TranSOP: Transformer-based Multimodal Classification for Stroke Treatment Outcome Prediction

Zeynel A. Samak[1,*] Philip Clatworthy[2,3] Majid Mirmehdi[1]

[1] Department of Computer Science, University of Bristol, Bristol, UK,
[2] Translational Health Sciences, University of Bristol, Bristol, UK,
[3] Stroke Neurology, Southmead Hospital, North Bristol NHS Trust, Bristol, UK

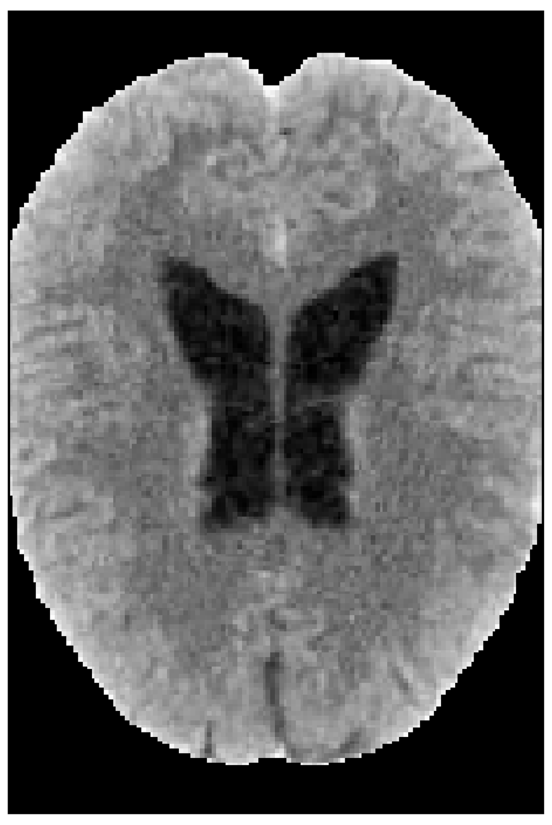
*zeynel.samak@bristol.ac.uk

1. Motivation

- Acute ischaemic stroke, caused by an interruption in blood flow to brain tissue, is a leading cause of disability and mortality worldwide.
- The selection of patients for the most optimal ischaemic stroke treatment is a crucial step for a successful outcome, as the effect of treatment highly depends on the time to treatment.
- Avoiding treatment where risks are highest

3. Dataset

- MR CLEAN trial dataset [1]
- 500 patients from 16 medical centers, NCCT volumes
- Clinical metadata comprises, such as patient demographics, medical history and the stroke metrics



2. Contributions

Problem: predict the successful rate (functional outcome) of ischaemic stroke treatment (thrombectomy) from baseline 3D non-contrast computed tomography (NCCT) volume (the first scan when the patient was admitted to hospital) and clinical metadata.

- A transformer-based multimodal network (TranSOP) to predict the functional outcome of stroke treatment.
- A fusion module to efficiently combine NCCT features and clinical information.
- Achieve a state of the art AUC score of 0.85.

