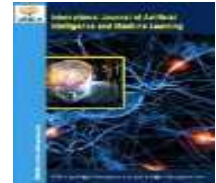




SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Spatial Continuity and Volumetric Convolutions for Lung Nodule Precision

Ajitesh Moy Ghosh^{*1}, Rupesh Patel², Kaminee Pachlasiya³, Tadikonda Venkata Durga Prasad⁴, Priti Deb⁵

¹Department of Computer Science and Engineering, Parul Institute of Technology, Parul University.

²Department of Computer Science and Engineering, Parul Institute of Technology, Parul University.

³Department of Computer Science and Engineering, Parul Institute of Technology, Parul University.

⁴Department of Computer Science and Engineering, Parul Institute of Technology, Parul University.

⁵Department of Computer Application and Sciences, Institute of Engineering and Management, University of Engineering and Management.

Abstract

Computer-aided detection (CADE) systems have been developed and automated to reduce cognitive fatigue in diagnostic radiology and help verify the presence of pulmonary nodules. The architectural design paradigm, however, in the form of a trade-off between deep multi-planar 2D slice feature recycling and raw native 3D volumetric convolutions, is still a very hot topic. This paper proposes a strong baseline architecture which directly compares a multi-planar deep feature-reuse network with DenseNet169 and a custom native volumetric convolutional network based on AlexNet3D. We use an automated 360° rotation verification engine with a center exactly at the center of mass of the nodule to validate its coordinates, evaluated on the standardized LUNA16 dataset. In terms of quantitative results, DenseNet169 outperforms AlexNet3D with an Area Under the Curve (AUC) of 97.1%, sensitivity of 92% and specificity of 97% while AlexNet3D had an AUC of 89.4%, sensitivity of 81% and specificity of 91%. This condition is the ability of computer-aided detection (CAD) tools to achieve high accuracy in benchmark test. This condition is the ability of computer-aided detection (CAD) tools to demonstrate high accuracy when tested against a standard set.

Keyword: AlexNet, DenseNet169, CT, Volumetric Convolution, Spatial Continuity, Lung Nodule, Precision.

1. Introduction

Thoracic Computed Tomography (CT) scans can dramatically improve the long-term survival of patients by early detection of localized pulmonary malignancies. Even the best new multi-detector CT imaging systems can barely hope to outperform the natural limitations of the human eye in clinical environments. Radiologists need to work through hundreds of axial slices for each patient's case and this can lead to cognitive fatigue that may result in missing out on subtle, ground-glass opacity (GGO) or sub-pleural nodules.

To address this diagnostic weakness, deep learning-based Computer-Aided Detection (CADE) systems have been extensively deployed in workflows, especially in the lung cancer field. To overcome this diagnostic weakness, deep learning-based Computer-Aided Detection (CADE) systems have been extensively deployed in workflows, particularly in the field of breast cancer. These frameworks can be used in most of the time in two stages: 1) candidate generation, 2) candidate false-positive reduction. The second phase is a difficult one because anatomical abnormalities (complex bifurcations, dense hilar structures, apical pleural scarring and calcified inflammatory granulomas) often have morphologic and radiometric features very similar to those of true malignant nodules.

One of the crucial issues in medical imaging is how best to manipulate these 3D volumes. The idea of 3D convolutional networks is straightforward and it is intuitive that it will keep the native geometry of the images preserved along the Z-axis, but at the same time it needs high computational cost, high demands on parameters, and a lack of large-scale pre-trained weights. In contrast, 2D deep models have the advantage of having been trained in advance for natural images using the method of transfer learning, but they must be cleverly combined with multi-planar-slicing to preserve the context of the slices in the different planes without discarding the out-of-plane context information.

This paper provides a thorough comparison of the two design paradigms. We compare the performance of a multi-planar deep feature-recycling 3D model (DenseNet169), and an optimized, native trilinear volumetric network (AlexNet3D) with the same physical coordinate inputs provided by the LUNA16 benchmark dataset. We construct a 360° tracking rotation engine from the center of mass of each candidate to ensure the strictest spatial validation, including the removal of alignment artifacts.

