2nd PANHELLENIC CONGRESS OF MEDICAL PHYSICS
4-6 OCTOBER 2024| EUGENIDES FOUNDATION

Differences between 2D CNNs and 3D CNNs in brain lesion image segmentation

Konstantia Moirogiorgou[1], Ellas Spyratou[2], Efstathios P Efstathopoulos[2], Antonis Sideris [1]

[1]R&D dept., Seven Red Lines company, Vrilissia, Greece

[2]2nd Department of Radiology, Medical School, Attikon University Hospital, National and Kapodistrian University of Athens, 11527 Athens

1. Background-Aim

- Accurate diagnosis and treatment of brain stroke affects directly the recovery process and the patient's wellness after the incidence.

- Accurate lesion segmentation using stroke images using magnetic resonance imaging (MRI) scans enables the volumetric quantification of lesion load and, so helps on monitoring the progression over time along with response to treatments.

- Moreover, in case of acute and early subacute ischemic strokes, the detailed representation of stroke progression and the advanced features of the lesion can be crucial for patient's health.

- 2D CNN architectures like ResNet50, MobileNet, Efficient Net, etc. are able to classify brain image scans, like computed tomography (CT) or MRI scan images, into hemorrhagic stroke, ischemic stroke and normal with high precision[1,2].

- Recent studies present optimized results for acute stroke lesion detection using 3D CNNs utilizing CT or MRI images, able to provide solutions that reduce false positive findings[3,4].

- Following the above context, present study explores the differences between 2D and 3D CNN architectures for early detection and classification of brain tumors.

2. Materials & Methods

- In general, 2D CNNs are able to process RGB images, i.e. 3 channels of image data (2D frames) and many pre-trained models like U-Net and ResNet are usually used for MRI segmentation.
- 3D CNNs take as input a 3D volume or a temporal sequence of 2D frames, where MRI scans are considered as an input class of this case.

- In the present study, a pre-trained (weights from the ImageNet dataset) ResNet50 model was used to perform image classification for brain tumor MRI images, a common practice due to the computational cost of training such models[5].
- Additionally, a 3D CNN was implemented and performed over the same input dataset in order to compare their performance. In this case, a reshape layer has been added (store data in rank-3 tensors) in order to be able to perform 3D convolutions on the data. The additional dimension is needed to take into account the number of image channel which in this case is just 1[6].

- This brain tumor dataset[7] that was used contains three kinds: meningioma (705 dicom images), glioma (54) and pituitary tumor (7), so the dataset is quite unbalanced. For the split of the dataset samples, for both training and validation, we used 60%-40% ratios.
- Training samples - Total 55 patients:
  - Case_1_Meningioma: 48 patients, Case_2_Glioma: 5 patients, Case_3_Pituitary_Tumor: 2 patients
- Validation samples - Total 34 patients:
  - Case_1_Meningioma: 31 patients, Case_2_Glioma: 2 patients, Case_3_Pituitary_Tumor: 1 patients
- As an output, the ROC AUC is used to provide a representation of the model's performance, since the validation set is unbalanced.

