Classification of Intracranial Hemorrhages Using Vision Transformers and Hybrid ViT-RNN Architectures

Noah Márquez Vara, Felix Bjerhem Aronsson

Chalmers University of Technology
Electrical Engineering Department

CHALMERS UNIVERSITY OF TECHNOLOGY

Problem Statement & Motivation

Intracranial hemorrhages (ICH) are medical emergencies requiring rapid diagnosis. We developed deep learning models for classifying ICH subtypes (epidural, intraparenchymal, intraventricular, subarachnoid, subdural) from CT scans using a subset of the originial RSNA ICH dataset (≈874,000 slices) [4]. This aims to address radiologist shortages by enabling AI triage in resource-limited settings.

Previous studies have used advanced machine learning techniques for ICH detection, including deep learning models achieving expert-level performance [1, 3], hybrid 3D/2D convolutional neural networks (CNNs) [2], and three-dimensional joint convolutional and recurrent neural networks [5].

We mainly innovate in the way we preprocess the dataset while using it with Vision Transformers (ViTs) for end-to-end feature learning on medical data and hybrid ViT+RNN for sequential volumetric analysis, extending beyond 2D slice-wise classification.

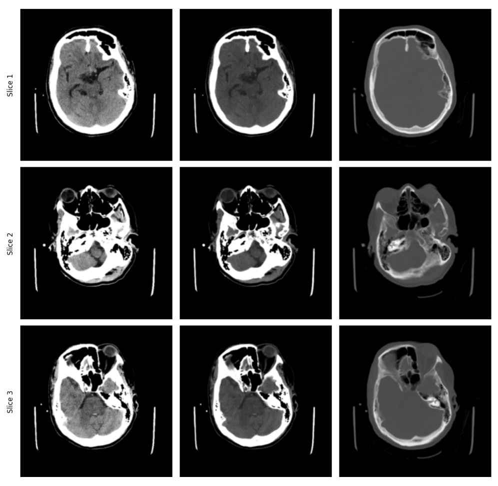

Figure 1: Examples of ICH subtypes in CT scans (Intraventricular type).

Dataset & Preprocessing

Dataset: Subset of RSNA ICH [4] (≈ 2,285 samples), 80% train, 10% val, 10% test.

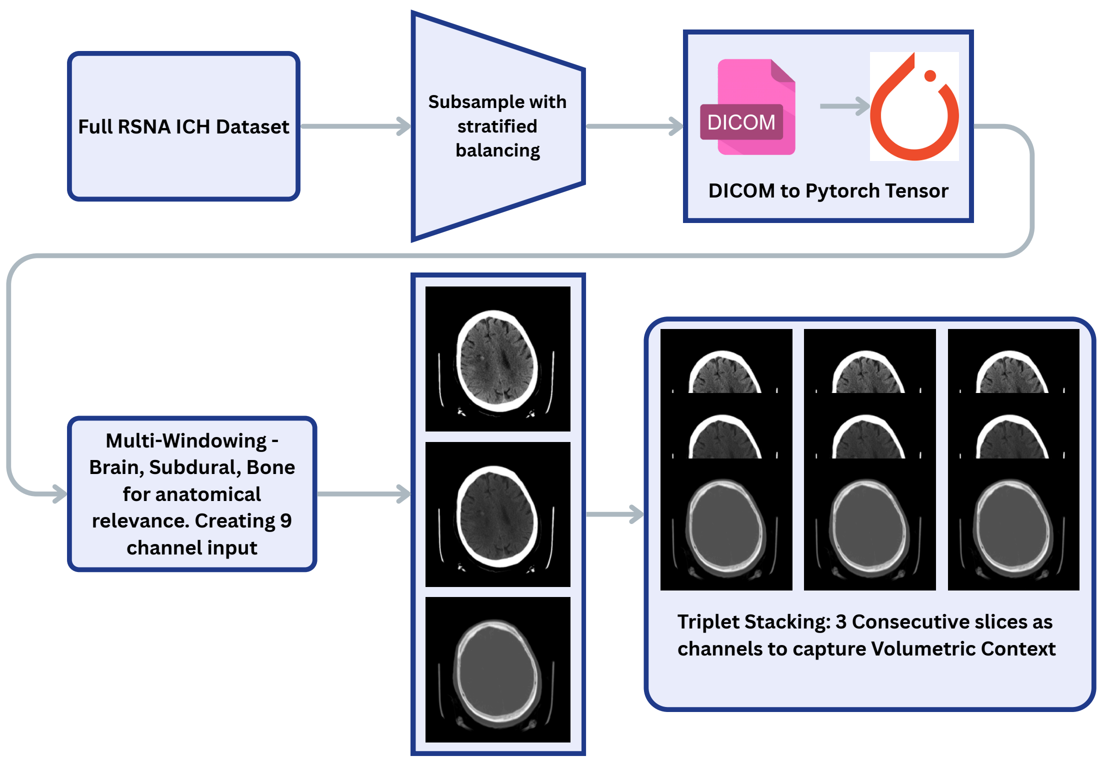

Figure 2: Preprocessing Pipeline.

Model architecture

ViT Implementation:

- Pretrained ViT-B/16 adapted for 9-channel inputs and multi-label binary classification output per slice.
- FCNN classifier on ViT features.

ViT+RNN (Hybrid):

- Bidirectional LSTM processes sequence across patient volume
- Sum-of-products aggregation for multi-slice decisions.

