Fundus Image Analysis and Grading for Diabetic Retinopathy

Team: Pooja Krishnan, Veronika Kotova, Shraddha Beedu, Ali Kareem Raja
Tutor: Hasan Sarhan

Technische Universität München + Johns Hopkins Whiting School of Engineering

Agenda

- **Problem Statement**
- **Proposed Approach**
- **MNet**
- **ANet**
- **CNet**
- **Integration**
- **Learning Outcomes**
- **Impediments**

Problem statement

Diabetic retinopathy (DR)

- Leading cause of vision loss in the general population
- Affects up to 80 % of people who have had diabetes for 20 years or more

Challenges

- Detecting the subtle pathological lesion characteristic of early-stage disease
- No large well annotated datasets

Source: https://nei.nih.gov/health/diabetic/retinopathy, accessed on 11.12.2018

Proposed approach

Goal

Convolutional Neural Network based algorithm for:

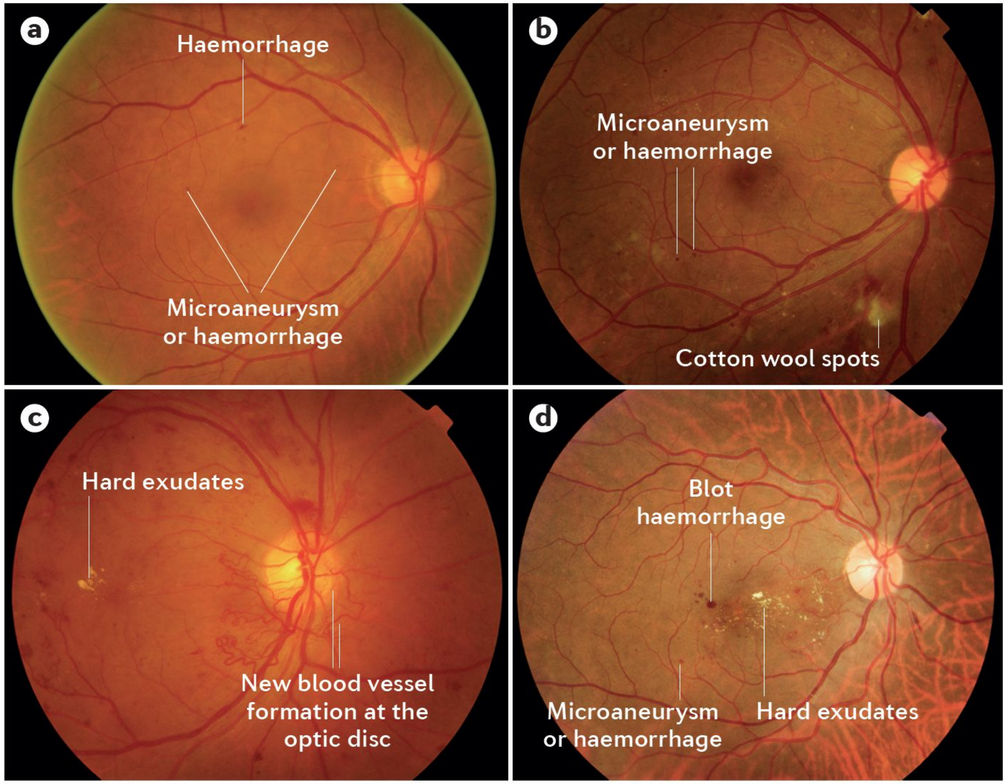
- diagnosing 5 stages of DR
- highlight suspicious regions

DR stages

- 0 – no DR
- 1 – mild
- 2 – Moderate
- 3 – Severe
- 4 – Proliferative

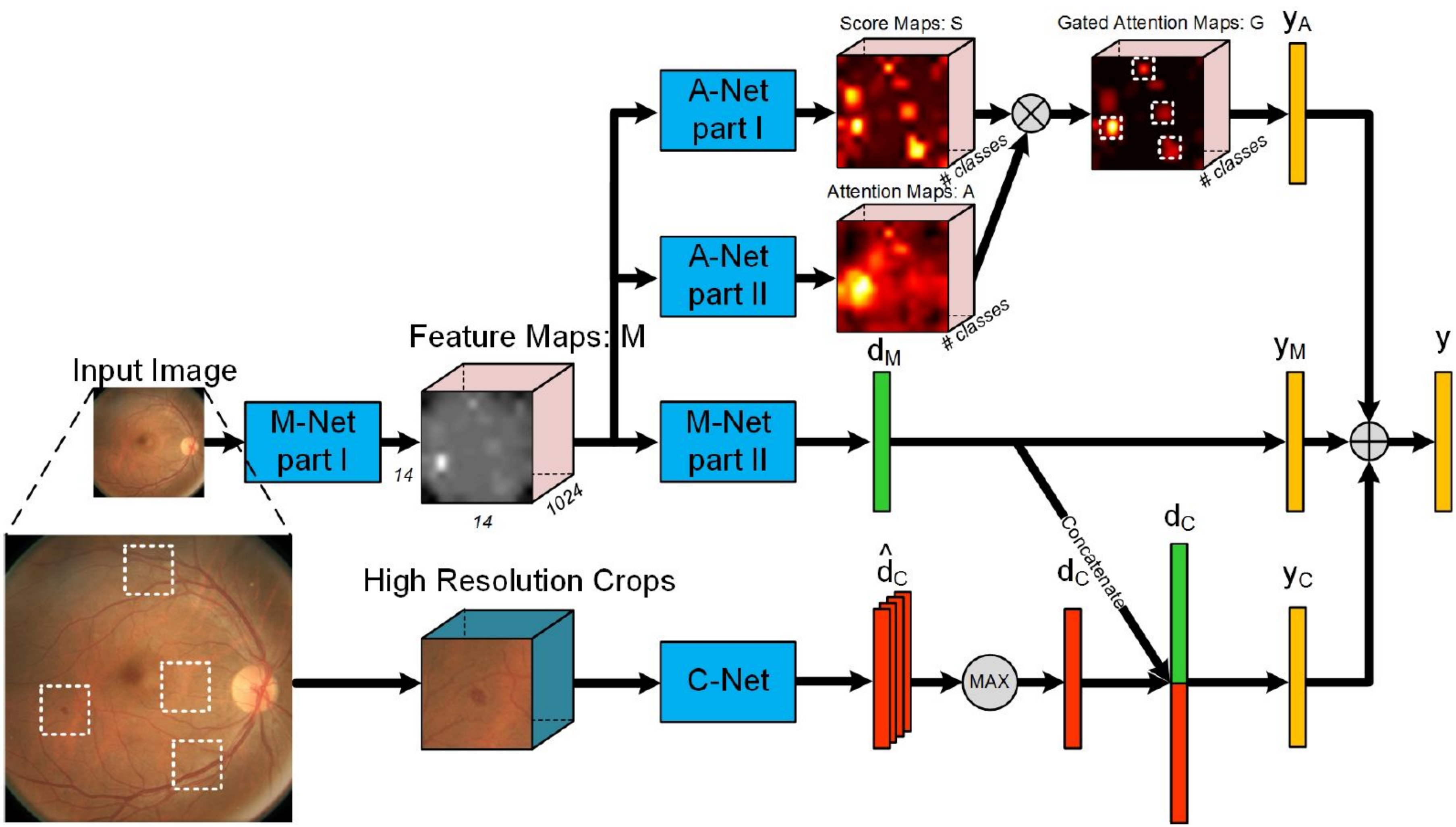
Datasets

- EyePACS
- IDRiD



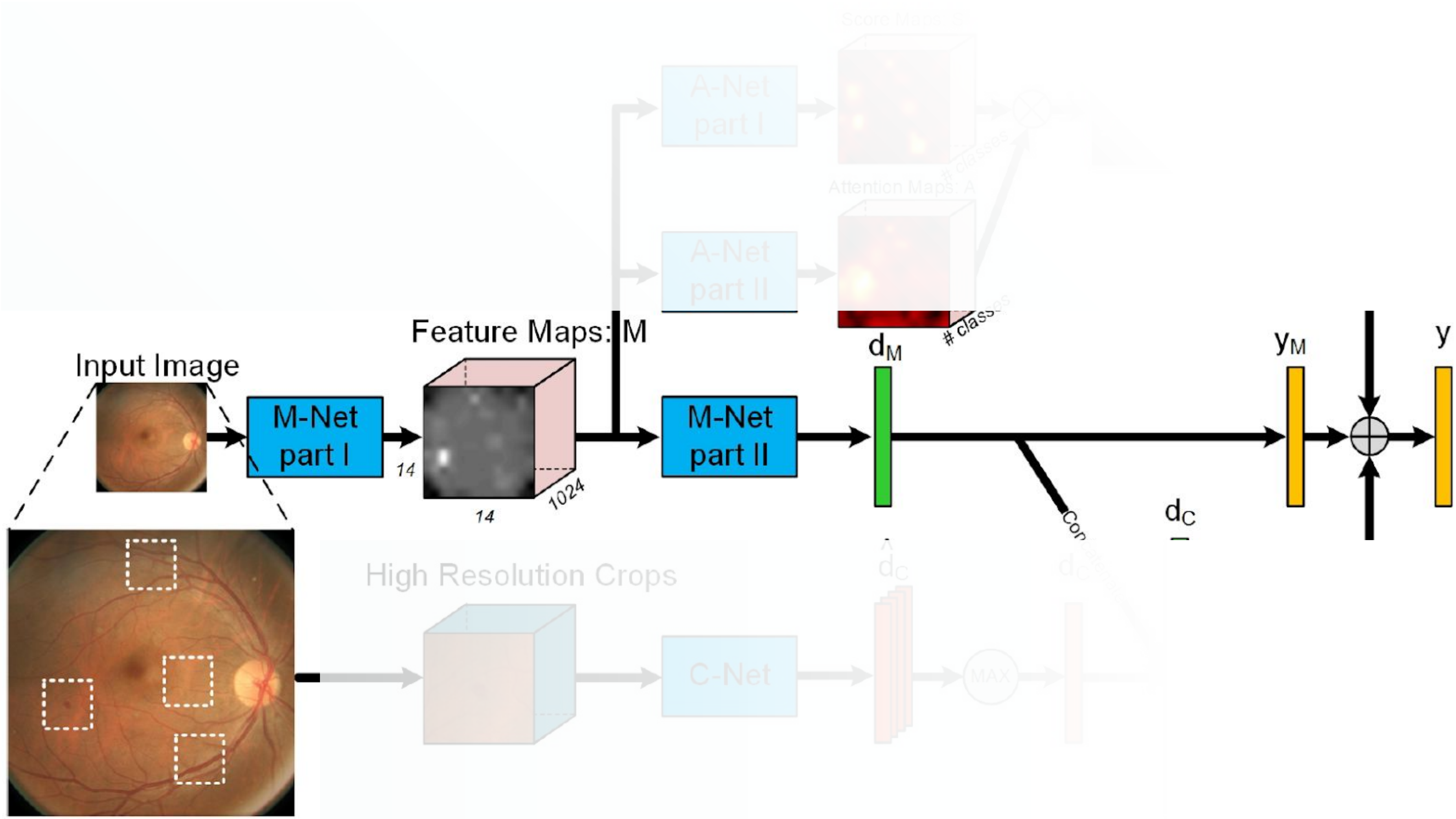
Source: T.Y. Wong, C.M. Cheung, M. Larsen, S. Sharma, and R. Simó. *Diabetic retinopathy*, 16012. Nature Reviews Disease Primers, 2016