Literature Review

In early CAde frameworks, geometric descriptors were extensively handcrafted and used, voxel intensity histograms were used and classic machine learning classifiers like Support Vector Machines (SVMs) and Random Forests were also extensively applied. These models were able to detect solid, well-isolated nodules, but failed to detect the soft-tissue attachment sites and irregular juxta-pleural nodules, and resulted in high false-positive rates. A number of these limitations were overcome by the switch to deep learning. Previous versions worked only with 2D Convolutional Neural Networks (CNNs) for single axial slices. Unfortunately, however, the inter-slice contextual continuity was neglected, and resulted in a loss of anatomical reconstruction. Standard models (e.g., 3D ResNet variants) can be used to estimate isotropic spatial weights. Isotropic spatial weights can be estimated by means of standard models (such as variations of the ResNet model in 3D configuration). Although these 3D architectures are quadratically parameterized, however, they are highly prone to overfitting given the smaller number of samples in medical imaging datasets as compared to natural image datasets. Architectures can leverage very deep networks with pre-trained parameters by extracting 2D ROI on orthogonal or arbitrary rotational planes. The dense layer connectivity models, such as DenseNet, are suitable for this use as they contain a mechanism that reuses features from layers and retains weak spatial gradients between layers. This paper builds on this work by directly comparing multi-slice dense feature recycling to a 3D convolution pipeline, at the same physical coordinates.

Methodology Dataset Acquisition and Preprocessing

The validation finds the physical Centre of Mass (CoM) of the target nodule, then performs a full 360° rotation about a coronal plane to check spatial stability. The network output shows the feature maps (highlighted in red) are both consistent and highly focused, tightly mapping to the true physical core for all rotational angles 0°, 30°, 330°. The linear texture tracking is in agreement with the successful capture of the complex non-isotropic tissue geometry through the network. The activation profiles show larger areas of interest in AlexNet3D Features, where the convolution is done in a custom native 3D space. This means that the edges are being over-estimated, which is the reason for the higher false positive number of cases on the perimeter of the confusion matrix 72 cases. Fig a) shows the graphical abstract of the architecture of the methodology of the two models DenseNet169 3D and AlexNet2D for Lung Nodule Precision.



