W-net architecture for bone age prediction

MohammadAmin Farajzadeh

RSNA Pediatric Bone Age Challenge (2017)

As part of its efforts to spur the creation of artificial intelligence (AI) tools for radiology, in 2017 RSNA conducted a challenge to assess bone age from pediatric hand radiographs, a routine task that determines an important developmental indicator.

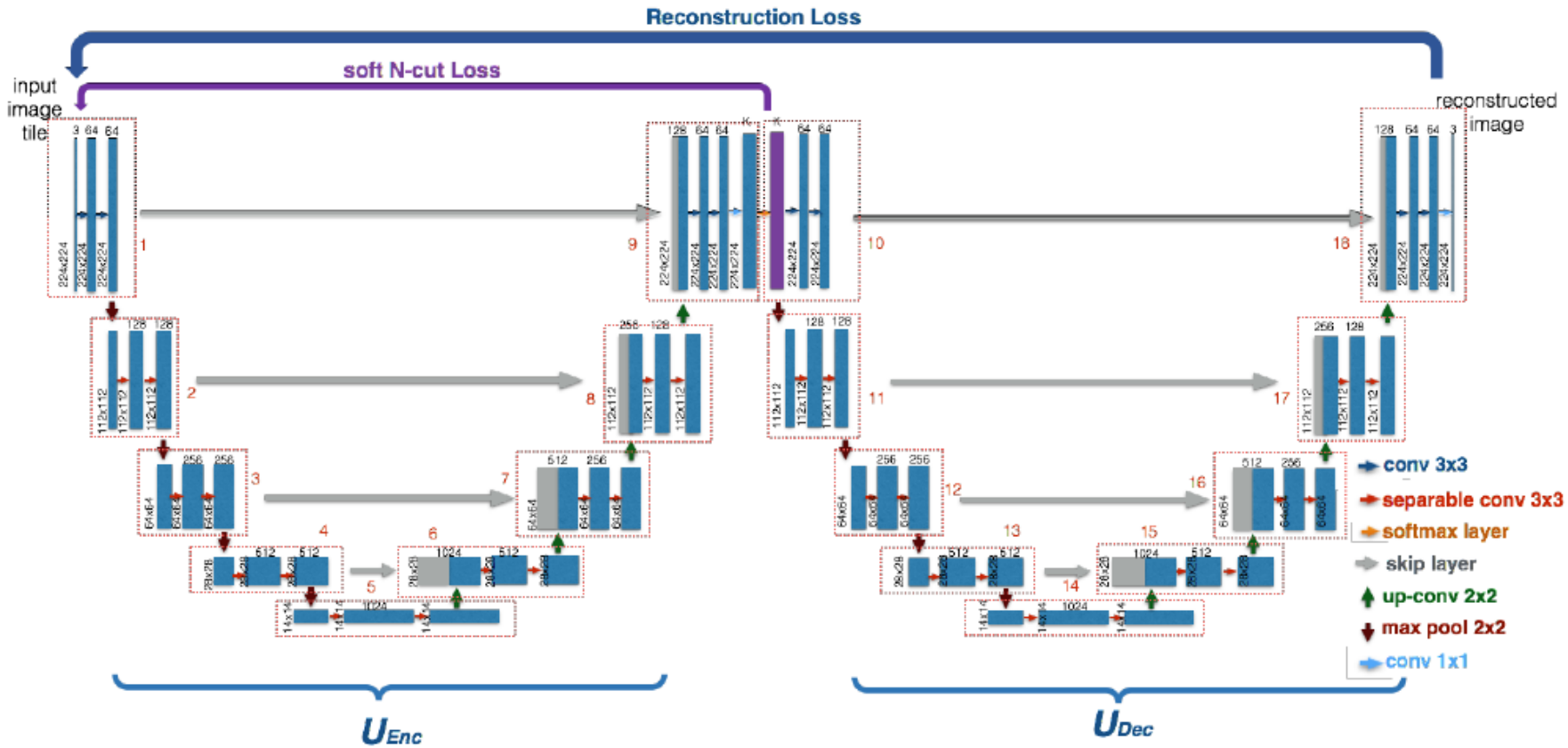
Challenges

- The important challeng in many models for this task is that how we can tell the model specifically which region of the radiographs contain more information and model incorporates the segments of wrist bones .
- There is different solutions for this purpose for example we can make mask for subset of dataset and train a unet for segmentation or we can use unsupervised information theory ased methods for this outcome .
- See
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10017763/

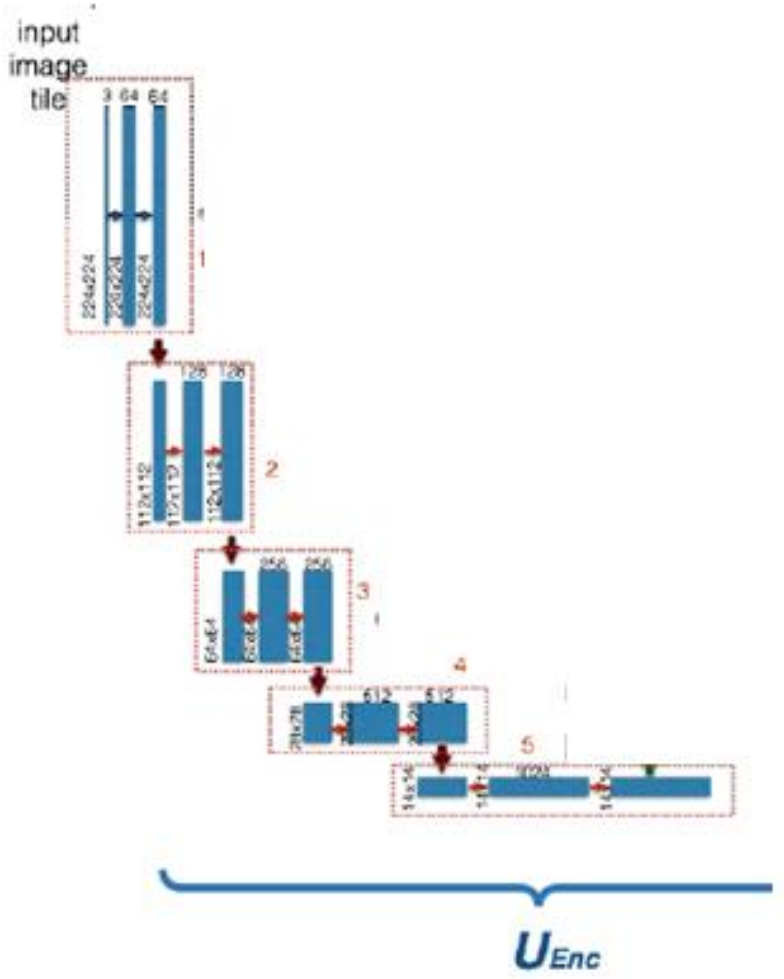
What I propose

- I believe that we can use a unsupervised images segmentation for coming up with a backbone to map the data to latent space and then connect the module to some FCN layers to prediction the architecture we use here is W-net.
- The min idea is that there is two parts and two losses one for auto-encodeing and other for segmentation we try to connect a pretrained U-enc module of the 2-net to 6 FCN.

W-net


Reconstruction Loss
soft N-cut Loss
input image tile
reconstructed image
conv 3x3
separable conv 3x3
softmax layer
skip layer
up-conv 2x2
max pool 2x2
conv 1x1
U_Enc
U_Dec

Our stucture

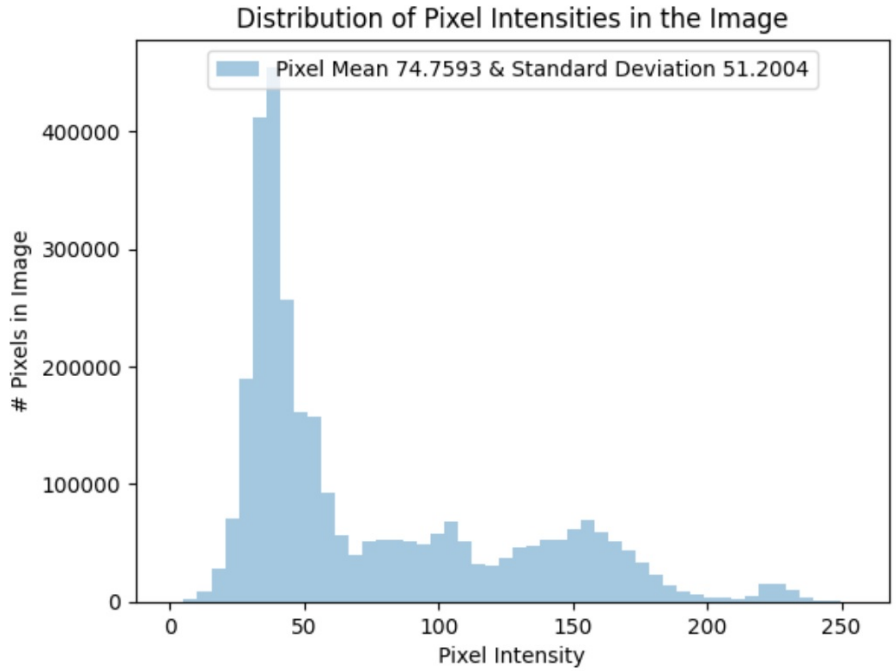


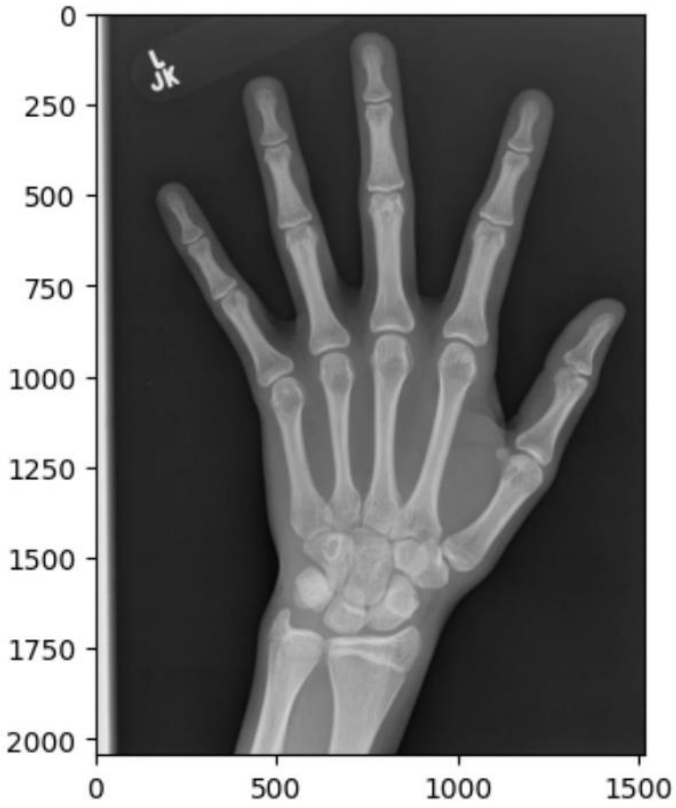
Let’s go to implmemnt

- We use two repositories to implement this model one we use for model architecture and other one for N-Cut-Loss.
- See references and notebook

Data EDA and preprocessing

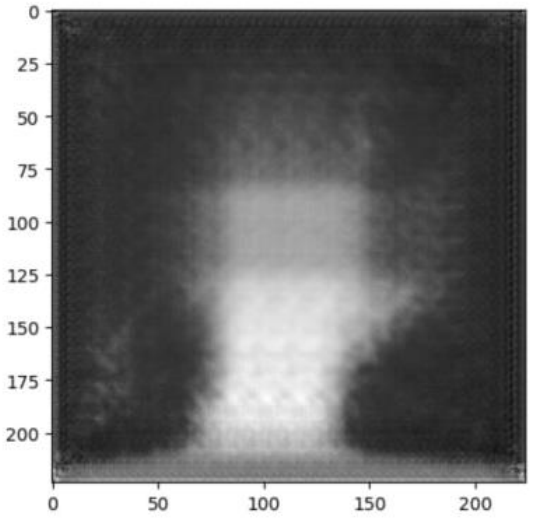
- We normalize the pixel intensity and resixe all images to 224*224:



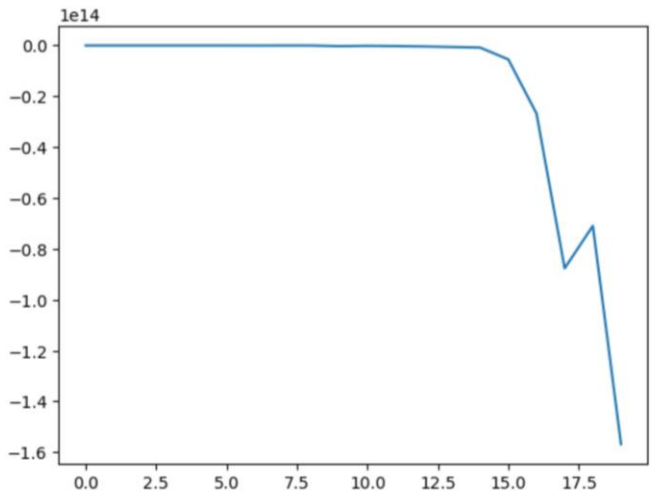


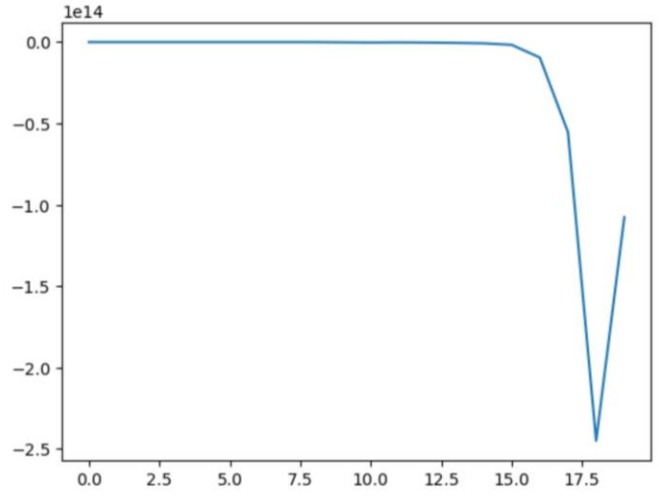
Training eval for w-net (left train set-rightcrossval)

]: <matplotlib.image.AxesImage at 0x7a4ef9f1b550>

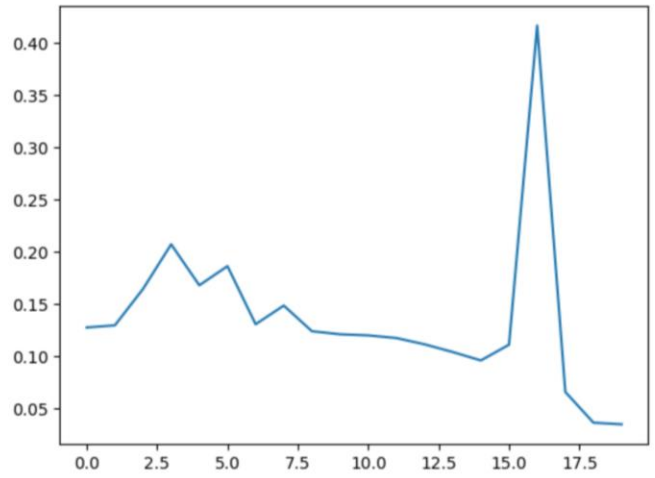


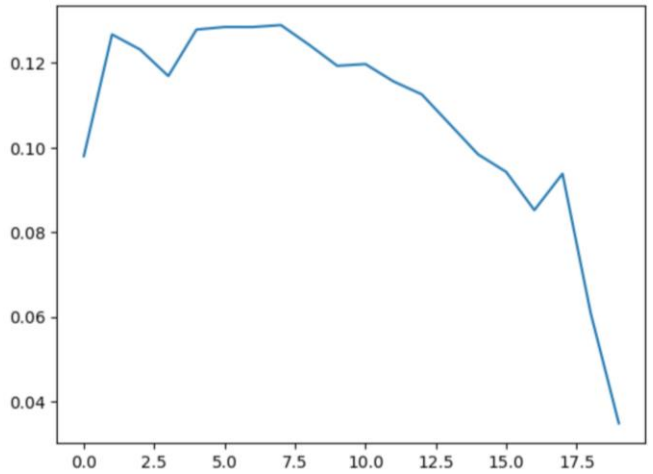
N-cut loss



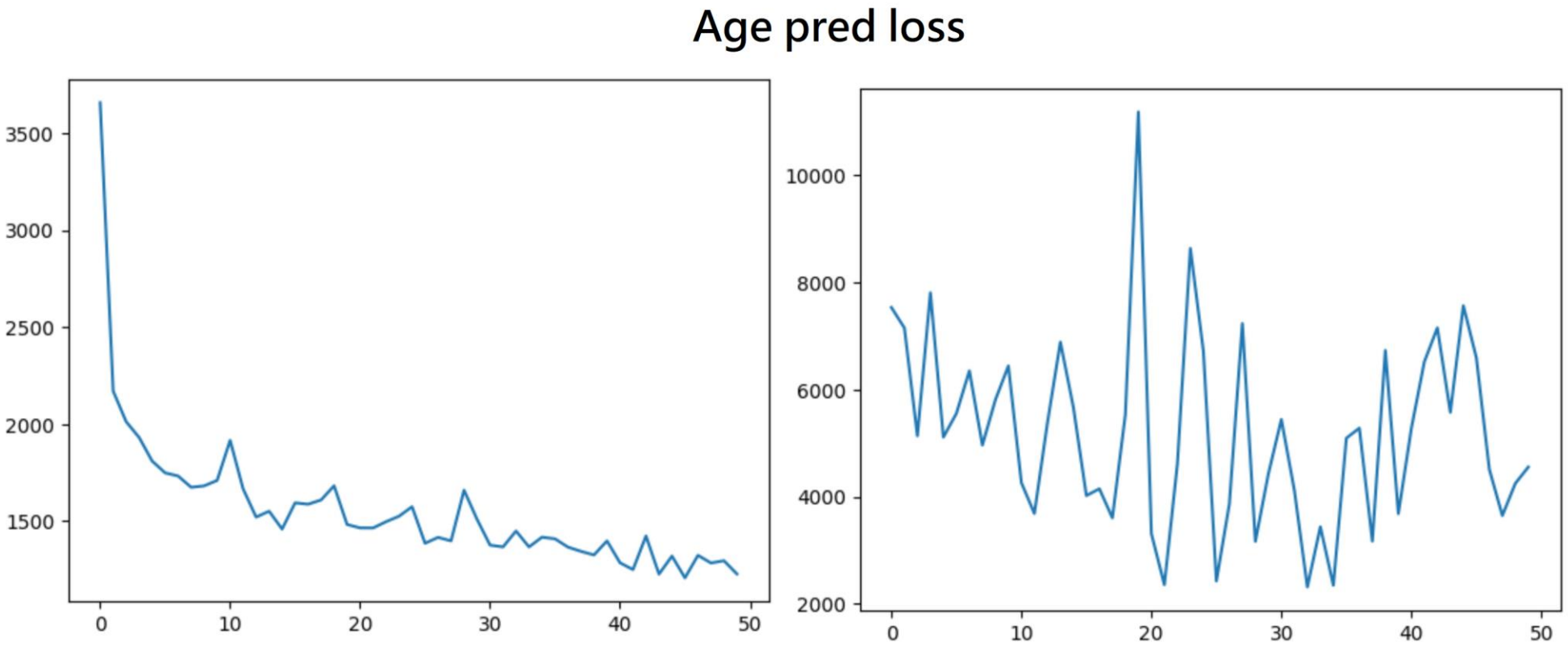


Reconstruction loss





Training eval for age pred



Why is not so good and next steps:

- First of all for a model like w-net takes about 500 epoch to train in many examples but our training is just for 20 eppoches and the same issues for the noisy training of age predictor.
- We can use some high pass filter to filter nonhands parts of radiographs and use them as mask for training a supervised u-net and then apply the method like the paper we mentioned.
- We need augumentation
- Because of limitation in resources and trying to have something quicke to implement we ran the model on 2000 image but we should scale it.

References:

- https://github.com/Andrew-booler/W-Net
- https://github.com/taoroalin/WNet
- Bone age assessment based on deep neural networks with annotation-free cascaded critical bone region extraction
- **Generalizability and Bias in a Deep Learning Pediatric Bone Age Prediction Model Using Hand Radiographs**
- Ridge regression neural network for pediatric bone age assessment