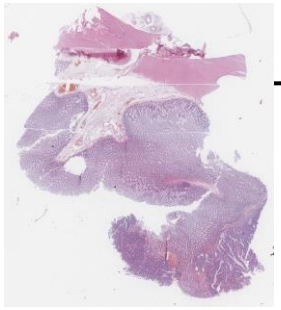
HARDNET-DFUS FOR MULTICLASS COLORECTAL CANCER TUMOR SEGMENTATION FROM METU CHALLENGE DATASET

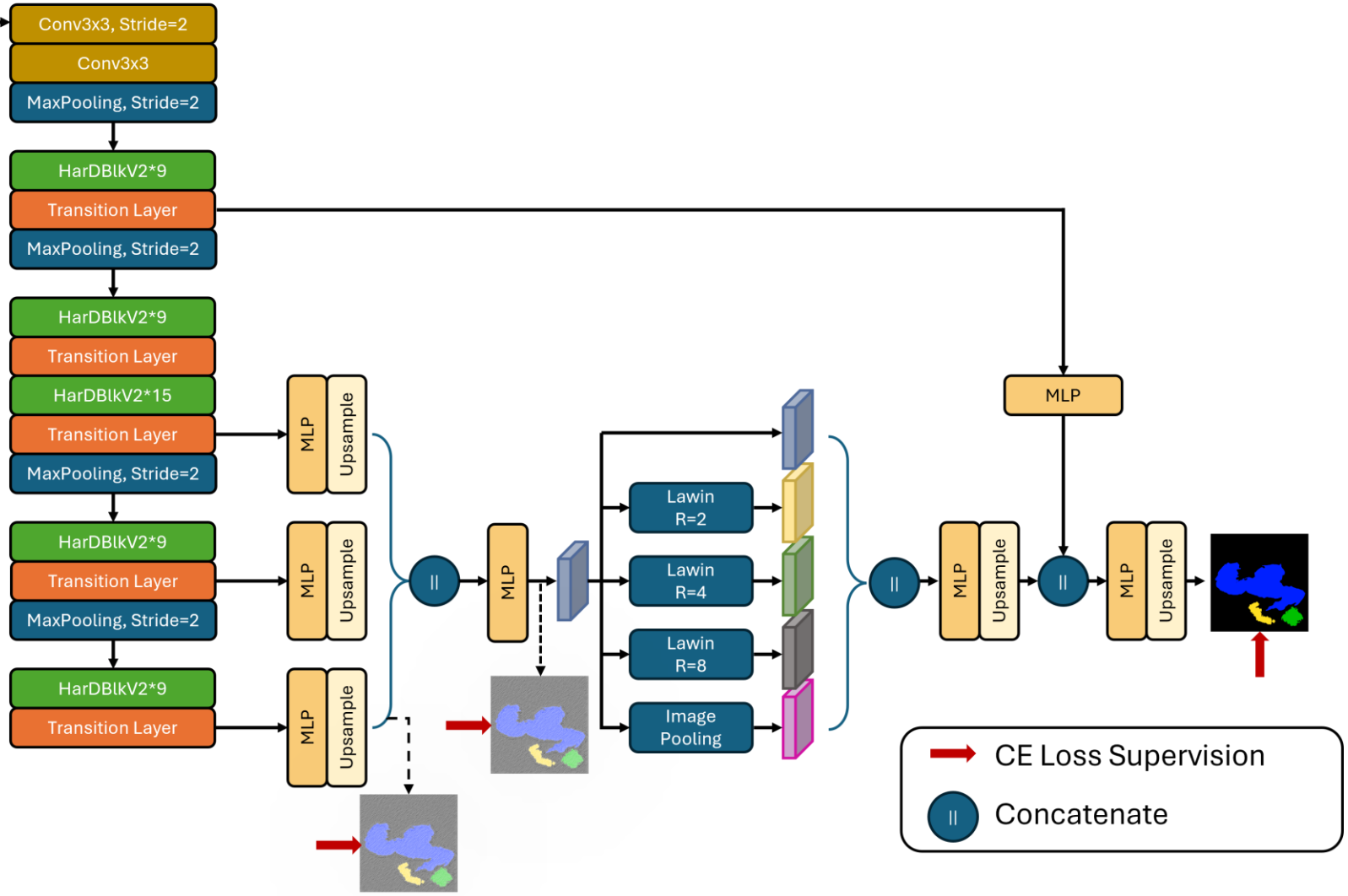
Guillaume PICAUD[1,3], Marc CHAUMONT[1,2], Gérard SUBSOL[1], Luc TEOT[3]
{guillaume.picaud, marc.chaumont, gerard.subsol}@lirmm.fr, l.teot@cicat-occitanie.org