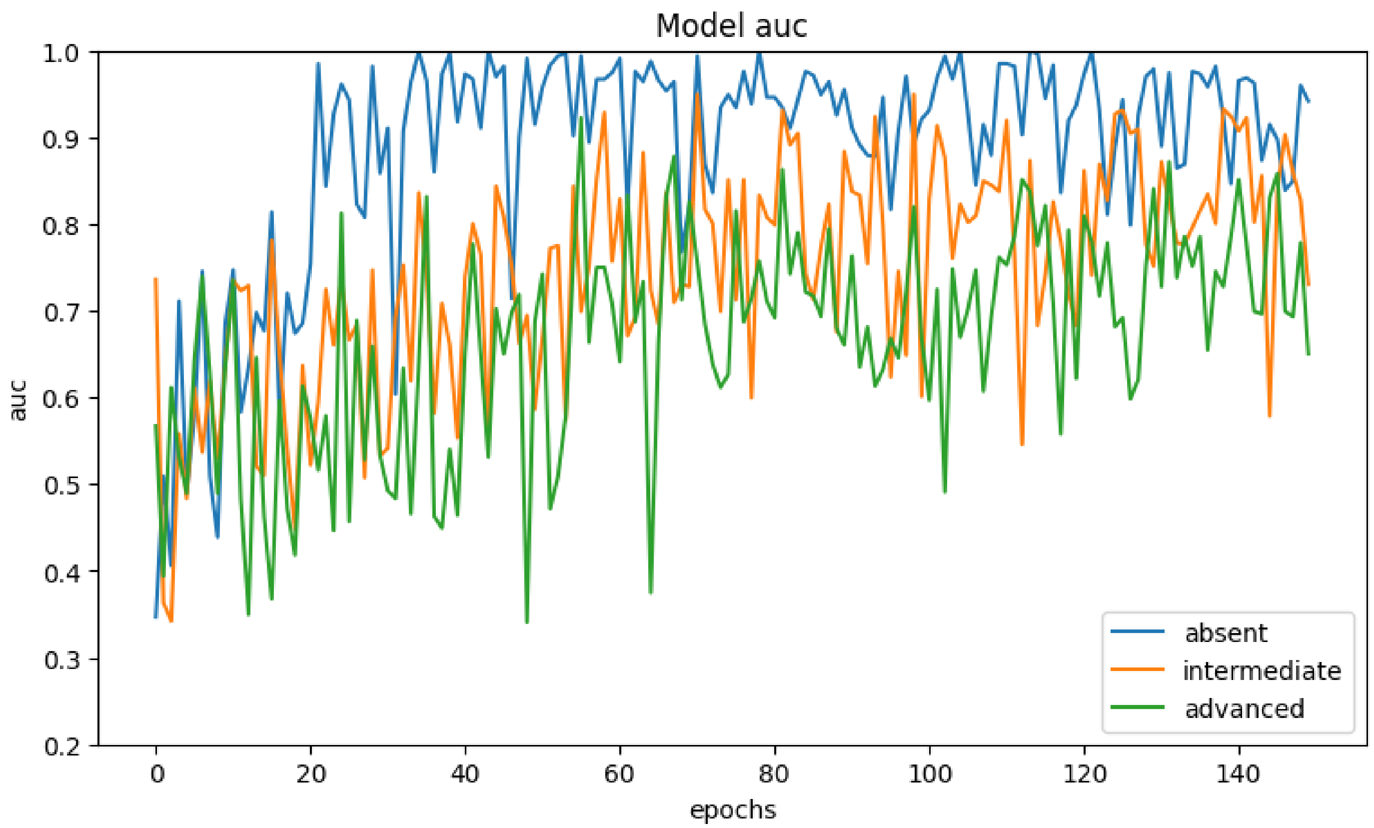
3D CNN model:

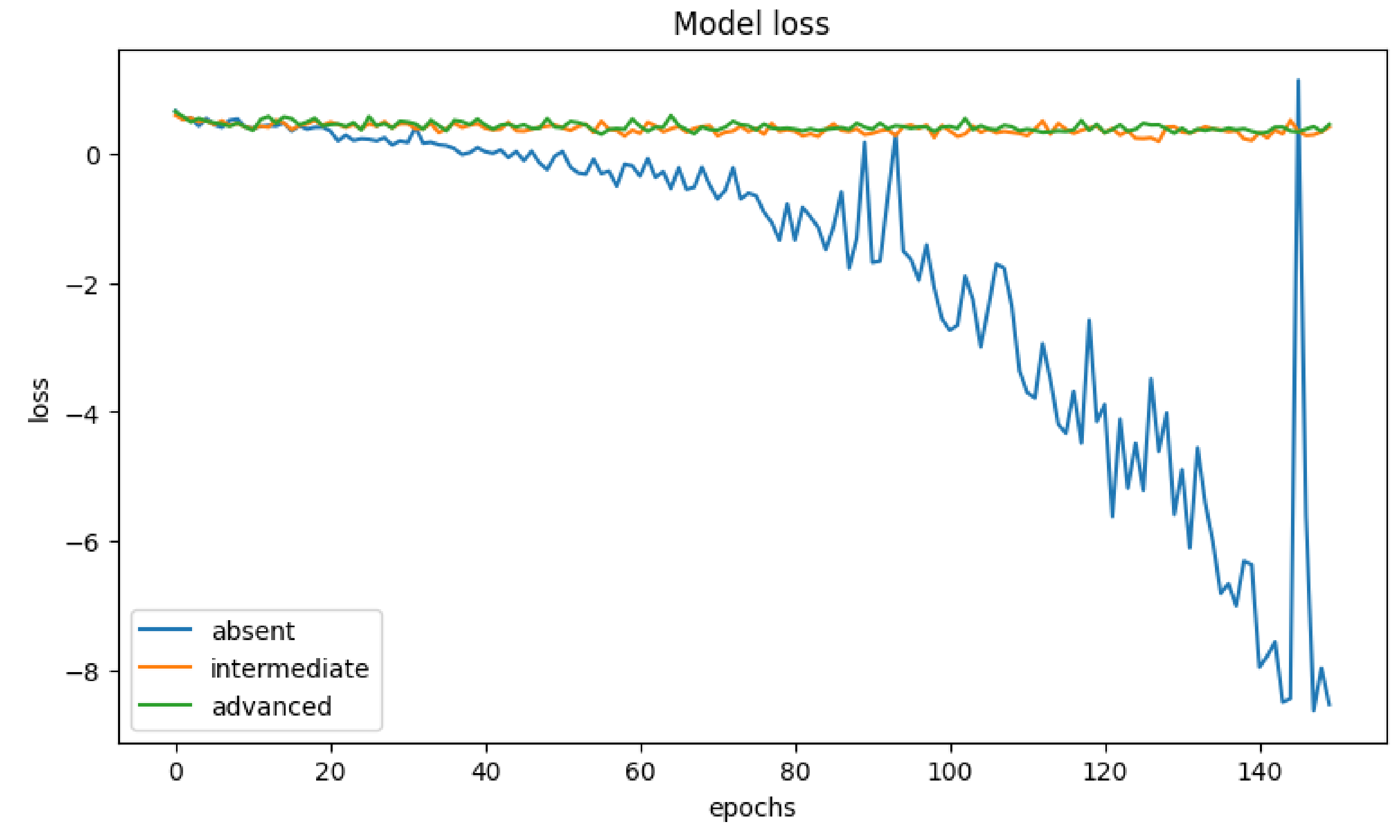
→ Training for 150 epochs in 4 different conditions:

→ Absent augmentation: all augmentation layers disabled

→ Intermediate augmentation: brightness, contrast, rotation, flip and shift layers enabled

→ Advanced augmentation: all augmentation (i.e. also zoom and shear) layers enabled





3. Results

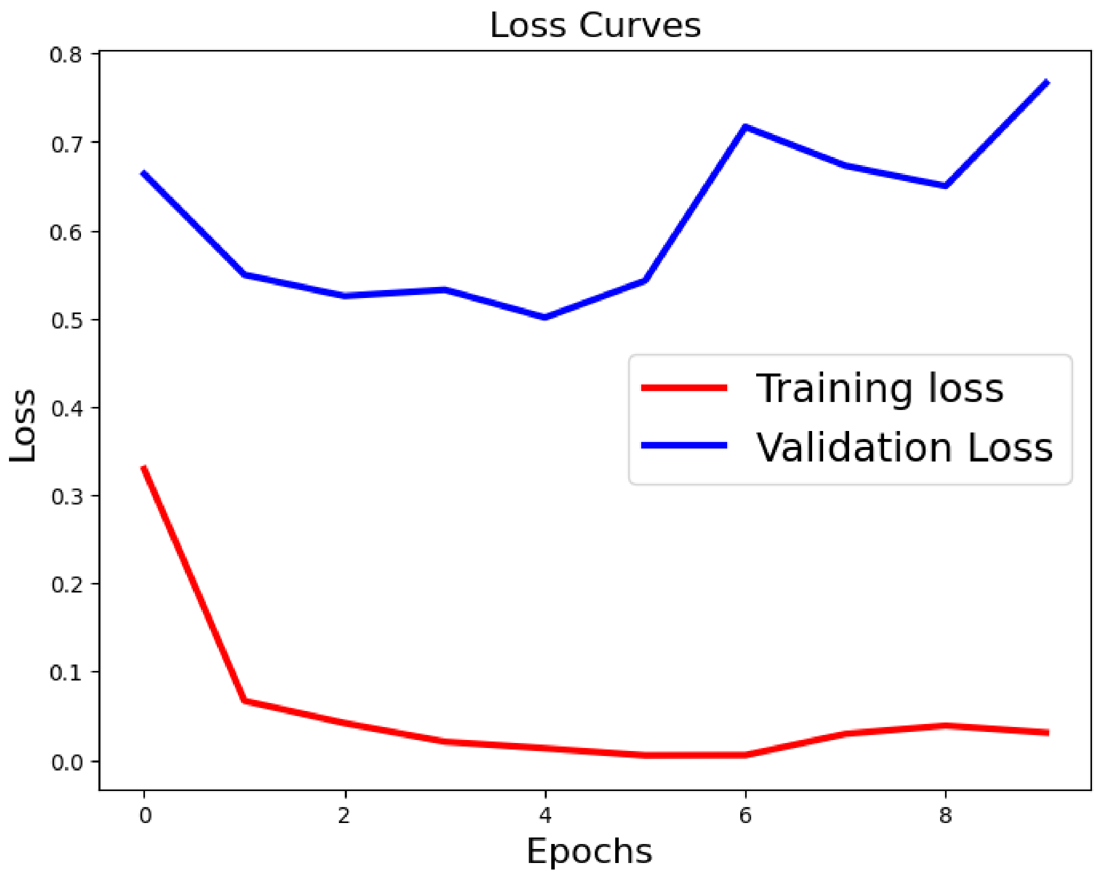
2D ResNet50 model:

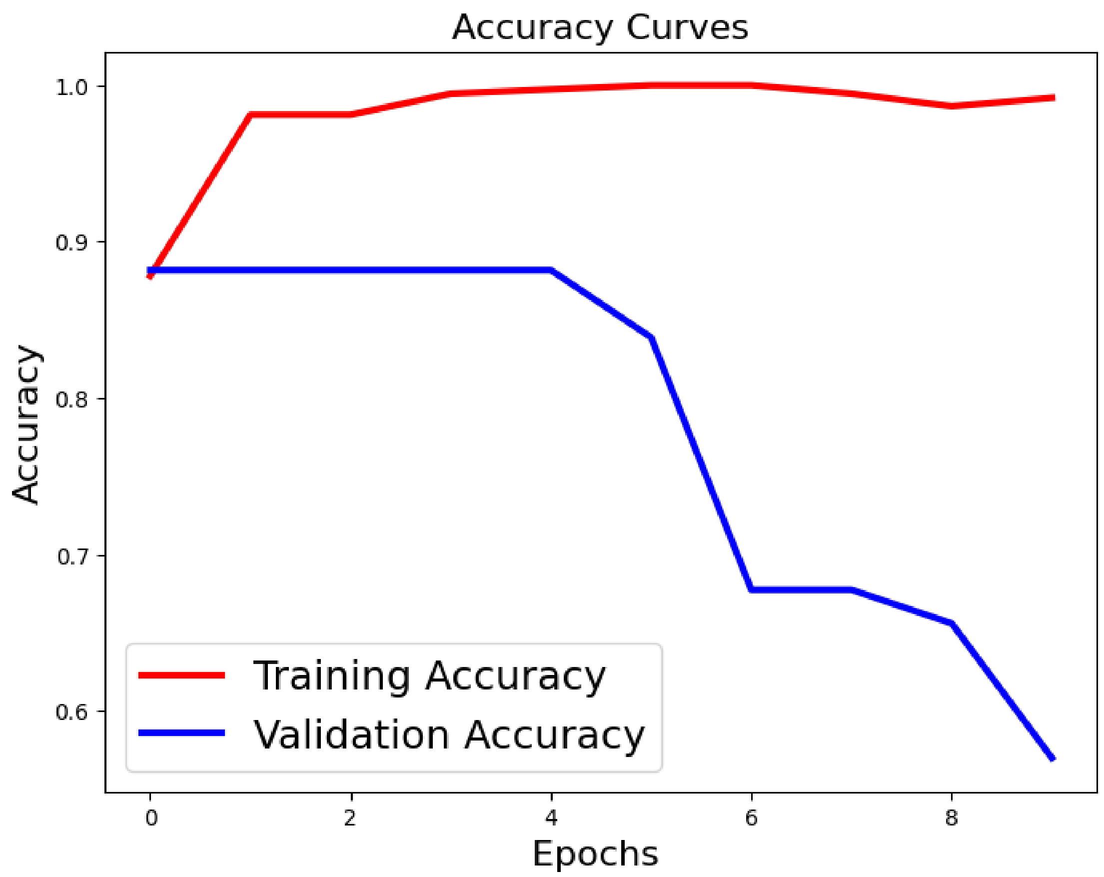
→ Same human brain MRI images dataset classified into 3 classes {meningioma (705), glioma (54), pituitary tumor (7)}

→ 60% training – 40% testing

→ Data augmentation through random transformations, because of the class imbalance (avoid overfitting since our model will never see twice the exact same picture) / BATCH_SIZE = 64, EPOCHS = 50

→ Early Stopping function, able to stop training when a monitored metric stops improving: in our case, early stopping was enabled at epoch 10, restoring model weights from the end of the best epoch: 5.

