Scopeformer: n-CNN-ViT Hybrid Model for Intracranial Hemorrhage Classification
-MIDL 2021-

Barhoumi Yassine
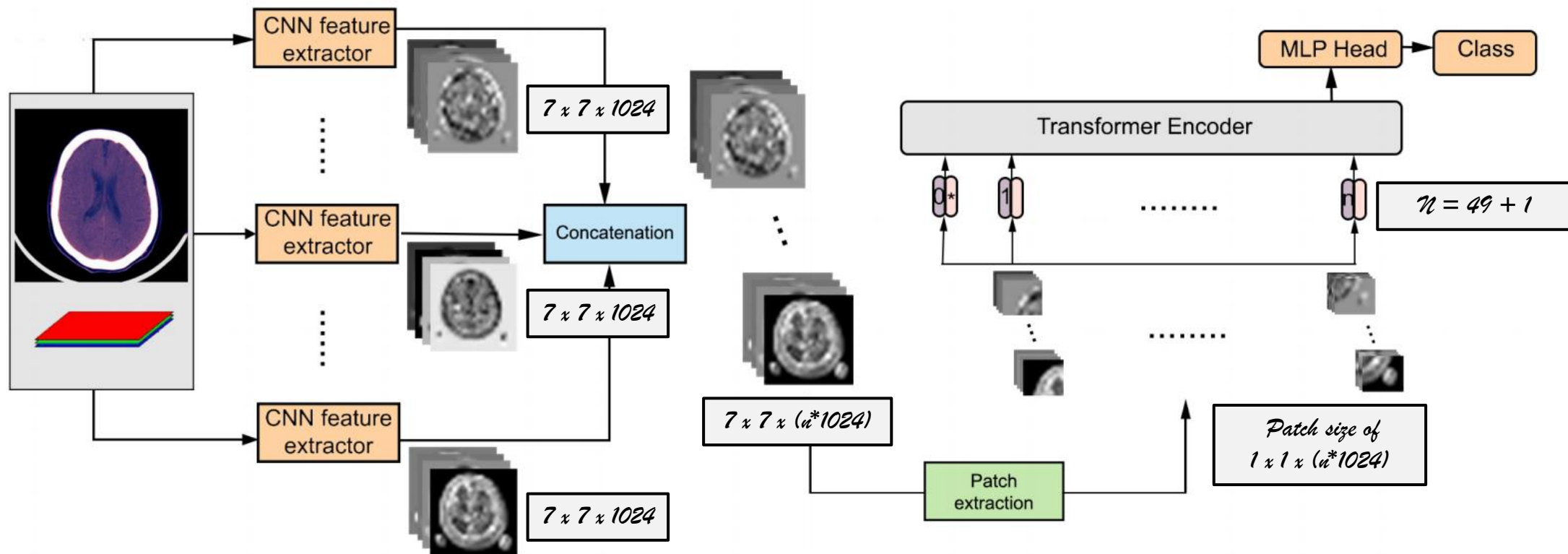
Ghulam Rasool

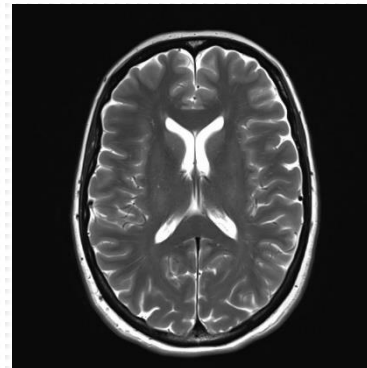Rowan University, New Jersey, USA

Scopeformer Architecture

Pretraining datasets

ImageNet dataset example; Pond fish

- ❑ Two types of datasets were used for pretraining:
  - ❑ ImageNet
  - ❑ Data generated using Style Transfer method applied on ImageNet

Applying style transfer on ImageNet dataset with a brain MRI picture as a target

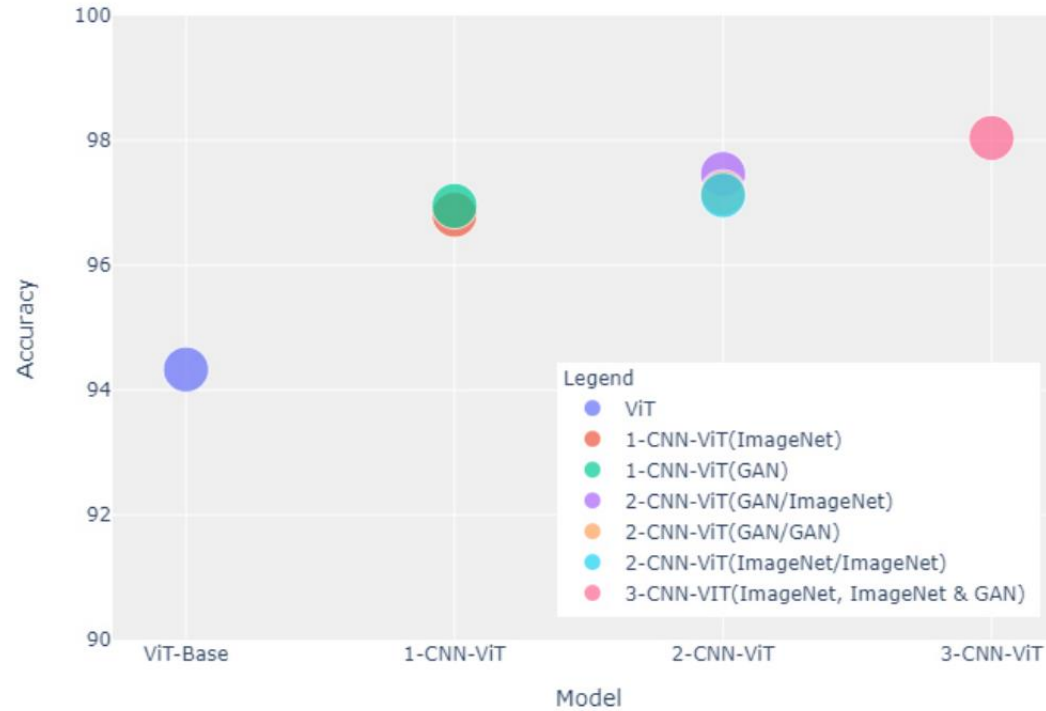Style transferred ImageNet dataset example

Results & Discussion

Table 1: Classification performance of ViT based models on the RSNA validation dataset

Model	ViT input dimension	Validation accuracy	Loss
ViT	$256 \times 256 \times 3$	94.33%	0.1822
1-CNN-ViT (GAN)	$7 \times 7 \times 1024$	96.95%	0.08272
2-CNN-ViT (ImageNet/GAN)	$7 \times 7 \times 2048$	97.46%	0.07754
3-CNN-ViT (ImageNet/ImageNet/GAN)	$7 \times 7 \times 3072$	98.04%	0.07050