wu C, Jiang Y, “A Novel Pulmonary Nodule Detection Model Based on Multi-Step Cascaded Networks”, *Sensors* 2020. [\[CrossRef\]](#)[\[PubMed\]](#)
- [7] Xiao Z, Liu B, Geng L, Zhang F, Liu Y, “Segmentation of Lung Nodules Using Improved 3D-UNet Neural Network”, *Symmetry* 2020. [\[CrossRef\]](#)
- [8] Nasrullah N, Sang J, Alam M.S, Mateen M, Cai B, Hu H, “Automated Lung Nodule Detection and Classification Using Deep Learning Combined with Multiple Strategies”, *Sensors* 2019. [\[CrossRef\]](#)[\[PubMed\]](#)
- [9] Alakwaa, Wafaa & Nassef, Mohammad & Badr, Amr, “Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN)”, *International Journal of Advanced Computer Science and Applications*, 2017. [\[CrossRef\]](#)
- [10] Jalali Y, Fateh M, Rezvani M, Abolghasemi V, Anisi M.H, “ResBCDU-Net: A Deep Learning Framework for Lung CT Image Segmentation”, *Sensors* 2021. [\[CrossRef\]](#)[\[PubMed\]](#)
- [11] Lyu J, Bi X, Ling S.H, “Multi-Level Cross Residual Network for Lung Nodule Classification”, *Sensors* 2020. [\[CrossRef\]](#)[\[PubMed\]](#)
- [12] Perez G, Arbelaez P, “Automated lung cancer diagnosis using three-dimensional convolutional neural networks”, *Med Biol Eng Comput* 2020. [\[CrossRef\]](#)[\[PubMed\]](#)
- [13] Xiao Z, Du N, Geng L, Zhang F, Wu J, Liu Y, “Multi-Scale Heterogeneous 3D CNN for False-Positive Reduction in Pulmonary Nodule Detection, Based on Chest CT Images”, *Appl. Sci.* 2019. [\[CrossRef\]](#)
- [14] Bishop, Christopher M, “Mixture density networks”, *Technical Report, Aston University, Birmingham* 1994. [\[CrossRef\]](#)
- [15] He Kaiming & Zhang Xiangyu & Ren Shaoqing & Sun Jian, “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2014. [\[CrossRef\]](#)[\[PubMed\]](#)
- [16] Xavier Rafael-Palou, Anton Aubanell, Ilaria Bonavita, Mario Ceresa, Gemma Piella, Vicent Ribas, Miguel A. González Ballester, “Re-Identification and growth detection of pulmonary nodules without image registration using 3D siamese neural networks”, *Medical Image Analysis* 2021. [\[CrossRef\]](#)[\[PubMed\]](#)