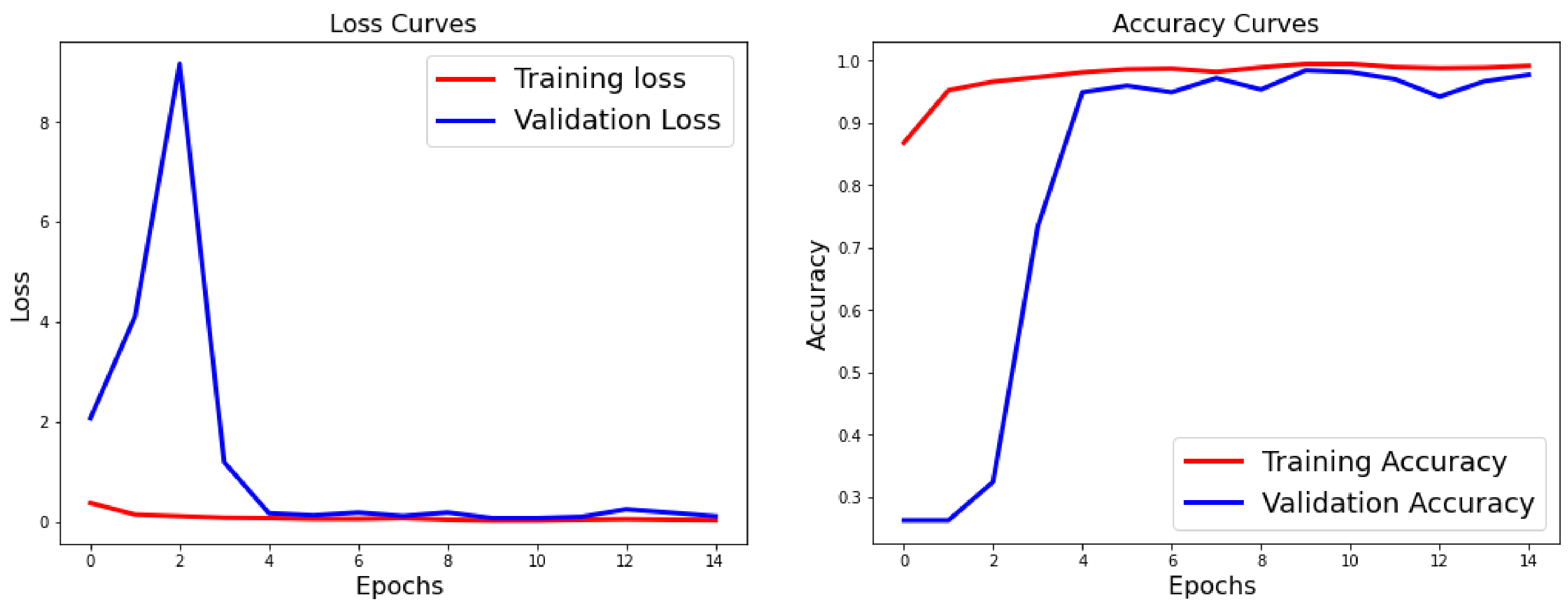




4. Conclusions

- The purpose of the present work is to highlight the differences between 2D CNNs and the learning representation design of 3D CNNs through quantitative analysis and finally to differentiate the use of each single solution.
- The findings are promising and, so, as a next step, we plan to apply the present work in the domain of ischemic strokes, where the accurate detection of the events is a more demanding process as well as the evolution of these events over time.

- Also, if we expose 2D ResNet50 to a wider dataset where 7022 images of human brain MRI images are classified into 4 classes: {glioma, meningioma, no tumor, pituitary}, the curves differentiate as follows:



- So, we intend to validate that our 3D CNN approach performs better in this extended dataset too.

5. References

1. Anjali Gautam, Balasubramanian Raman, 2021. Towards effective classification of brain hemorrhagic and ischemic stroke using CNN. Biomedical Signal Processing and Control, Volume 63, 102178, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2020.102178.
2. Lee, KY., Liu, CC., Chen, D.YT. et al., 2023. Automatic detection and vascular territory classification of hyperacute staged ischemic stroke on diffusion weighted image using convolutional neural networks, Sci Rep 13, 404, https://doi.org/10.1038/s41598-023-27621-4.
3. Öman, O., Mäkelä, T., Salli, E. et al. 2019. 3D convolutional neural networks applied to CT angiography in the detection of acute ischemic stroke. Eur Radiol Exp 3, 8. https://doi.org/10.1186/s41747-019-0085-6.
4. Kumar, S. and Kumar, D., 2023. Human brain tumor classification and segmentation using CNN. Multimedia Tools and Applications, 82(5), pp.7599-7620.
5. He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
6. Ahmed, E., Saint, A., Shabayek, A.E.R., Cherenkova, K., Das, R., Gusev, G., Aouada, D. and Ottersten, B., 2018. A survey on deep learning advances on different 3D data representations. arXiv preprint arXiv:1808.01462.
7. https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset