SEA-Net: Squeeze-and-Excitation Attention Net for Diabetic Retinopathy Grading

Ziyuan Zhao[1], Kartik Chopra[2], Zeng Zeng[1*], and Xiaoli Li[1]

[1] Institute for Infocomm Research (I2R), A*STAR, Singapore
[2] National University of Singapore, Singapore

Institute for Infocomm Research
I²R

NUS
National University of Singapore

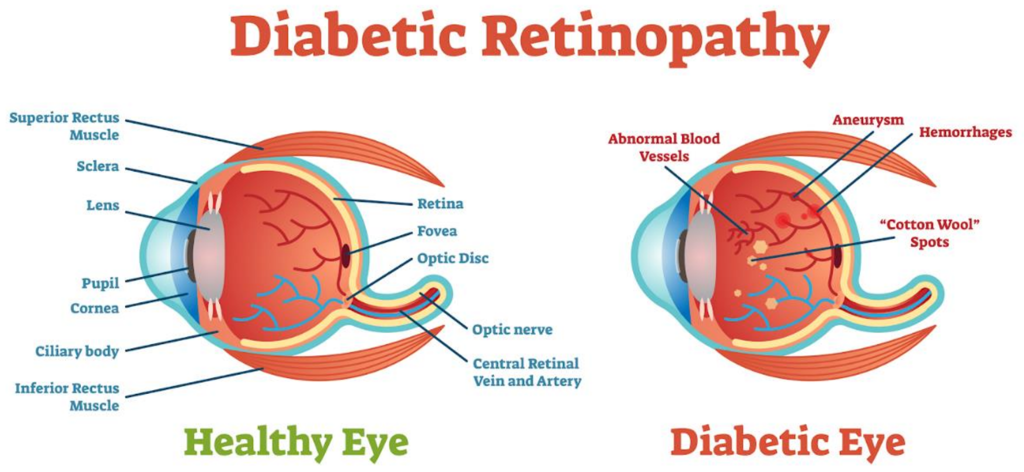
ARS-02: Biomedical and Biological Image Analysis #2630

Background

- Diabetic retinopathy (DR) is a common retinal disease that leads to blindness.

- In Singapore, around 1 out of 12 people aged from 19 to 69 years are affected by diabetes, and 43.5% among them suffer from different severity of DR *.

- It Augments the blood pressure in small vessels and influence the circulatory



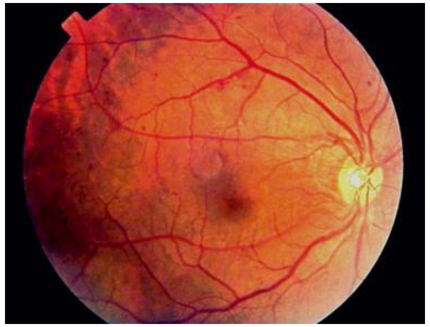
[Posted on: July 12, 2019 Home Health Care]

* https://www.singhealth.com.sg/news/medical-news-singhealth/updates-in-detection-and-treatment-of-diabetic-retinopathy

Challenges

Manual Inspection

- There are no early warning symptoms for DR.
- Difficulties in timely diagnosis and early treatment.
- DR grading also suffers from high intra- and inter-observer variability.



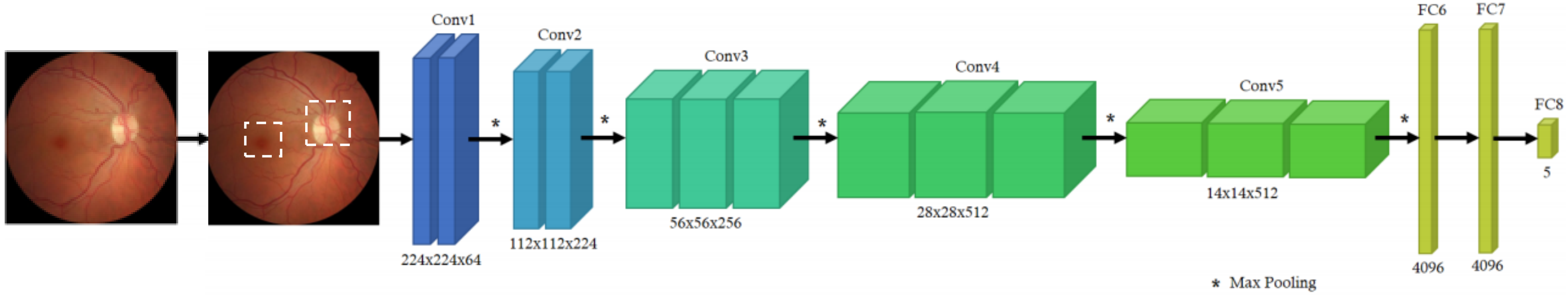
Challenges

Automatic Method

- Extracted features from photos are hand-crafted features.

- Feature localization and segmentation can not be well embedded into the whole DR detection framework.

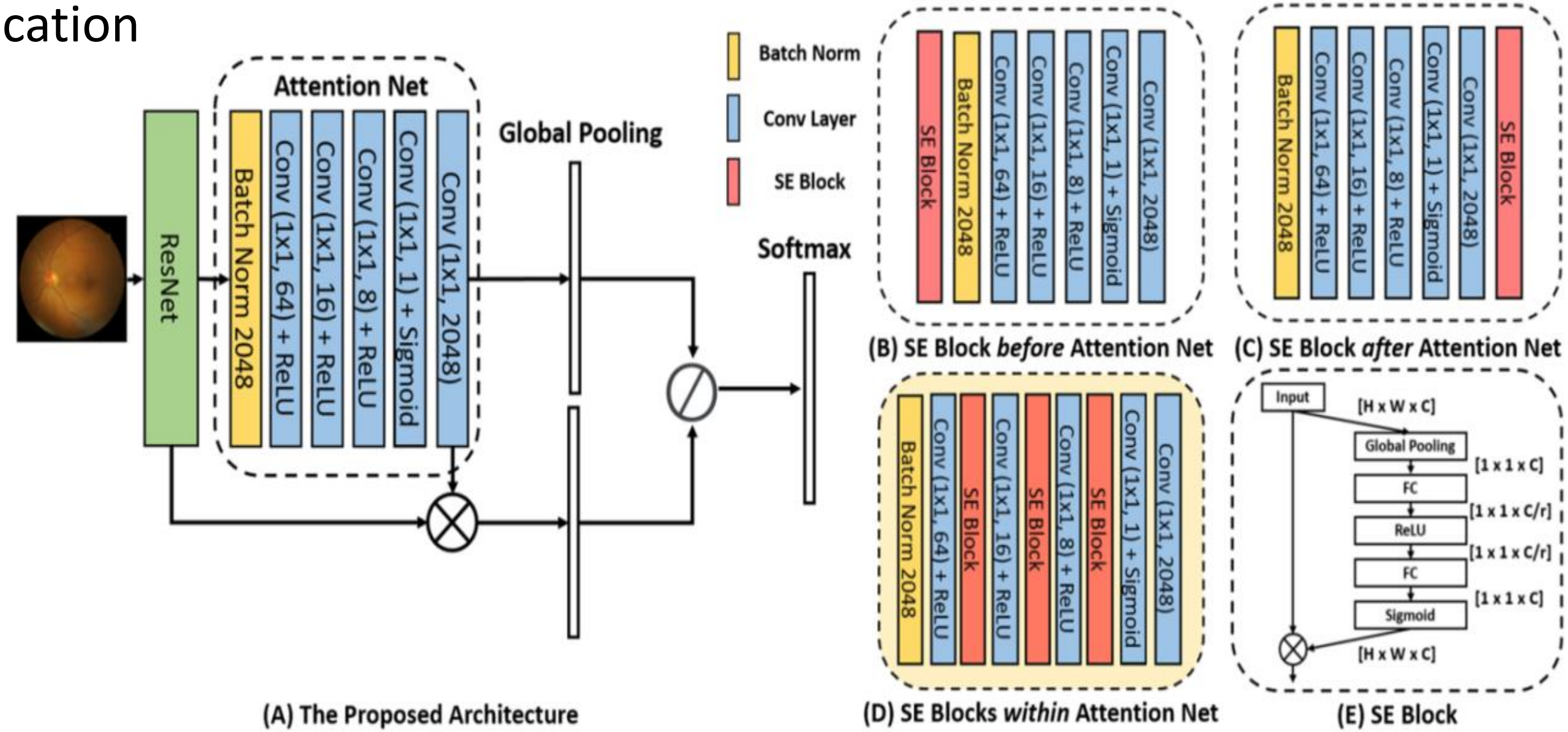
- Most about Binary classification (DR / no DR)



SEA-Net: Squeeze-and-Excitation Attention Net

- **Attention Net** is extended from BiRA-Net* for spatial attention

- **SE blocks** are introduced to recalibrate channel-wise feature maps for fine-grained classification



[* https://ieeexplore.ieee.org/document/8803074](https://ieeexplore.ieee.org/document/8803074)

Attention Net

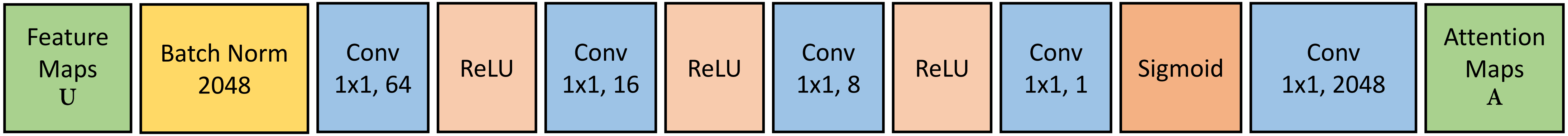
- **ResNet-50** is implemented first for deep feature extraction (I → U)

- Through a sequence of 1 × 1 convolution layers and pooling layers, the refined feature map is obtained (U → A)

- The global average pooling (GAP) layer provides a receptive field of whole spatial extent

- An element-wise division is used followed by a softmax layer

- Output: $Output = GAP\left(A^{l}\right) \oslash GAP\left(A^{l} \otimes \mathrm{U}^{l}\right)$

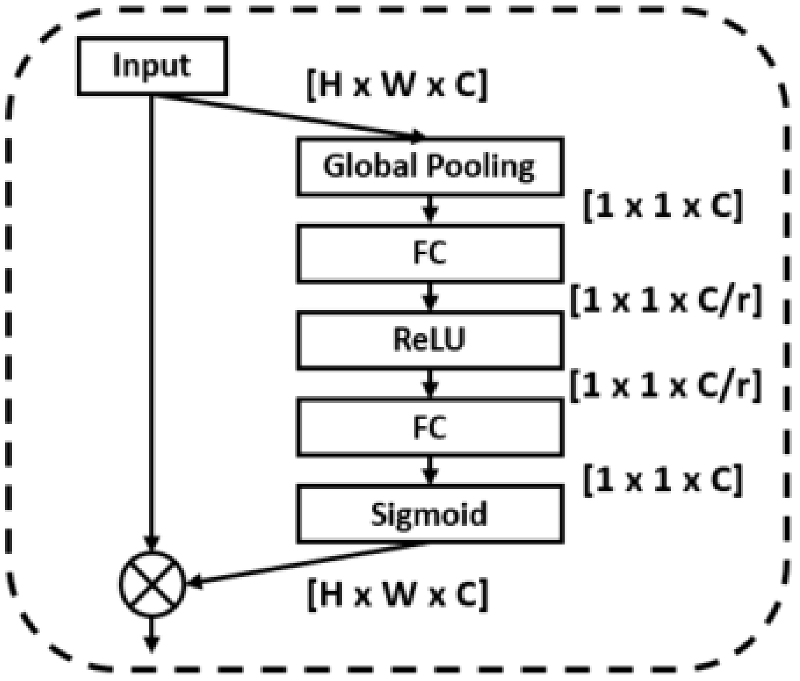


A^l and U^l are l-th attention map and l-th feature map.
$\otimes$ and $\oslash$ denote element-wise multiplication and element-wise division.

Squeeze-and-Excitation Block

- SE Block is borrowed from **Squeeze-and-Excitation Networks***

- To exploit channel dependencies and contextual information, we propose to incorporate the SE block into the proposed architecture



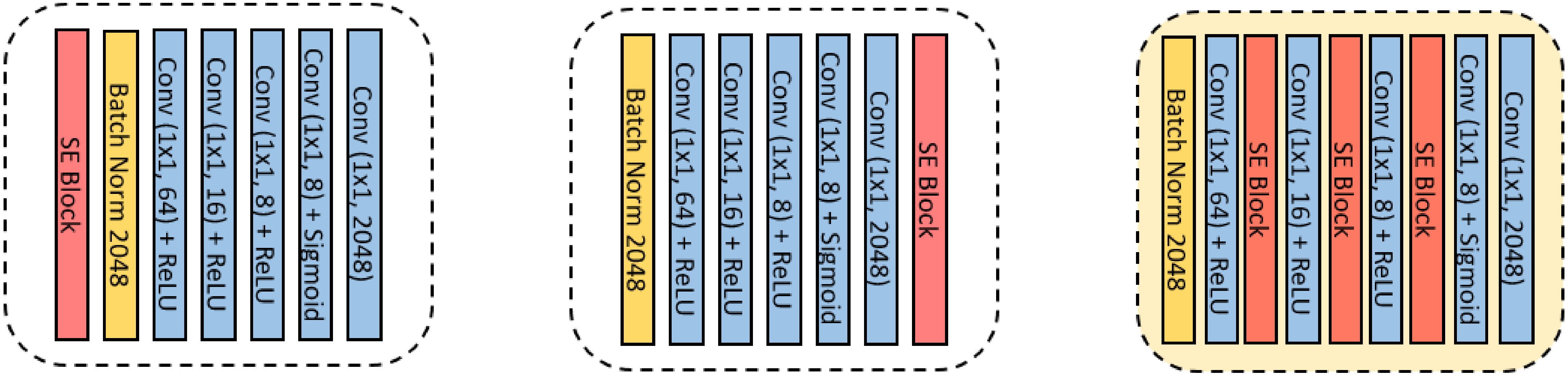
SE Block

[* https://ieeexplore.ieee.org/document/8578843](https://ieeexplore.ieee.org/document/8578843)

Squeeze-and-Excitation Block

- The positions of SE blocks in the network influence the performance of DR grading.

- To find the optimal position, we explore three different positions of SE blocks.



(a) SE Block before Attention Net
SE-AT-Net

(b) SE Block after Attention Net
AT-SE-Net

(c) SE Block with Attention Net
SEA-Net

Hybrid Loss Function

- Implement center loss to reduce the loss-accuracy discrepancy and get an improved convergence.

- The weighted cross entropy loss is used to alleviate data imbalance

$$\mathfrak{L}_{ct} = \frac{1}{2}\sum_{i=1}^{m} \|\boldsymbol{x}_i - \boldsymbol{c}_{y_i}\|_2^2 \qquad \mathfrak{L}_{ce} = \text{weight}_y \left(-\log \left(\frac{\exp(x[y])}{\sum_j \exp(x[j])} \right) \right)$$

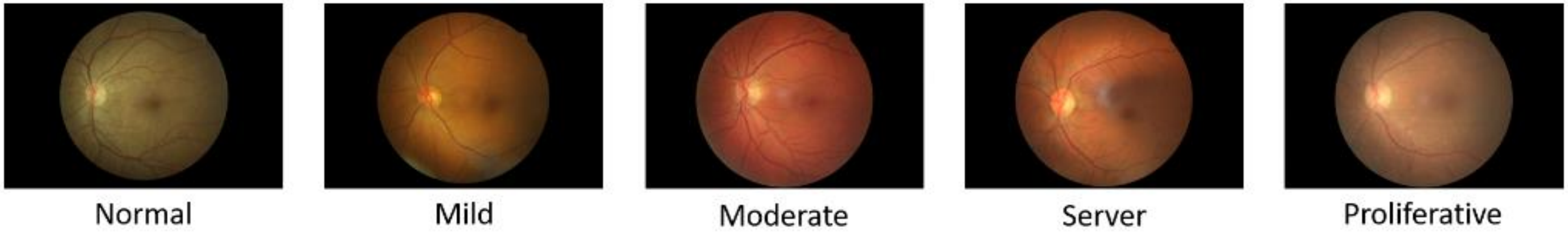
$$\mathfrak{L} = \mathfrak{L}_{ce} + \lambda \mathfrak{L}_{ct}$$