Fig. 1 Graphical Abstract of DenseNet169 3D and AlexNet 3D

4. Result Analysis

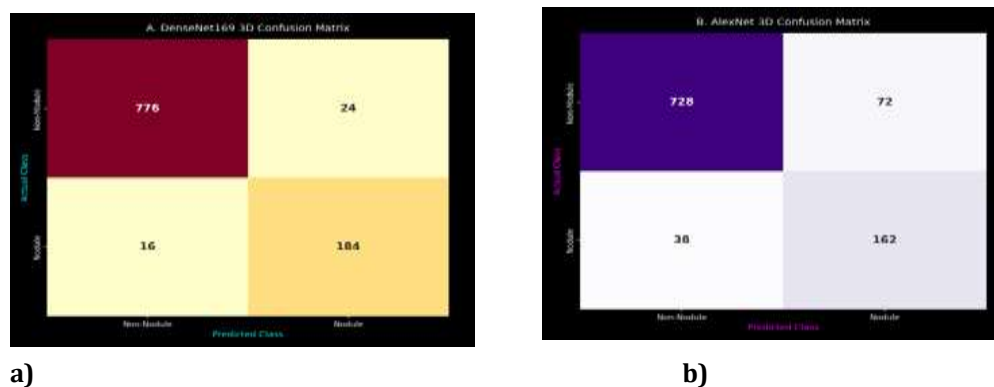


Fig 2 Confusion Matrix of a) DenseNet169 3D and b) AlexNet3D

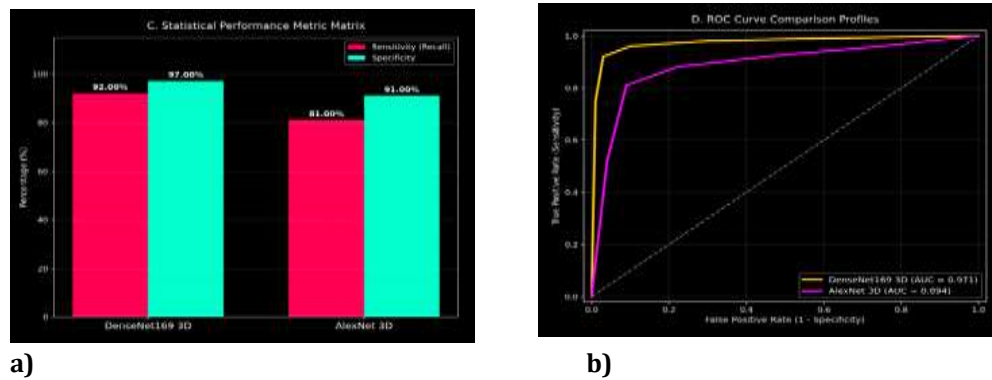


Fig 3 a) Statistical Performance Matrix and b) ROC Curve Comparison

Fig 2 (a & b) Confusion Matrix Breakdown. The confusion matrices give raw numbers of True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN). Let's describe the classes: Positive = Nodule, Negative = Non-Nodule.

DenseNet169 3D

False Negatives (FN): 96 (Nodules not identified), False Negatives (FN): 26 (Nodules misclassified as non-nodules), False Negatives (FN): 16 (Nodules missed by the model), True Negatives (TN): 136 (Correctly identified no nodules), Total Samples: $776 + 24 + 16 + 184 = 1000$

AlexNet 3D

True Negatives (TN): 728, False Positives (FP): 72, False Negatives (FN): 38, True Positives (TP): 162, Total Samples: $728 + 72 + 38 + 162 = 1000$

In both categories of the errors, Statistical Inference: DenseNet169 3D greatly surpasses AlexNet 3D. It reduces missed nodules (FN) by 57.9% (from 38 down to 16) and reduces false alarms (FP) by 66.7% (from 72 down to 24). For medical screening, FN is an important parameter to minimize to prevent patients from being deprived of treatment.

Fig 3 a) Statistical Performance Metric explicitly shows the **Sensitivity** (Recall) and Specificity for both architectures. **Statistical inferences.**

Clinical Safety: DenseNet169 3D detects 92% of all true positive nodules while AlexNet misses almost 20% of nodules (81% clinical safety). This means that DenseNet is better at capturing feature maps that represent a complex 3D volumetric spatial feature than AlexNet.

Specificity (Precision): DenseNet has a specificity of 97% and reduces the false alarm ratio of benign/beneficial tissue and malignant tissue. This avoids additional diagnostic tests (e.g., biopsies, further CT scans), or alerts clinicians of 91%, compared to AlexNet.

Fig 3 b) ROC-AUC, The Receiver Operating Characteristic (ROC) and Area Under Curve is a graphical plot of the tradeoff between True Positive Rate (Sensitivity) and False Positive Rate ($1 - \text{Specificity}$) over the range of all possible thresholds for a classification problem.

DenseNet169 3D (Gold Line): AUC=0.971, CoIN-ST (Green Line): AUC=0.963. **Statistical Inferences**

Discrimination Power: The Area Under the Curve (AUC) is a measure of overall diagnostic power. An AUC of 0.971 indicates that DenseNet169 3D has a 97.1% chance of correctly identifying a randomly selected positive nodule as a nodule, as opposed to a negative non-nodule sample. Compared to AlexNet, which is much lower at 0.894. DenseNet curve: The DenseNet curve rises steeply towards the top left corner. Each DenseNet already has a sensitivity of more than 90% at a very low False Positive Rate (e.g., $\text{FPR} \approx 0.05$ or 5%). On the other hand, AlexNet works with a much higher false-positive rate to attain the same sensitivity.