[1] LIRMM, ICAR team, Univ. Montpellier, CNRS, France
[2] Univ. Nîmes Place Gabriel Péri, France
[3] Cicat-Occitanie, Montpellier, France

An architecture using both convolutions and transformers

- A high-performing segmenter [1]
  - Encoder: HarDNet [2]
    Convolution, lightweight 3M
  - Decoder: Lawin [3]
    Transformer-based
- Multi-scale feature maps
- Deep Supervision

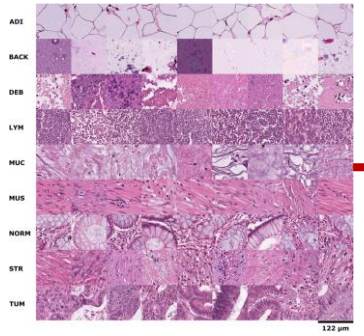




[1] Ting-Yu L. et al. 2022. "Hardnet-dfus: Enhancing backbone and decoder of hardnet-mseg for diabetic foot ulcer image segmentation" Diabetic Foot Ulcers Grand Challenge. Springer
[2] Yan, H. et al. 2022. "Lawin transformer: Improving semantic segmentation transformer with multi-scale representations via large window attention ". *arXiv*
[3] Chao, P. et al. 2019. " Hardnet: A low memory traffic network ". ICCV

Learning procedure

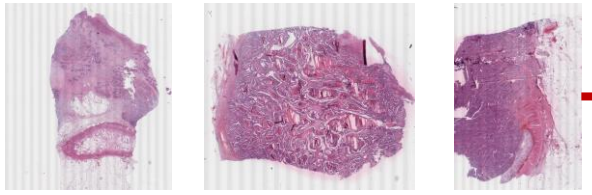
RTX A6000 48G0



external dataset :
NCT-CRC-HE-100K-NONORM [4]

Classification
Pre-training

Keep only encoder weights



Downsized METU : train/val splits

5 folds

DA

Fine-tuning

300 Epochs
lr = 10^{-4}
Cosine annealing
warmup
AdamW & EMA
Size = 1536*1536
Batch=2

5 models
Best epoch F1 score on val split

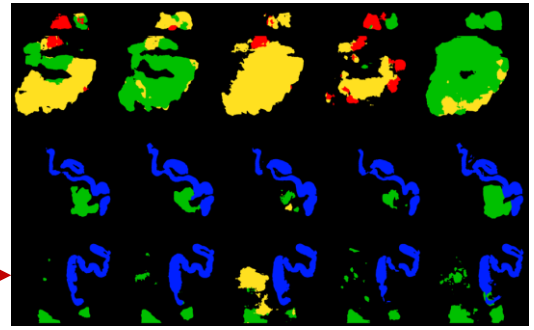
METU : test split

TTA

Inference

Ensembling

Hole filling



[4] Kather, J. et al. 2018. 100,000 histological images of human colorectal cancer and healthy tissue

Results & perspectives

Method	Notes	mFscore	mIoU	mPrecision	mRecall
VAN+UperNet	WSI & ensemble	70.2	56.5	67.2	74.4
DPT+MaxViT	WSI & ensemble	69.8	55.8	68.6	71.7
HardNet+Lawin	ensemble	66.7	52.8	63.0	72.4
Segmenter-L		65.2	51.3	63.7	67.5
Segformer-B1		65.1	50.7	61.2	70.6
PathVTA		64.2	50.2	63.3	65.7
SwinTransformer [11]	baseline	62.9	-	60.9	69.6

Best downsized method of this challenge

- Annotation inconsistency in downsized version?

B-0391(2-3).png

8612-20.png

- Perspectives:
  - Downsized / WSI ?
  - hyperparameters optimization
  - deformable convolution ref
  - Other pre-training datasets
  - Semi or self supervised learning