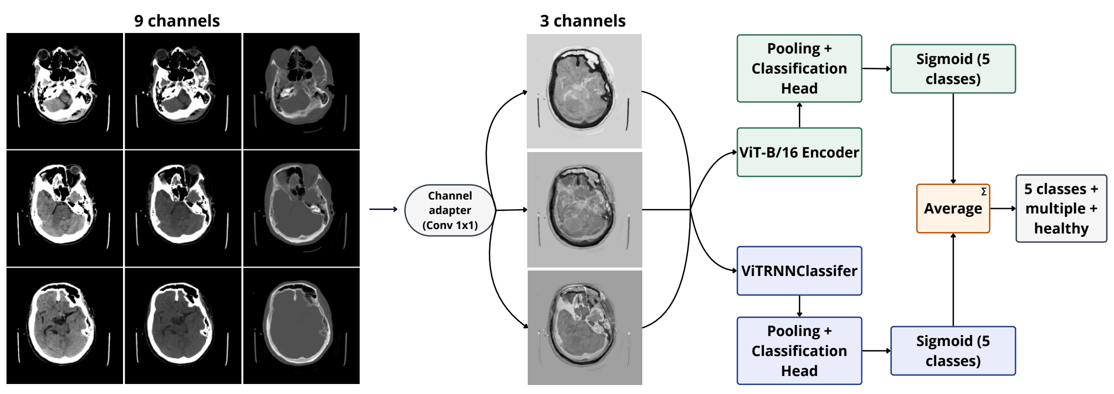

Figure 3: ViT and ViT+RNN architectures with channel adapter.

Results & Evaluation

Metric	Score
Macro AUC-ROC	0.9489
Micro AUC-ROC	0.9610
Weighted AUC-ROC	0.9523
Macro F1	0.7578
Micro F1	0.7845
Weighted F1	0.7867
Macro Precision	0.6870
Macro Recall	0.8476
Exact Match Accuracy	0.6201
Hamming Accuracy	0.8978

Table 1: Ensemble performance metrics on an imbalanced dataset.

Model	Macro AUC-ROC	Micro AUC-ROC
ViT	0.9389	0.9572
ViT+RNN	0.8767	0.8607
Ensemble	**0.9489**	**0.9610**

Table 2: Comparison of models by AUC-ROC.

The ensemble model (ViT + ViT+RNN) demonstrates robust performance in classifying intracranial hemorrhage subtypes, as shown in Table 1. It achieves a Macro AUC-ROC of 94.89% and a Micro AUC-ROC of 96.10%, with strong individual subtype performance: 99.34% for intraventricular, 97.36% for intraparenchymal, 93.47% for subarachnoid, 92.37% for epidural, and 91.90% for subdural. The model also excels in detecting healthy cases (97.13%) but shows a lower AUC-ROC of 75.97% for multiple hemorrhages, indicating challenges with complex cases. Table 2 highlights the ensemble's superior Macro and Micro AUC-ROC over standalone ViT (93.90%, 95.50%) and ViT+RNN (87.67%, 89.20%), underscoring the benefit of combining both architectures for enhanced diagnostic accuracy.

Conclusion & Clinical Impact

The ensemble model, integrating ViT and hybrid ViT+RNN architectures, achieves a macro AUC-ROC of 94.89%, surpassing individual models and establishing competitive performance in intracranial hemorrhage classification on the RSNA dataset. Our innovative multi-window preprocessing and channel adaptation techniques enable effective usage of pretrained ViTs on multi-channel CT imaging data, while the hybrid RNN component captures crucial volumetric and sequential dependencies across scan slices. These advancements extend beyond traditional 2D classification, providing a more comprehensive analysis of 3D medical volumes. Clinically, this system offers substantial potential for emergency department triage, accelerating diagnosis in time-sensitive scenarios, alleviating radiologist workload in underserved areas, and ultimately contributing to improved patient outcomes by minimizing delays in critical interventions.

Limitations & Future Work

The primary limitation of this study stems from the constrained dataset size, as limited computational resources restricted us to using only a small subset the original RSNA ICH dataset, potentially leading to overfitting in the Vision Transformer (ViT) due to insufficient diverse training samples. Additionally, class imbalance, particularly the rarity of epidural hemorrhages, may reduce model performance on underrepresented subtypes. The use of fixed windowing presets (brain, subdural, bone) limits flexibility in capturing anatomical variations, potentially affecting feature extraction. Future work will focus on utilizing the full dataset with enhanced computational resources, implementing synthetic augmentation to address class imbalance, exploring adaptive windowing techniques, employing Grad-CAM visualization to enhance model interpretability, and conducting clinical validation to ensure practical applicability in medical settings.

References

[1] M. R. Arbabshirani, B. K. Fornwalt, G. J. Mongelluzzo, J. D. Suever, B. D. Geise, A. A. Patel, and G. J. Moore. Advanced machine learning in action: Identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. *NPJ Digital Medicine*, 1:9, Apr 2018.

[2] P. D. Chang, E. Kuoy, J. Grinband, B. D. Weinberg, M. Thompson, R. Homo, J. Chen, H. Abcede, M. Shafie, L. Sugrue, and C. G. Filippi. Hybrid 3d/2d cnn for hemorrhage evaluation on head ct. *American Journal of Neuroradiology*, 39(9):1609–1616, Sep 2018.

[3] W. Kuo, C. Häne, P. Mukherjee, J. Malik, and E. L. Yuh. Expert-level detection of acute intracranial hemorrhage on head computed tomography using deep learning. *Proceedings of the National Academy of Sciences*, 116(45):22737–22745, Nov 2019.

[4] Radiological Society of North America. Rsna intracranial hemorrhage detection dataset, Sep 2019. Kaggle Competition.

[5] H. Ye, F. Gao, Y. Yin, D. Guo, P. Zhao, Y. Lu, S. Wang, J. Jin, and Y. Wang. Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. *European Radiology*, 29(7):3562–3571, Jul 2019.

noahmar@chalmers.se, felixbj@chalmers.se