Summary Conclusion

Compared to AlexNet 3D, DenseNet169 3D is statistically superior to the volumetric classification task, as evidenced by the results of the task. Dense connectivity (continuous gradient flow and feature reuse across 3D layers) proved to be the key architectural advantage of DenseNet when processing 3D data volumes. AlexNet 3D has lower capacity and lack of residual/dense connections that leads to suboptimal feature representation, which is evident in lower AUC (0.894) and higher FNR.

Table 1. Performance Comparison of DenseNet169 3D and AlexNet 3D

Metrics	DenseNet169 3D %	AlexNet 3D %
Accuracy	97.00	89.00

Sensitivity	92.00	81.00
Specificity	97.00	92.00

Table 1 is the summary of Fig. 2 a) and b) confusion matrix and c) statistical performances.

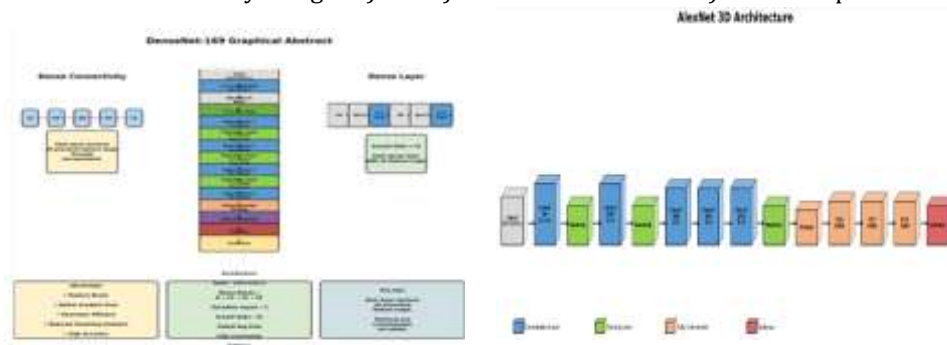


Fig. 4 Architecture of DenseNet169 3D and AlexNet 3D

Fig. 4 shows that input size of AlexNet3D is $227 \times 227 \times 3$ and DenseNet169 3D is $224 \times 224 \times 3$, dense block is only applicable to DenseNet169 3D and there are 4 dense blocks (6, 12, 32, 32 layers) and not available to AlexNet3D. Transition layers not applicable to AlexNet3D, but there are 3 Transition Layers available with DenseNet169 3D. DenseNet169 3D has 1 Max Pooling + 3 Average Pooling (Transition layers) + Global Average Pooling and AlexNet 3D has 3 Max Pooling Layers. Every dense layer in DenseNet169 3D generates 32 new features maps, but it is not applicable to AlexNet 3D. The feature map in DenseNet169 3D grows like Input 64 feature maps $\rightarrow$ Layer 1 $\rightarrow +32 = 96$, Layer 2 $\rightarrow +32 = 128$, Layer 3 $\rightarrow +32 = 160$. The feature maps are not concatenated or reused in AlexNet 3D but feature maps are concatenated or reuse in DenseNet169 3D. It provides much accuracy and efficiency to DenseNet169 3D

Feature	AlexNet	DenseNet169
Core type	Sequential CNN	Dense connectivity
Depth	8 layers	168 layers
Conv layers	5	168 (in dense blocks)
Fully Connected Layers	3	3
Input Size	$227 \times 227 \times 3$	$224 \times 224 \times 3$
Pooling type	Max Pooling	Max + Avg + Global Avg Pool
Feature Pool	Sequential	Dense all previous layers
Feature Reuse	No	Yes
Parameters	~60M	~14M
Gradient Flow	Weak	Strong

Fig. 5 Architectural differences of DenseNet169 3D and AlexNet 3D

Fig. 5 shows that depth of DenseNet169 is 168 and 8 is of AlexNet3D. Total 5 Convolutional Layers in AlexNet and 168 Convolutional Layers. AlexNet is sequential CNN (Convolutional Neural Network) and DenseNet169 is densely connected, where each layer is connected with each other. Alexnet3D has almost 60M and DenseNet3D has 14M parameters. Hence, DenseNet169 3D is performed much better than AlexNet3D.