Architecture (Zoom-in Net)



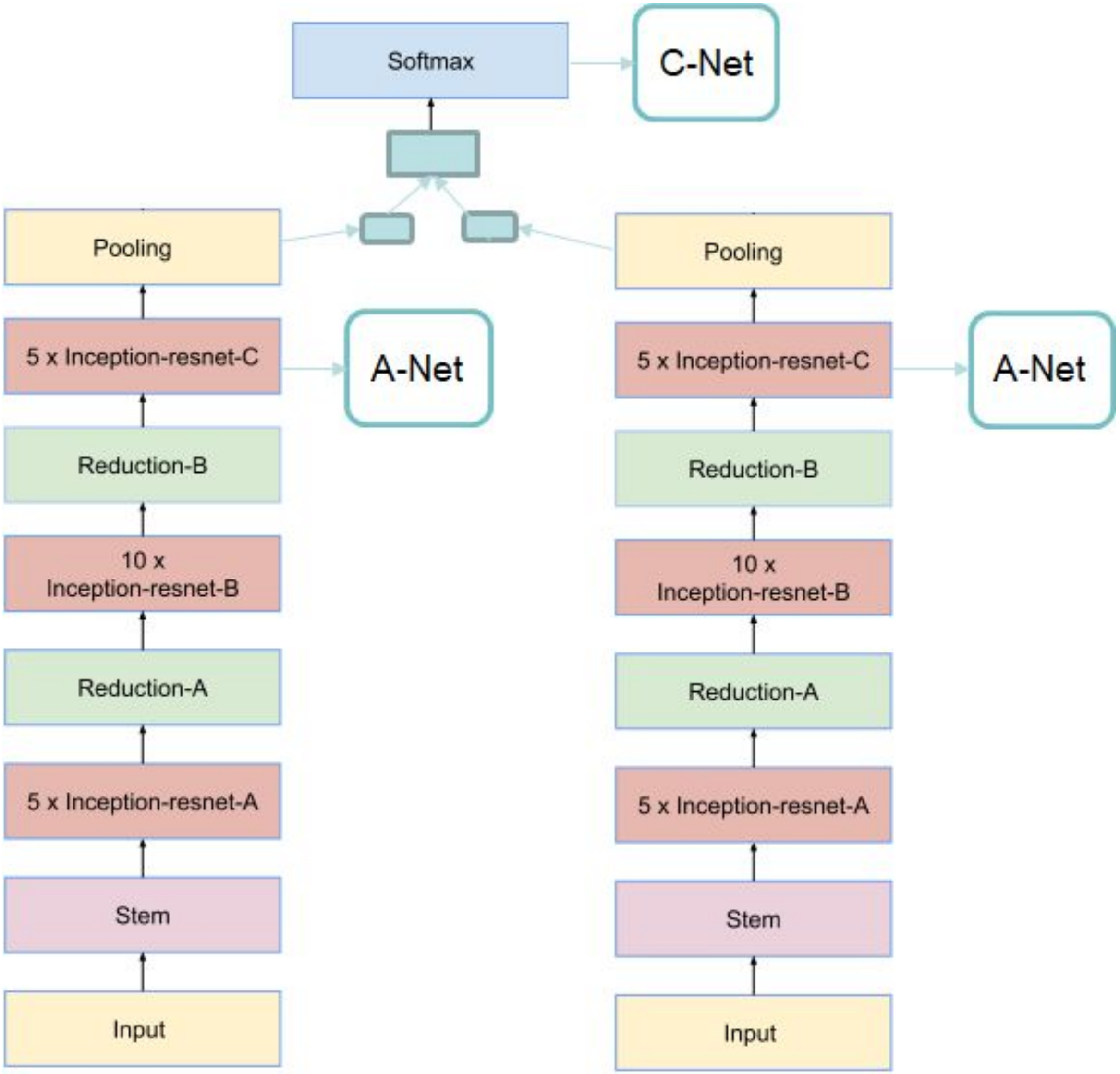
Source: Z.Wang, Y.Yin, J.Shi, W.Fang, H.Li, and X.Wang. *Zoom-in-Net: Deep Mining Lesions for Diabetic Retinopathy Detection*, pages 267–275. Springer International Publishing, 2017

Main Network (M-Net)



Source: Z.Wang, Y.Yin, J.Shi, W.Fang, H.Li, and X.Wang. *Zoom-in-Net: Deep Mining Lesions for Diabetic Retinopathy Detection*, pages 267–275. Springer International Publishing, 2017

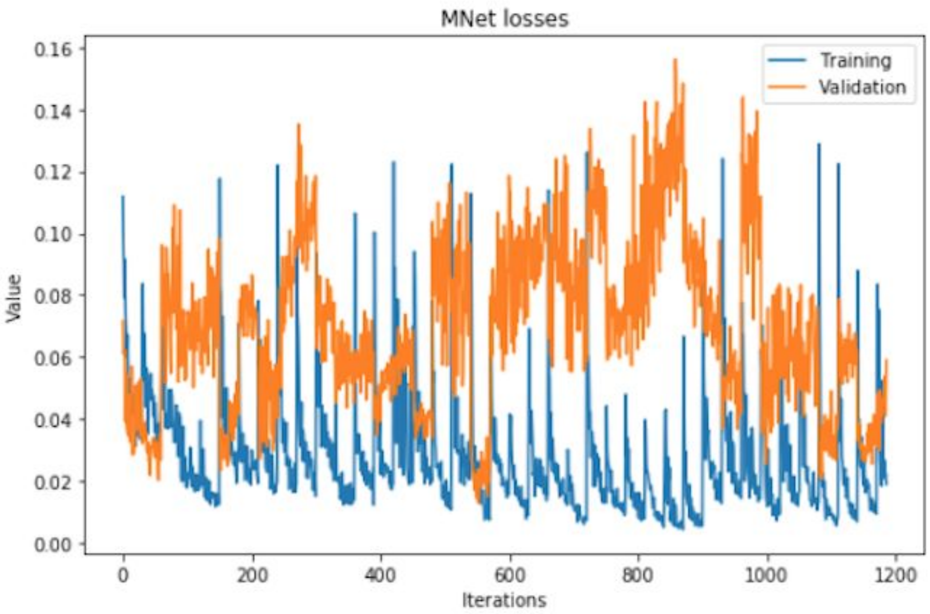
Main Network (M-Net)

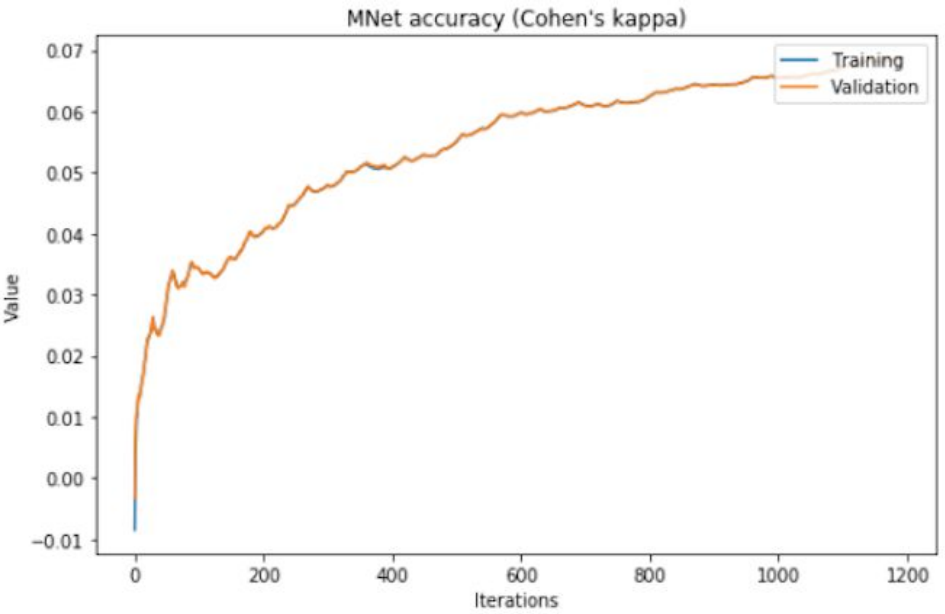


- **Input:** Pre-processed images of size 512x512
- **Output:** Probability of the image belonging to each disease level

- Output from layer *5xIncepetion-resnet-C* is taken as input for A-Net.
- Concatenate the left and right eye for classification.
- Weighted loss computation
- **Challenge:** getting the concatenation right

Mnet training. Original approach





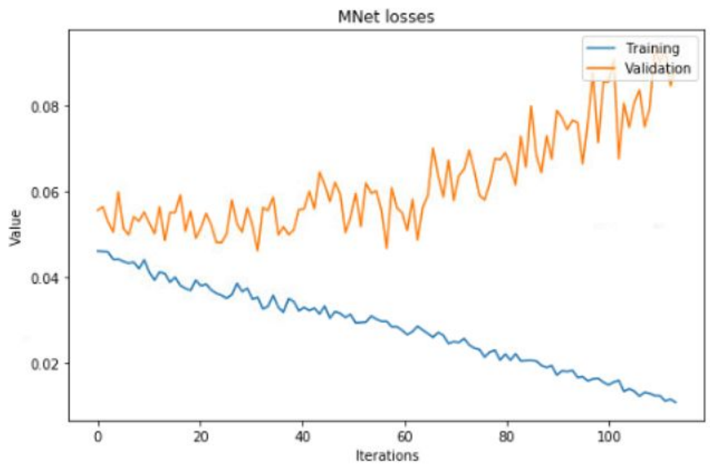
Advantages

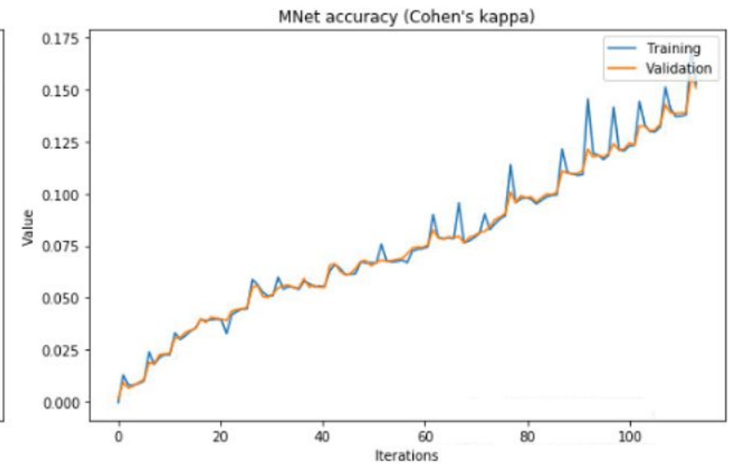
1. Quick
2. Computational resource friendly
3. Easily portable

Disadvantages

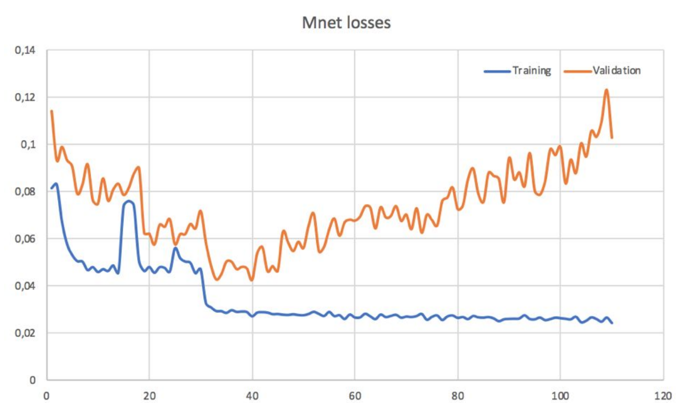
1. No understanding of what happening

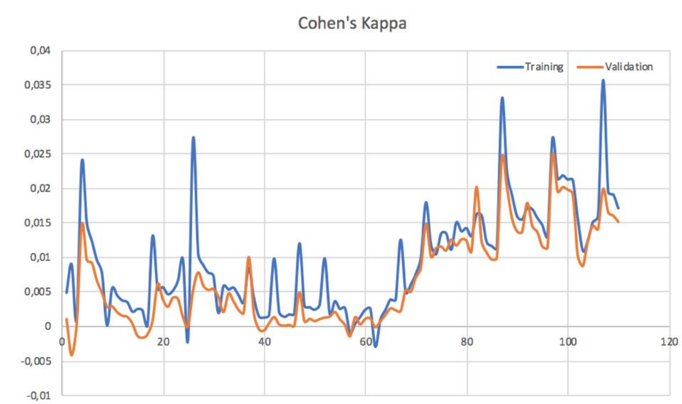
Mnet training. Subsetting



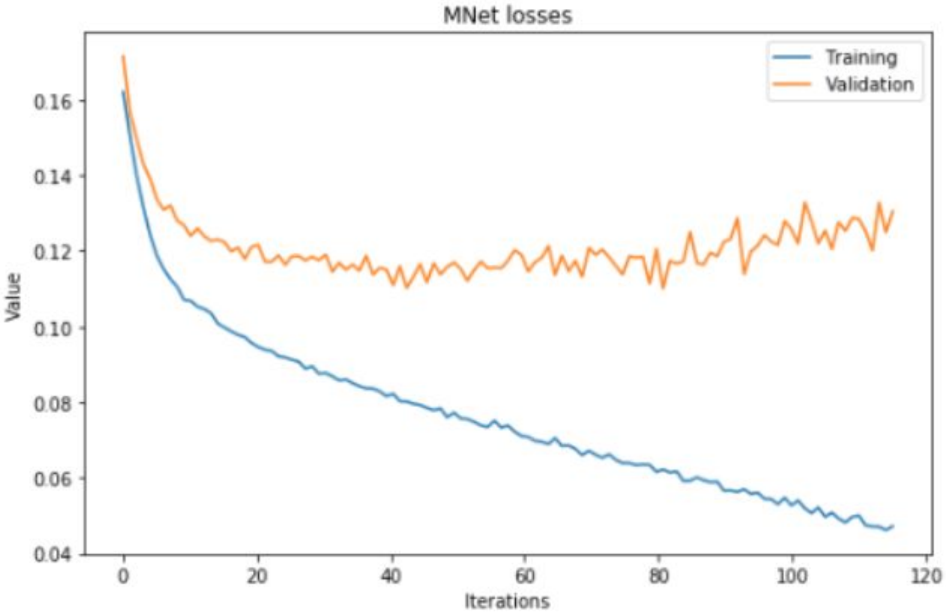


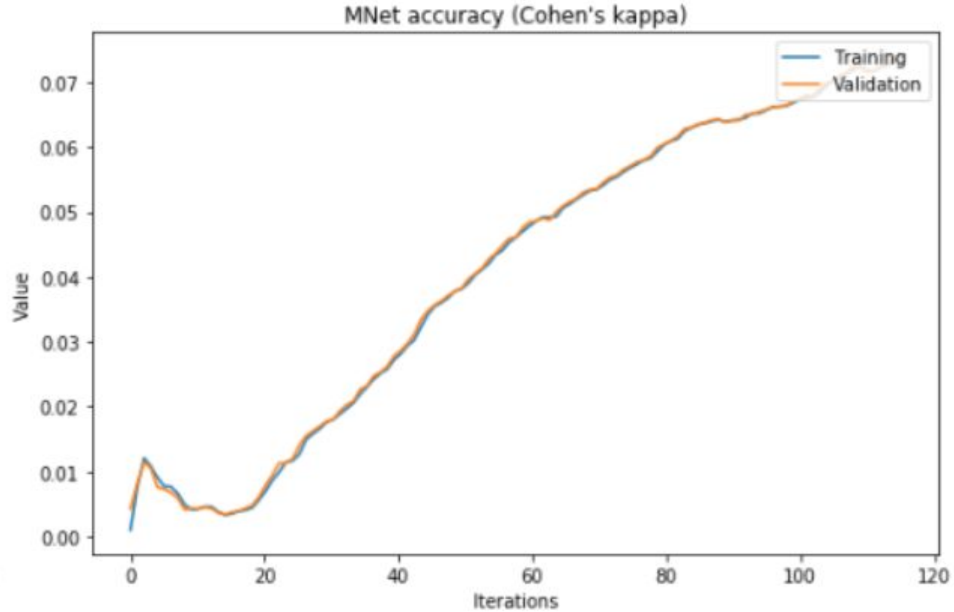
1. L2 regularization
2. Weighted Kappa Score





Mnet. 1 eye

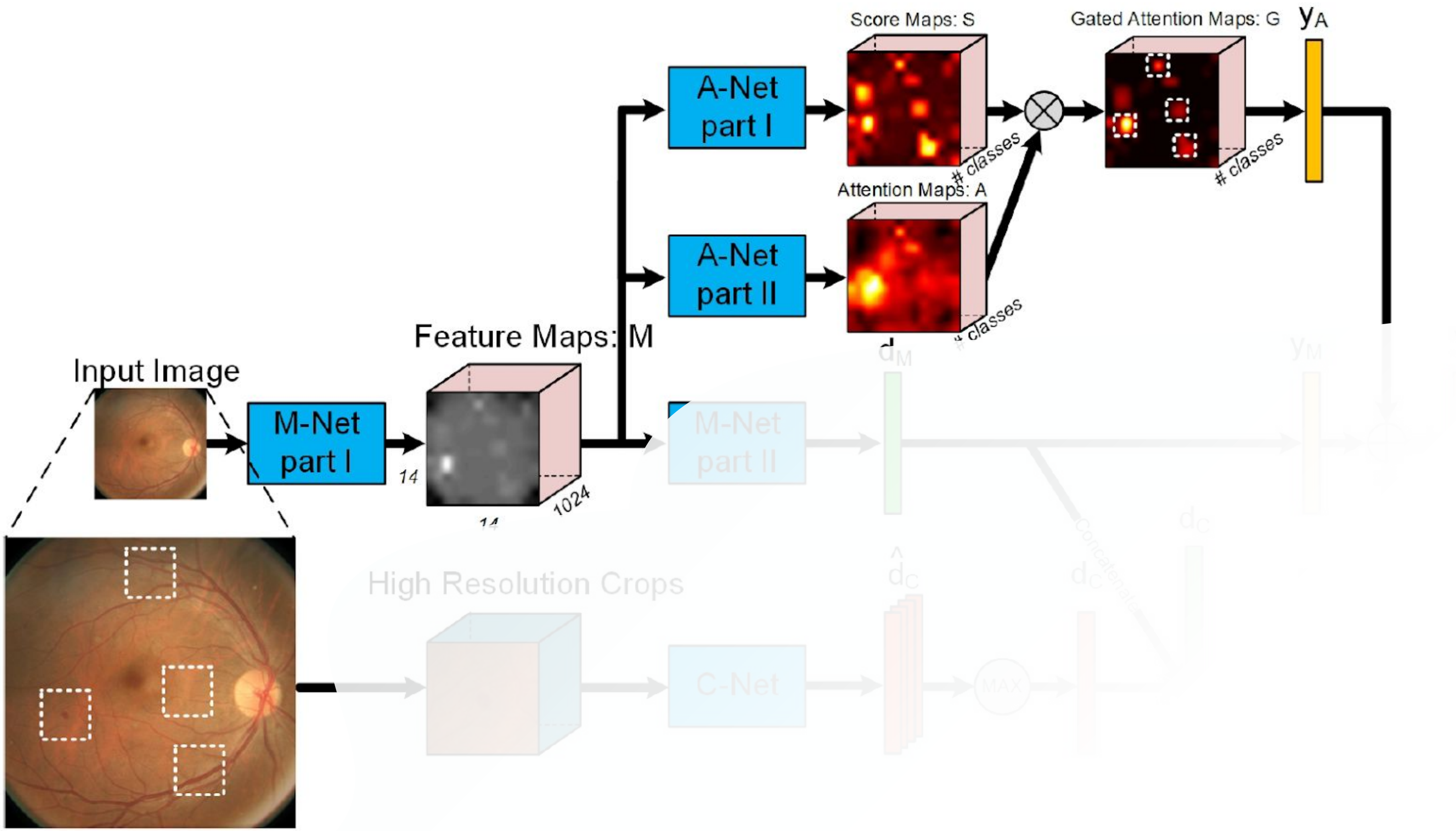




Further plans:

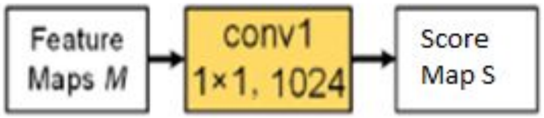
1. Train on more images
2. Double check Cohen's kappa score

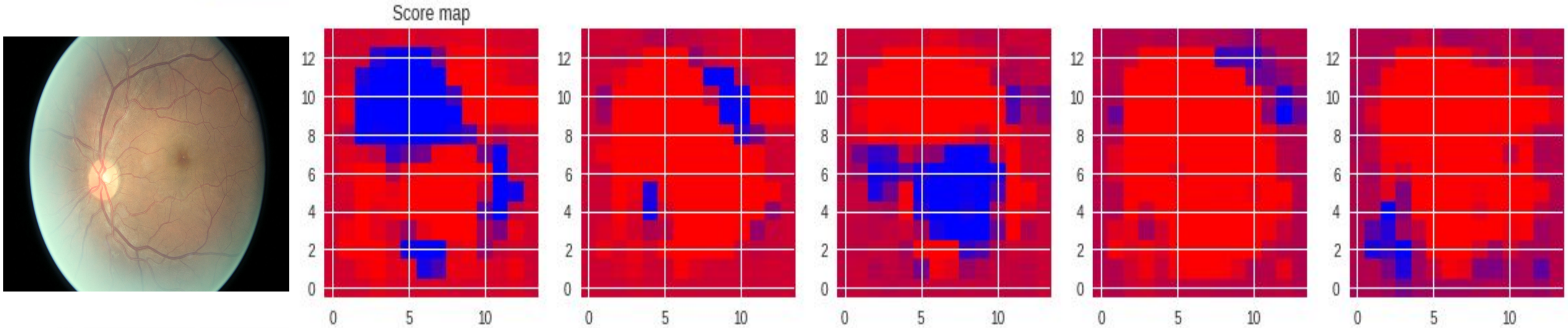
Attention Network (A-Net)

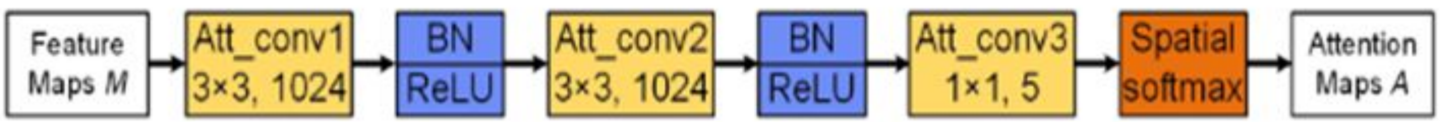


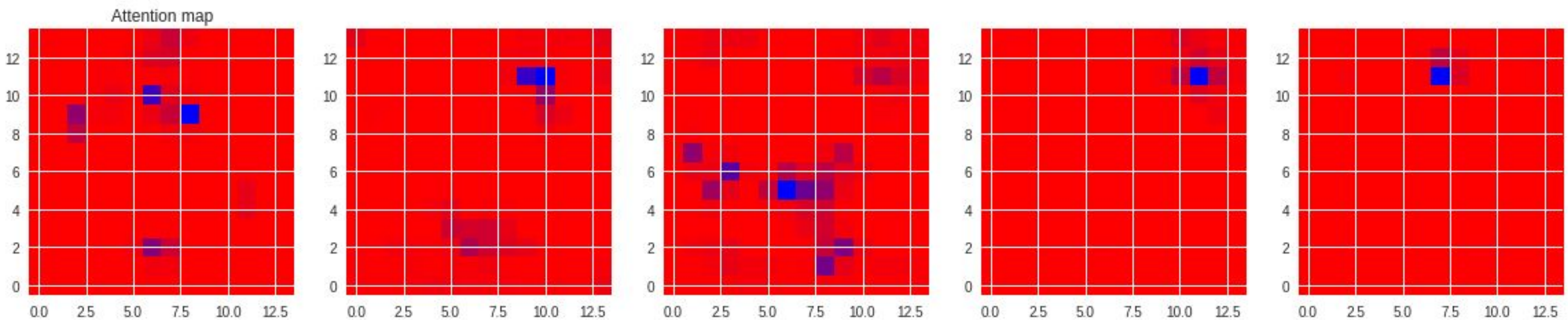
Source: Z.Wang, Y.Yin, J.Shi, W.Fang, H.Li, and X.Wang. *Zoom-in-Net: Deep Mining Lesions for Diabetic Retinopathy Detection*, pages 267–275. Springer International Publishing, 2017

Attention Network (A-Net)

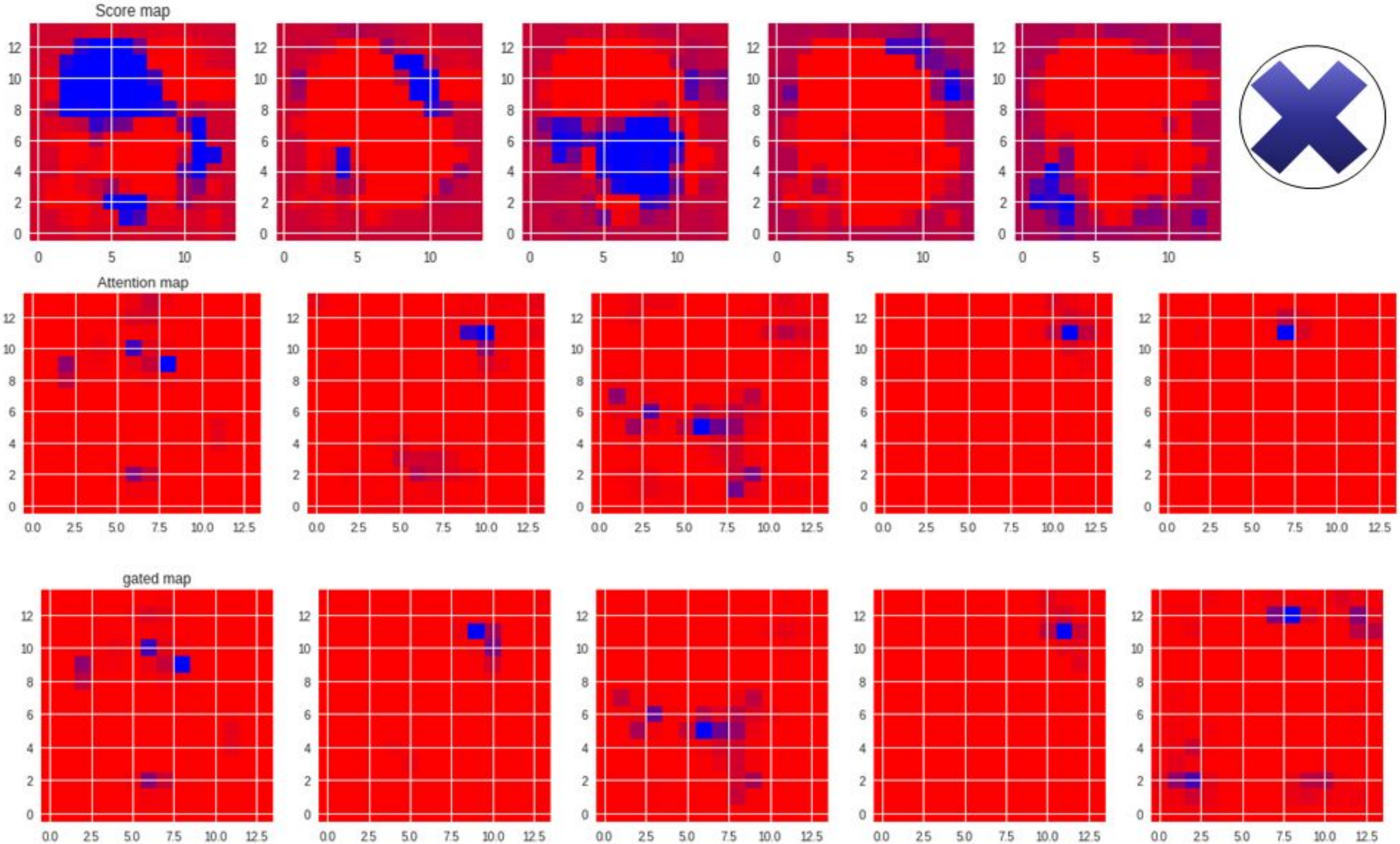




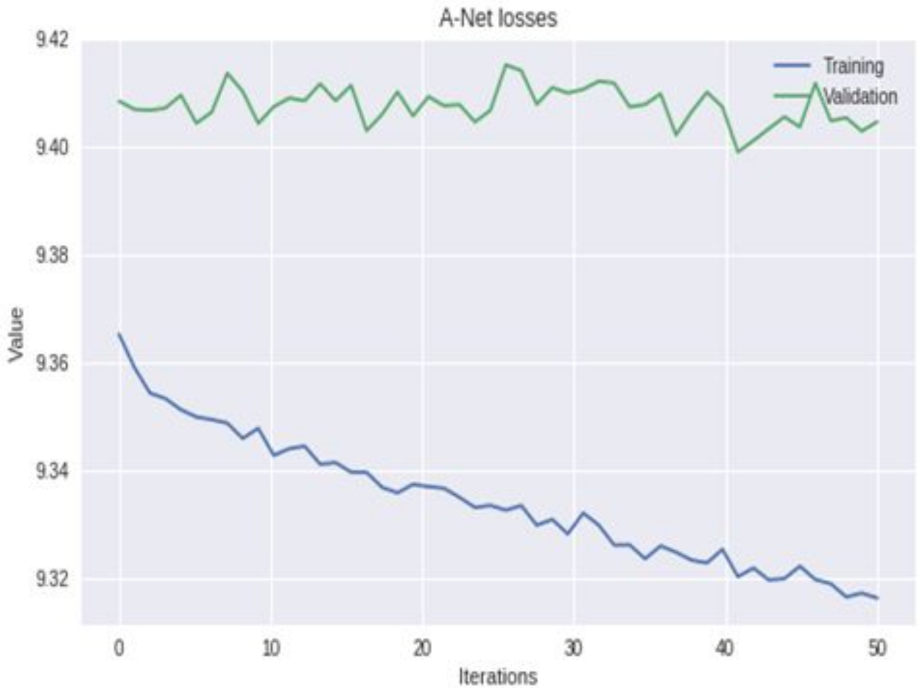


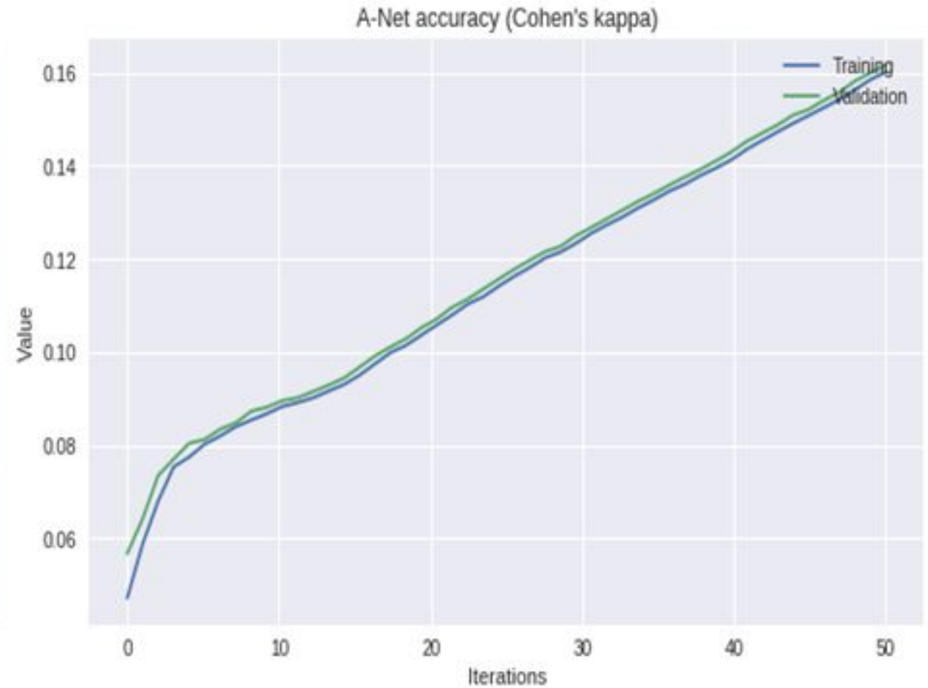


Gated map



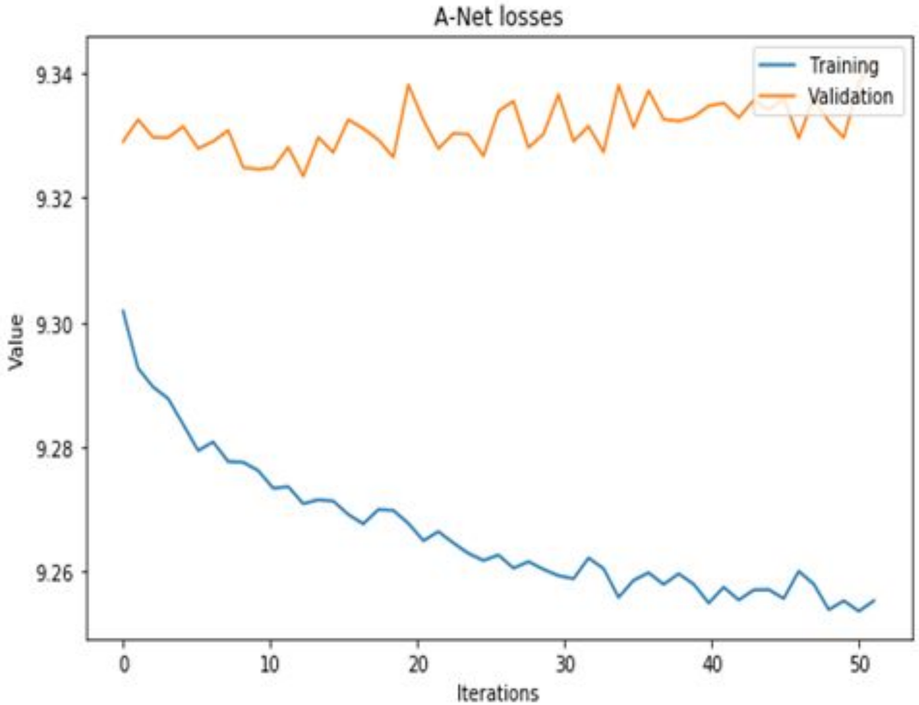
A-Net with 2-eye M-Net

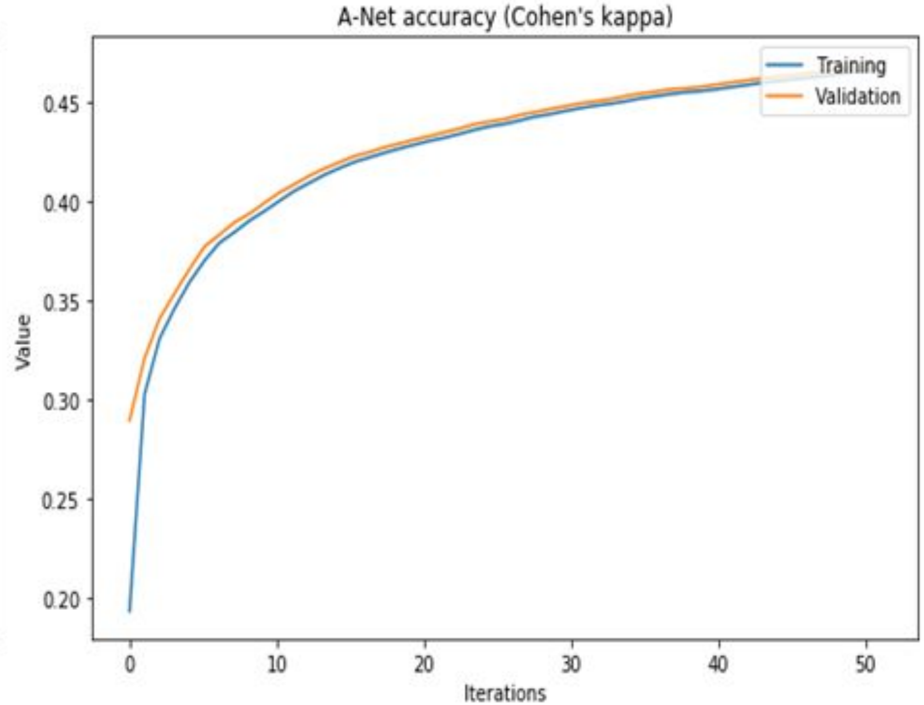




Weighted SoftMax Cross Entropy Loss (with L2 regularization) : 9.30
Kappa Score :0.16

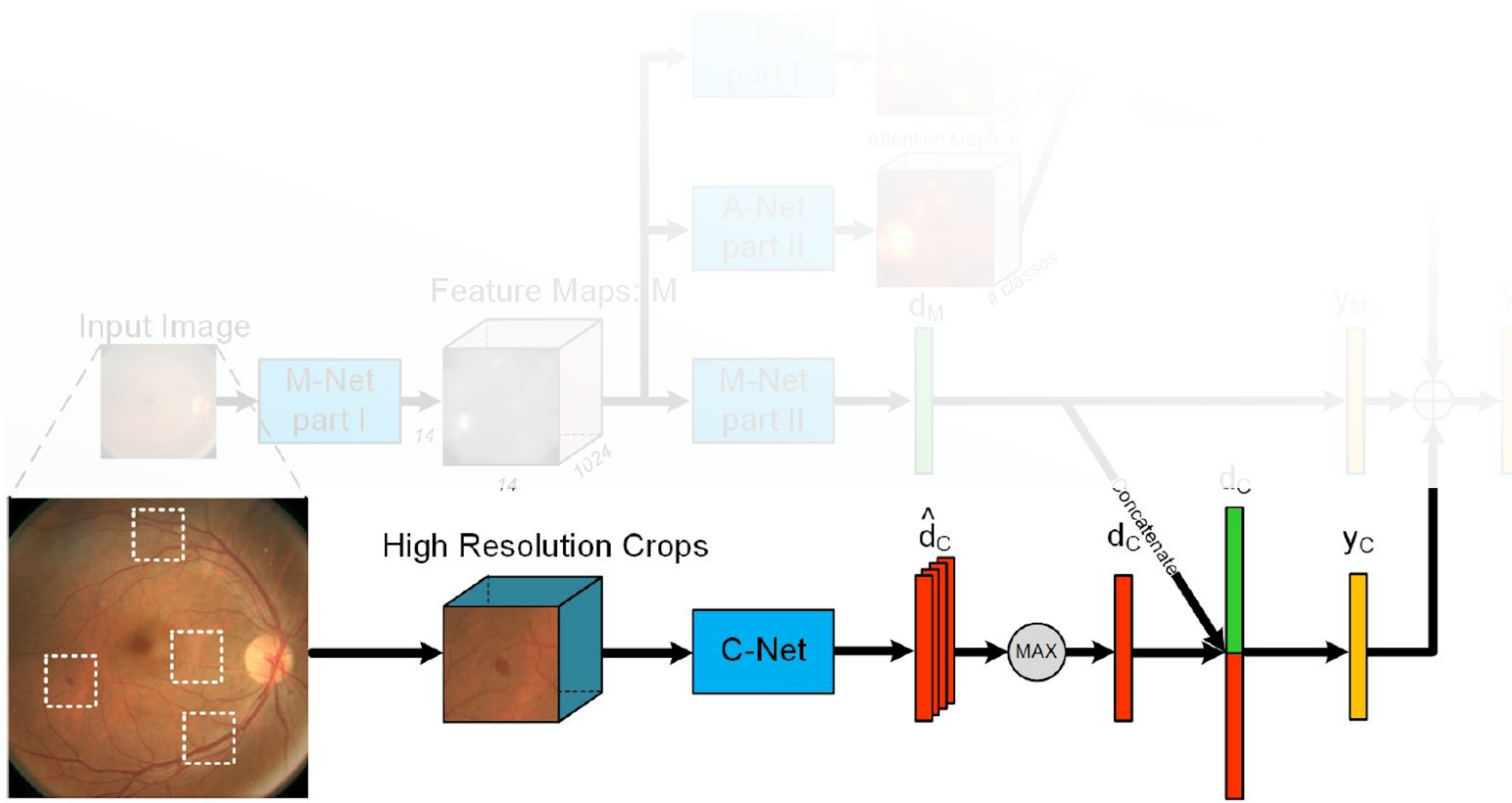
A-Net with 1-eye M-Net Model





Weighted SoftMax Cross Entropy Loss (with L2 regularization) : 9.25
Kappa Score :0.48

Crop Network (C-Net)



Source: Z.Wang, Y.Yin, J.Shi, W.Fang, H.Li, and X.Wang. *Zoom-in-Net: Deep Mining Lesions for Diabetic Retinopathy Detection*, pages 267–275. Springer International Publishing, 2017

Recap: Initializing Inception

Approach:

- Retrain all layers
- Using TF-Slim implementation
- Create TF-Records of all images then train
- Data set : IDRiD

Result:

- Trained the network up to 20k steps
- We have an initialized model for training with accuracy of 92% and loss of 0,70

Integration

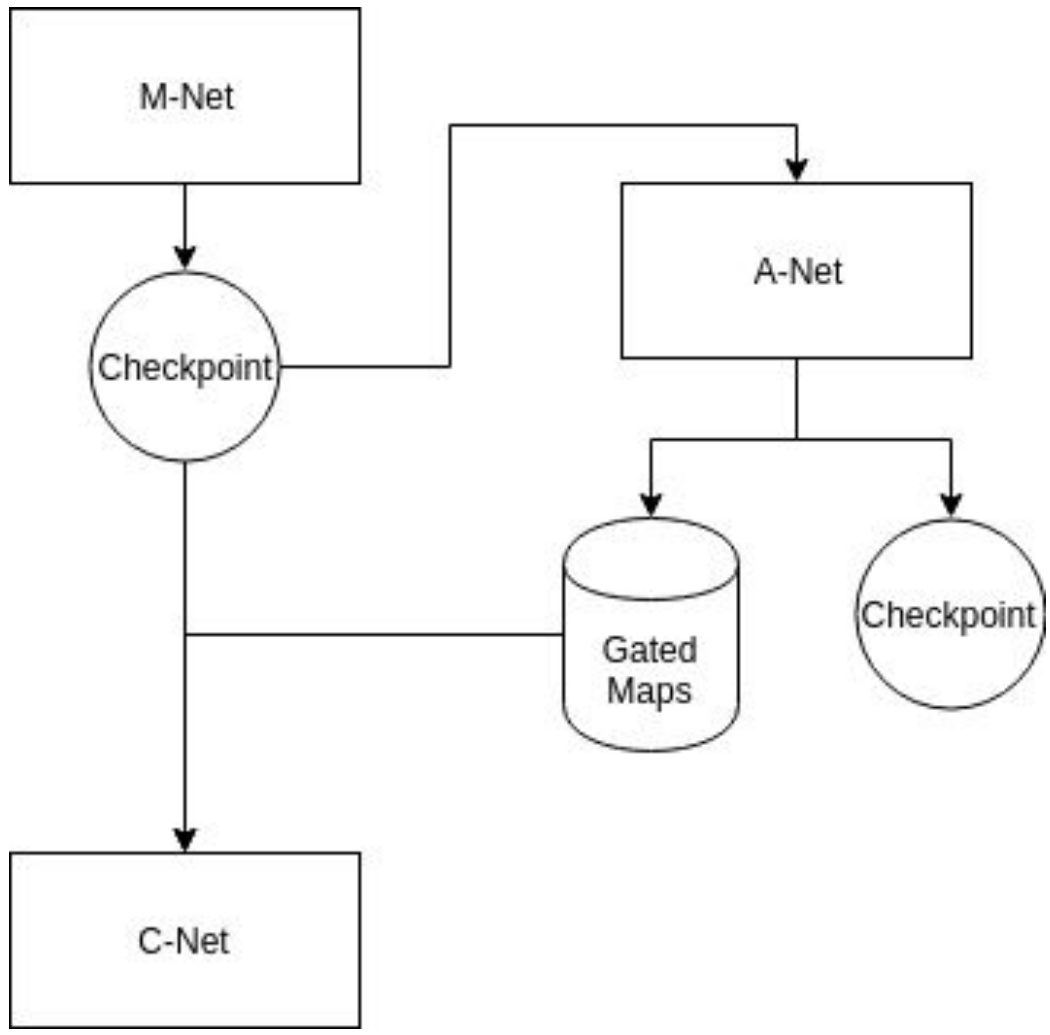
Planned:

- Initially planned on a single training pipeline for all 5 Neural Nets
- Unfeasible due to highly coupled nature of project

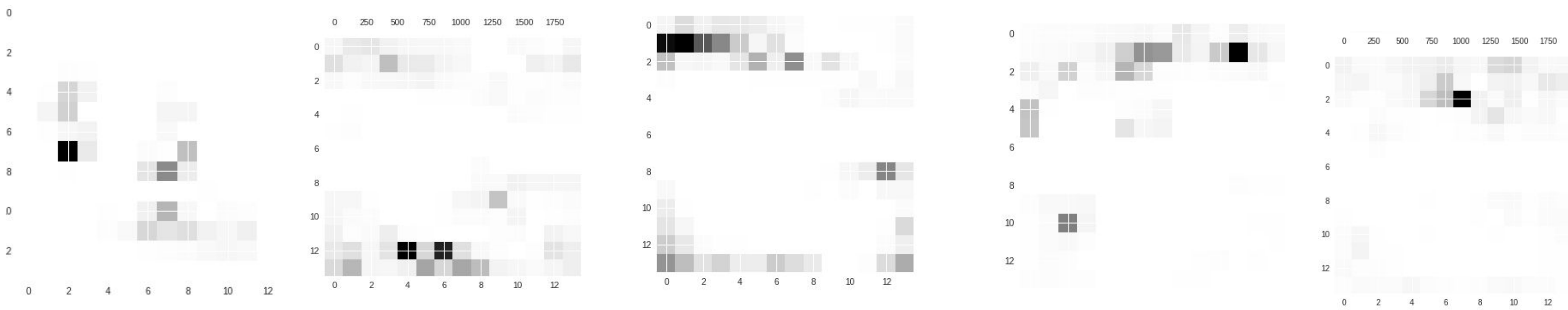
Implementation:

- Write each component independently
- Write output of each component into a binary file
- Saved checkpoints of each component used to get feature maps

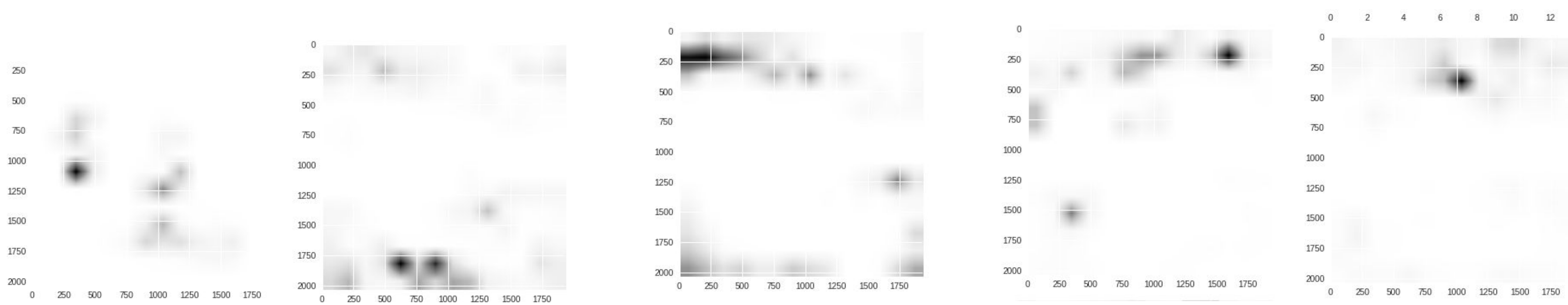
Integration



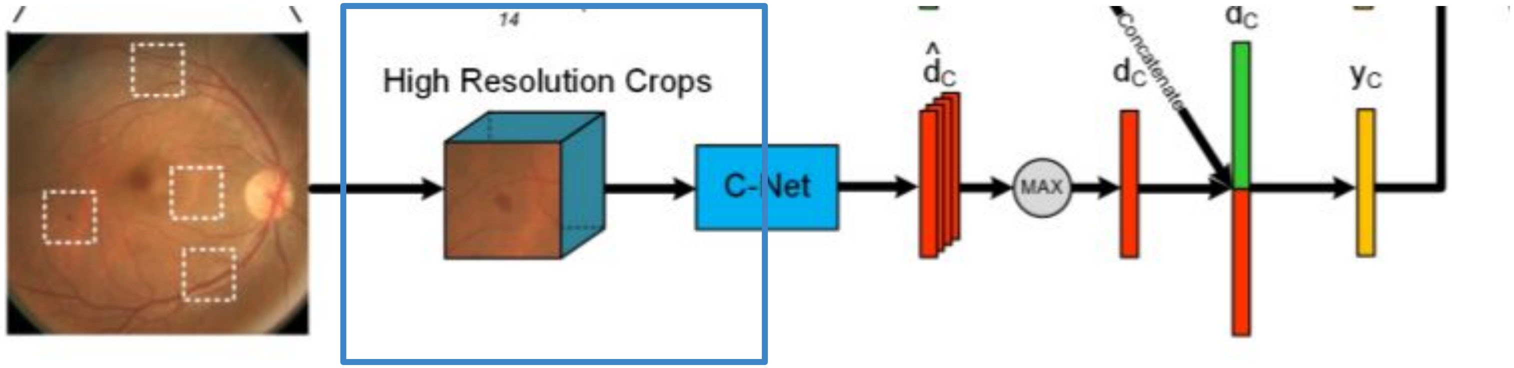
A-net Attention Maps



Resize to size of original Images



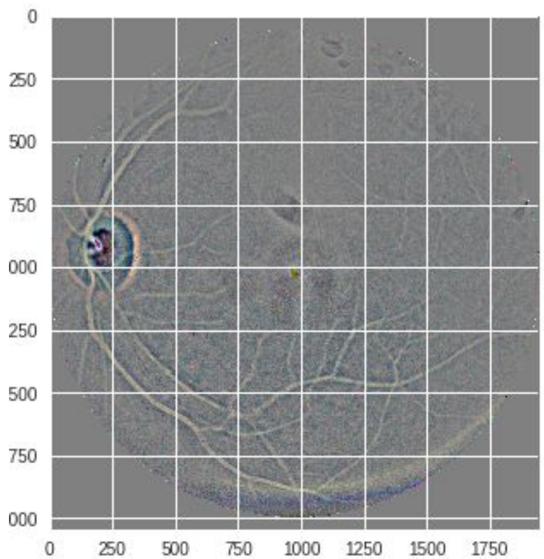
Crop Network (C-Net)

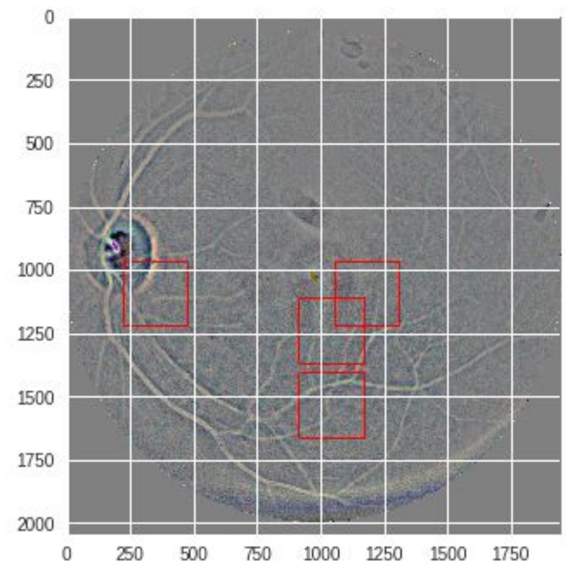


Process

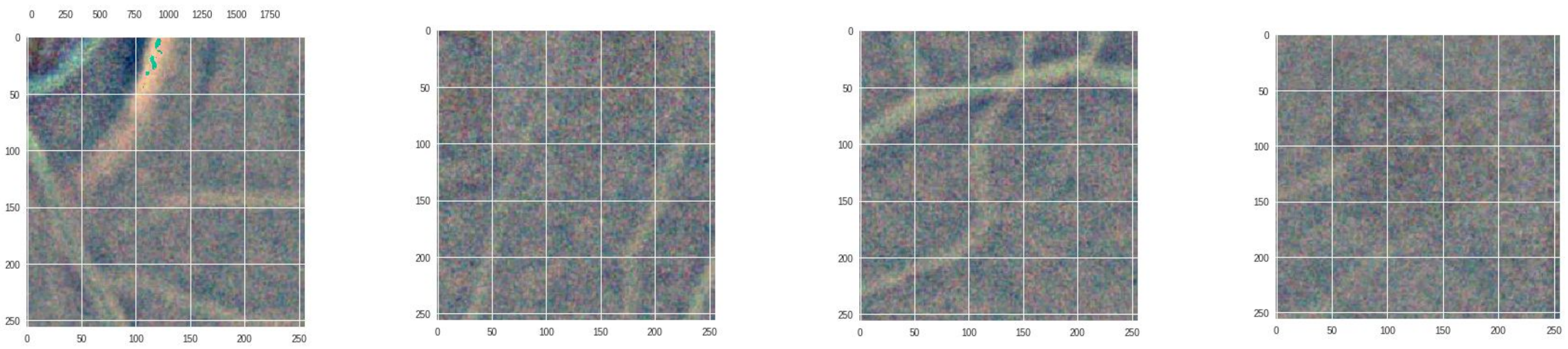
- **Zooming-in** suspicious attention regions
- Record the highest response in the attention map and crop the region
- Mask the region around the crop to avoid re-selection
- Repeat until total of N-coordinates are recorded

Crop Detection

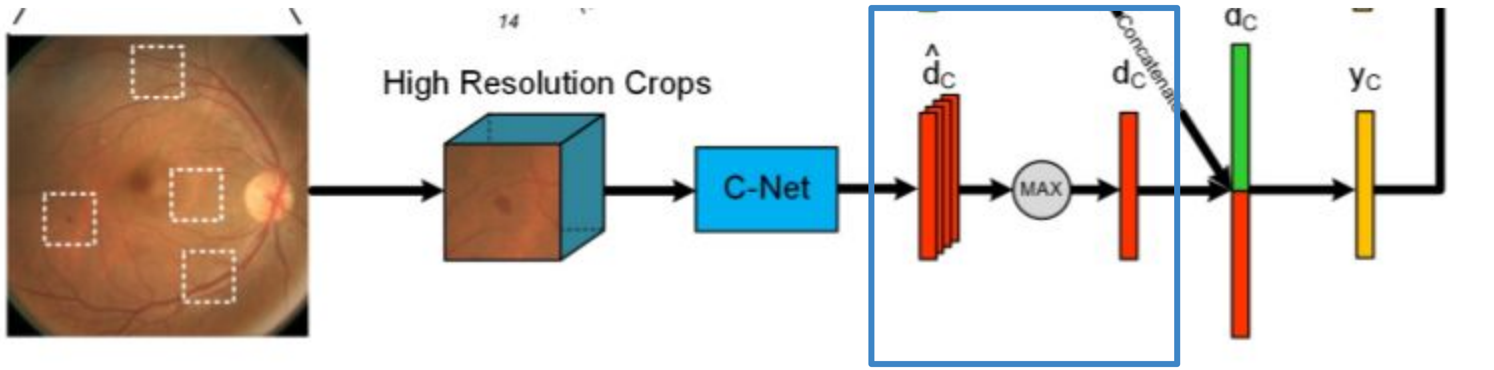




Crops


0 250 500 750 1000 1250 1500 1750
0 50 100 150 200 250
0 50 100 150 200 250
0 50 100 150 200 250
0 50 100 150 200 250

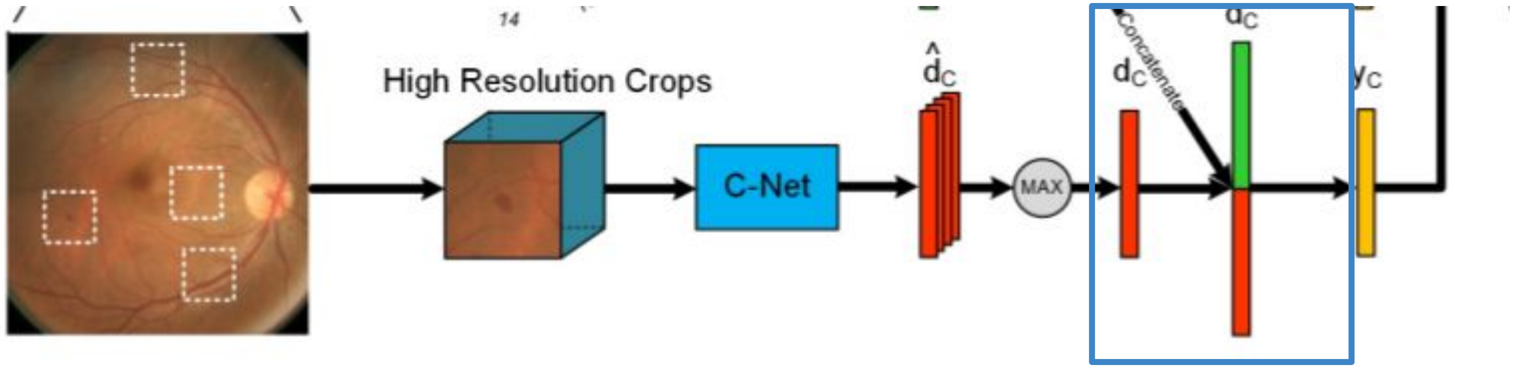
Crop Network (C-Net)



Output

- For all the crops of the same eye, take an element wise max

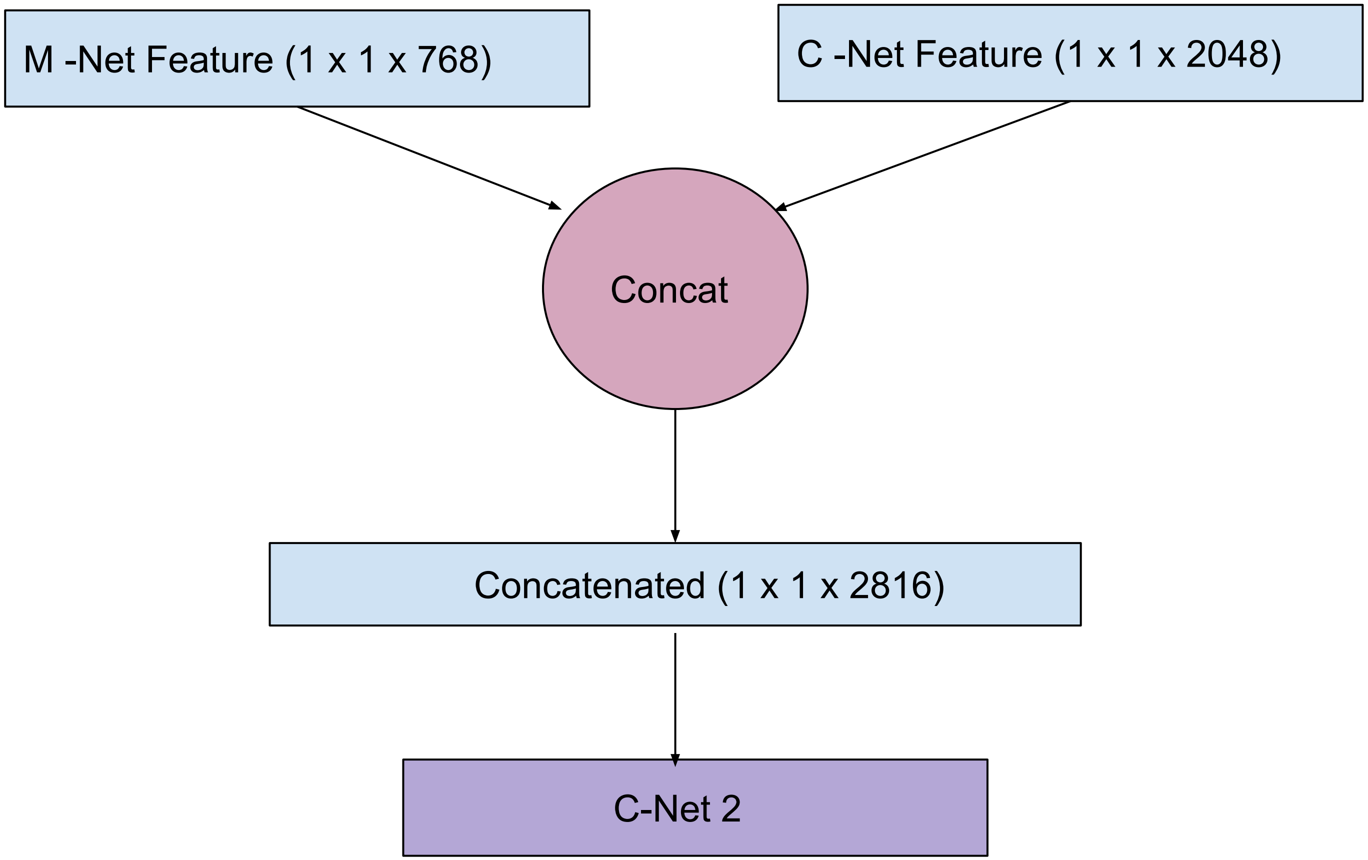
Crop Network (C-Net)



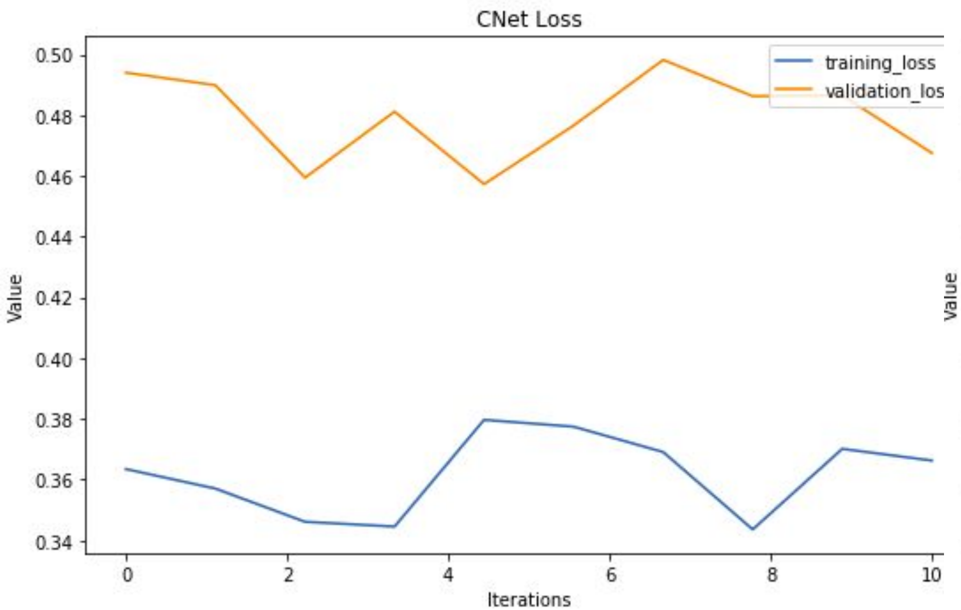
Concatenation:

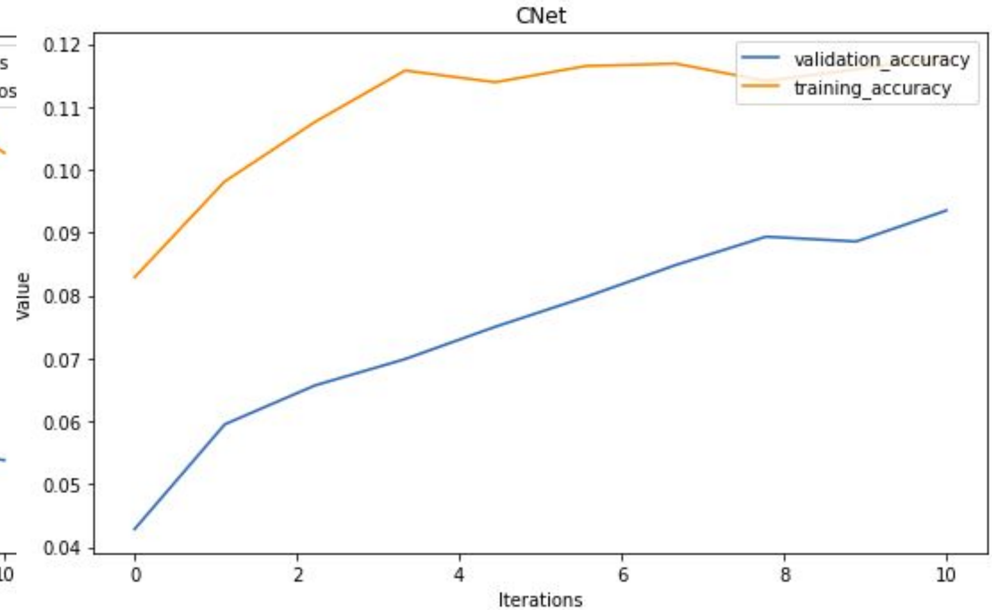
- Concatenate feature vector from C-Net with feature vector from M-Net
- From C-Net : “Global Pool” layer
- From M-Net: “Conv2d_2a_5x5” layer

Crop Network (C-Net)



C-Net Results





Training Loss: 0.372
Training Cohen's Kappa: 0.118
Checkpoint Used : Pre-trained from IDRiD Dataset

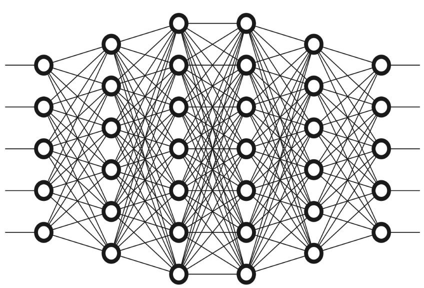
Summary

Network	Training Loss	Training Kappa	Validation Loss	Validation Kappa
M-Net 1 Eye	0.05	0.06	0.13	0.06
M-Net 2 Eye	0.03	0.07	0.05	0.07
A-Net	9.25	0.48	9.26	0.47
C-Net	0.372	0.118	0.47	0.08

Impediments

- Limitation of resources in proportion to size of the problem
  - 5 Networks
  - 85 GB
- Google Colab with multiple experiments, time out issues
- Strong dependency between groups, and models

Learning Outcomes



Questions?

Topics:

Problem Statement
Proposed Approach
Recap
MNet
ANet
CNet
Integration
Learning Outcomes
Impediments

Thank You

Technische Universität München

Johns Hopkins Whiting School of Engineering