λ is a scalar to control the strength of loss functions.

Dataset

- The retinal images are provided by EyePACS consisting of 35126 images. And each image is labeled as {0, 1, 2, 3, 4}, depending on the disease’s severity.

- following the data distribution adopted by Marıa A. Bravo et al, a balanced testing dataset of 1560 images was applied to our experiments for testing, and the rest were used for training



Marıa A. Bravo et al, “Automatic diabetic retinopathy classification,” in 13th International Conference on Medical Information Processing and Analysis, 2017

Results

- The proposed framework outperforms other methods in all metrics.

- SE blocks are proved to be effective in the proposed methods

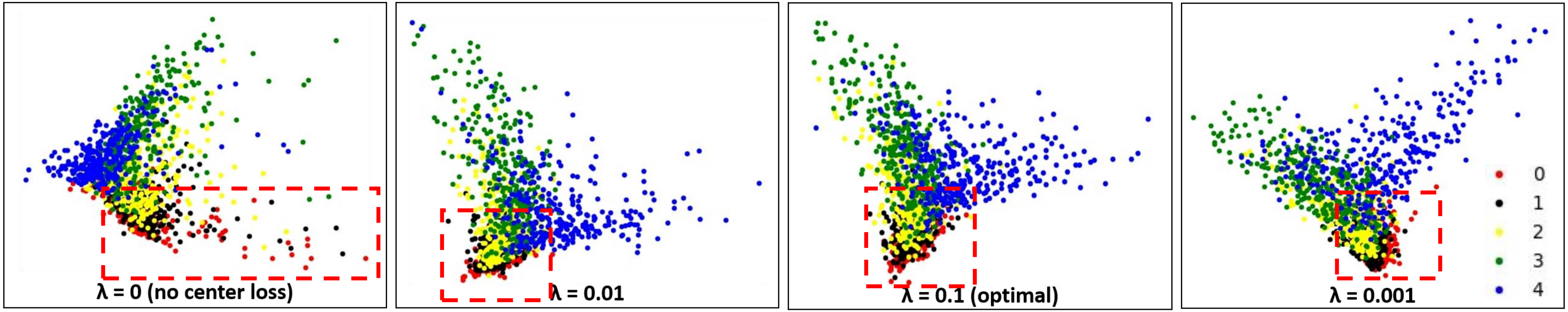
- In SEA-Net, the SE blocks are placed alternatively with convolution layers, recalibrating the learned feature maps in an adaptive manner.

Method	ACA	Marco-F1	AUC
Bravo et al.	50.51	50.81	-
BiRA-Net	54.31	57.25	-
AT-Net	54.42	49.51	86.99
SE-AT-Net	57.76	55.05	87.34
AT-SE-Net	5.83	58.92	87.21
SEA-Net	58.59	58.72	87.38
SEA-Net (λ = 0.1)	**59.94**	**60.47**	**87.6**

Results

- The propose method is further improved with the proposed hybrid loss function

- the proposed hybrid loss can learn better discriminative features, especially for confusing classes, i.e., class 0 and class 1.



Conclusion

- We proposed a novel deep learning architecture for DR grading.
- Spatial attention and channel attention are implemented to boost each other, recalibrating the attention maps adaptively.
- A hybrid loss function based on weighted cross entropy loss and center loss is implemented.
- Experimental results demonstrate the effectiveness of the proposed architecture.

THANKS

Zhao_Ziyuan@i2r.a-star.edu.sg