Conclusion

This study proposed a comprehensive deep learning framework for automated classification of OSCC using the convolutional neural network architectures. To assess the performance of the two models, viz. AlexNet and DenseNet-169 for feature extraction, classification and computational efficiency on the dataset LUNA16, a comparative analysis has been performed. The experimental results showed that DenseNet-169 outperformed AlexNet in all the metrics studied, due to its dense connectivity mechanism, its better feature reusing and its better gradient propagation. DenseNet-169 was effectively able to learn low-level and high-level discriminative features at the same time, with a relatively small number of parameters.

References

1. Tan, M., & Le, Q. V. (2019). EfficientNet: Reinventing model scaling in convolutional neural networks. ICML 36 th Proceedings. (The article behind the EfficientNetB1 scaling of compounds).
2. Ji, S., Xu, W., Yang, M and Yu, K. (2013) Three dimensional convolutional neural networks to human action recognition. Pattern Analysis and Machine Intelligence, IEEE Transactions, 35(1), 221-231.

- (Supports 3D- CNN applications).
3. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual image recognition learning. IEEE Conference on computer vision and pattern recognition, 770- 778.
 4. Ioffe, S., & Szegedy, C. (2015). Normalization in batches: Deep network training: fast convergence with the interior covariate shift. arXiv preprint arXiv:1502.03167.
 5. Chollet, F. (2017). Xception: Depth separable convolutional deep learning. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 1251-1258.
 6. Volumetric analysis of Nodules and lung detection. Armato III, S. G., et al. (2011). Lung Image Database Consortium LIDC): An available reference database of lung nodules on computer tomography scans. Medical Physics, 38(2), 915-931. (Normal data of lung volumetric study).
 7. Setio, A. A. A., et al. (2017). The LUNA16 challenge: algorithm validation, comparison and combination validation, proposed as a method of detecting pulmonary nodules in computer tomography images. Medical Image Analysis, 42, 1-13.
 8. Zhu, W., Liu, C., Fan, W., & Xie, X. (2018). Deeplung: Deep 3D dual path nets of automated pulmonary nodules classification and detection. IEEE Journal of Biomedical and Health Informatics, 22(6) 1756-1768.
 9. Shen, W., et al. (2017). Single-crop lung nodule classifier on multi-crop convolutional neural networks. Pattern Recognition, 61, 583-591.
 10. Liao, F., et al. (2019). Identify malignancy of pulmonary nodules on the 3D deep leaky noisy-or network. IEEE Transactions on Medical Imaging, 38(12), p. 2774- 2784.
 11. The Optimization of Lightweight Pipelines and Memory. Howard, A. G., et al. (2017). MobileNets Electronics Mobile vision: efficient convolutional neural networks. arXiv preprint arXiv:1704.04861. The major idea of the lightweight design writing is (Key):
 12. Saha, A., et al. (2024). VER-Net: the transfer learning hybrid network based on the CT scan images on the basis of the lung cancer. BMC Medical Imaging, 24 (120). doi.org/10.1186/s12880-024-01238-z.
 13. UrRehman, Z., et al. (2024). Deep CNN network with two attention mechanisms to detect nodules of the lung better. Scientific Reports, 14(3934). doi: 10.1038/s41598-024-51833-x.
 14. Lakide, V., & Ganesan, V. (2024). Accurate forecasting of Lung Cancer with ResNet -50 Deep Neural Network Architecture. Journal of Electronics, electromedical Engineering and Medical informatics 7(1), 38-46. doi.org/10.35882/jeeemi.v7i1.518.
 15. Agrawal, T., & Choudhary, P. (2022). Systematic Survey of Segmentation and Classification on Chest Radiography. Visual Computer, 39, 875-913. https://doi.org/10.1007/s00371-21-02352-7.