4. Proposed Method

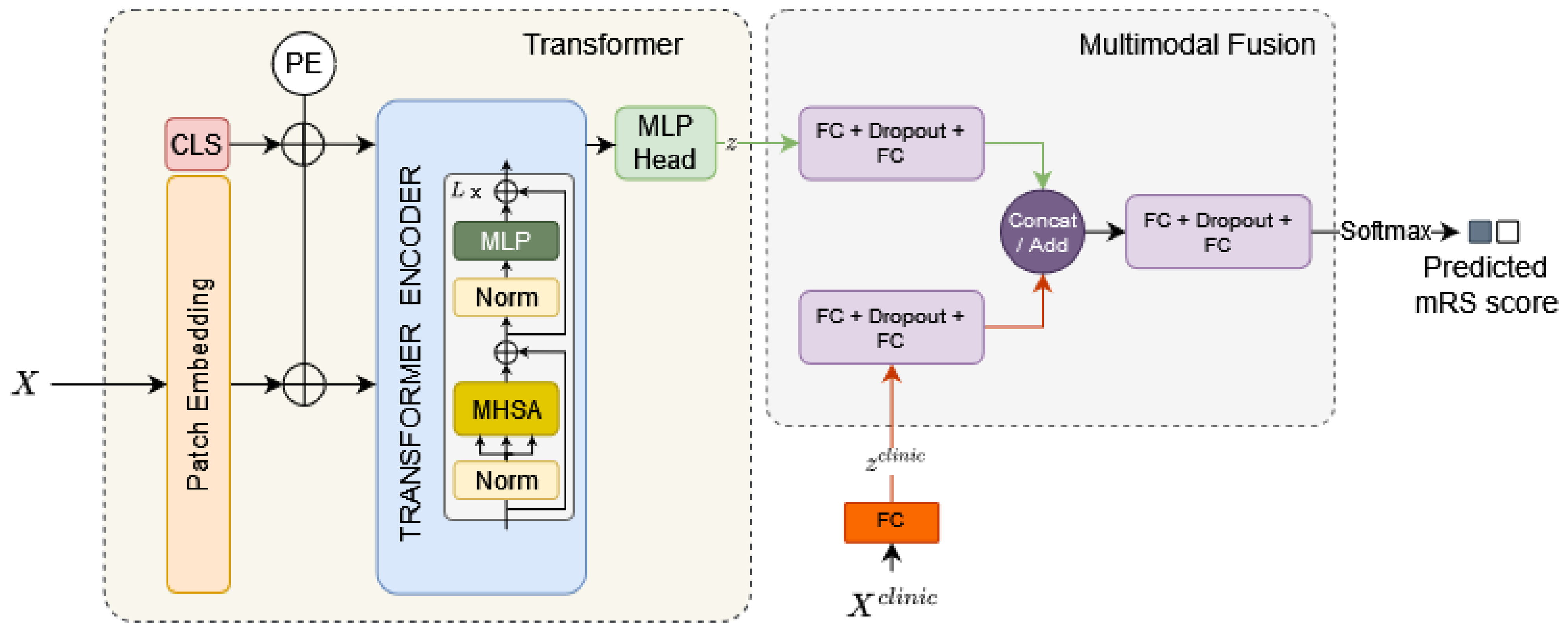


Figure 1: Overview of our proposed transformer-based multimodal architecture, TranSOP. PE: positional encoding, CLS: a token/vector that represents the input volume for classification, MHSA: Multi-head self-attention, MLP; multi-layer perceptron, FC: fully connected layer.

5. Results

	w/o Clinical Records				with Clinical Records		
Method	ACC (95% CI)	F1-score (95% CI)	AUC (95% CI)	Fusion	ACC (95% CI)	F1-score (95% CI)	AUC (95% CI)
ClinicDNN*	-	-	-	-	0.75 (0.65-0.85)	0.44 (0.19-0.64)	0.73 (0.57-0.86)
Samak et al[2]	<u>0.72</u> (0.62-0.82)	0.33 (0.09-0.53)	0.63 (0.44-0.81)	concat	0.77 (0.66-0.87)	0.47 (0.18-0.67)	0.78 (0.63-0.91)
				add	<u>0.79</u> (0.69-0.89)	0.44 (0.17-0.67)	0.71 (0.51-0.88)
Bacchi et al[3]	**0.75** (0.65-0.85)	0.40 (0.16-0.60)	<u>0.66</u> (0.48-0.80)	concat	0.73 (0.62-0.83)	0.51 (0.29-0.68)	0.78 (0.62-0.90)
				add	0.73 (0.62-0.83)	0.51 (0.29-0.68)	0.78 (0.62-0.90)
TranSOP$_{ConViT}$	0.58 (0.46-0.69)	0.40 (0.21-0.56)	**0.67** (0.46-0.85)	concat	0.77 (0.68-0.87)	<u>0.58</u> (0.36-0.74)	0.83 (0.72-0.93)
				add	0.77 (0.68-0.87)	<u>0.58</u> (0.36-0.74)	0.82 (0.71-0.92)
TranSOP$_{DeiT}$	0.58 (0.46-0.69)	0.40 (0.21-0.56)	0.63 (0.44-0.80)	concat	0.77 (0.68-0.86)	0.53 0.30-0.71)	0.82 (0.68-0.93)
				add	<u>0.79</u> (0.69-0.89)	0.52 (0.27-0.71)	<u>0.84</u> (0.71-0.94)
TranSOP$_{ViT}$	0.58 (0.46-0.69)	0.40 (0.21-0.56)	0.60 (0.40-0.78)	concat	**0.80** (0.70-0.89)	0.53 (0.28-0.74)	<u>0.84</u> (0.72-0.94)
				add	**0.80** (0.70-0.89)	**0.59** (0.35-0.76)	0.83 (0.71-0.93)
TranSOP$_{SwinT}$	0.58 (0.46-0.69)	0.40 (0.21-0.56)	0.64 (0.44-0.82)	concat	0.76 (0.66-0.86)	0.54 (0.32-0.71)	0.83 (0.71-0.93)
				add	<u>0.79</u> (0.69-0.89)	0.55 (0.31-0.73)	**0.85** (0.75-0.94)

* A method that uses only clinical metadata information.

Table 1: Results of the models with and without clinical records. The best and second best results are shown in bold and underlined respectively. The second and third rows are convolutional-based models. CI is confidence interval.

6. Conclusions

- Transformer models outperformed convolutional architectures in multimodal settings.
- The transformer models, although not performing as well on only imaging data, can learn better complementary imaging information when combined with clinical metadata.
- In future work, we plan to investigate and explore a data-efficient transformer model for small image datasets

7. References

[1] Olvert A. Berkhemer and et al. A randomized trial of intraarterial treatment for acute ischemic stroke. *NEJM*, 372(1):11–20, 2015.

[2] Zeynel A. Samak and et al. Prediction of thrombectomy functional outcomes using multimodal data. In *MIUA*, pages 267–279, Cham, 2020. SIP.

[3] Stephen Bacchi and et al. Deep learning in the prediction of ischaemic stroke thrombolysis functional outcomes: A pilot study. *Academic Radiology*, 4 2019.

8. Acknowledgements

The authors would like to thank the Principal Investigators of the MR CLEAN trial: Profs Aad van der Lugt, Diederik W.J. Dippel, Charles B.L.M. Majoie, Yvo B. W.E.M. Roos, Wim H. van Zwam and Robert J. van Oostenbrugge for providing the data. Z.A. Samak is funded by the Ministry of Education (1416/YLSY), the Republic of Türkiye.