Digital Retinal Images for Vessel Segmentation

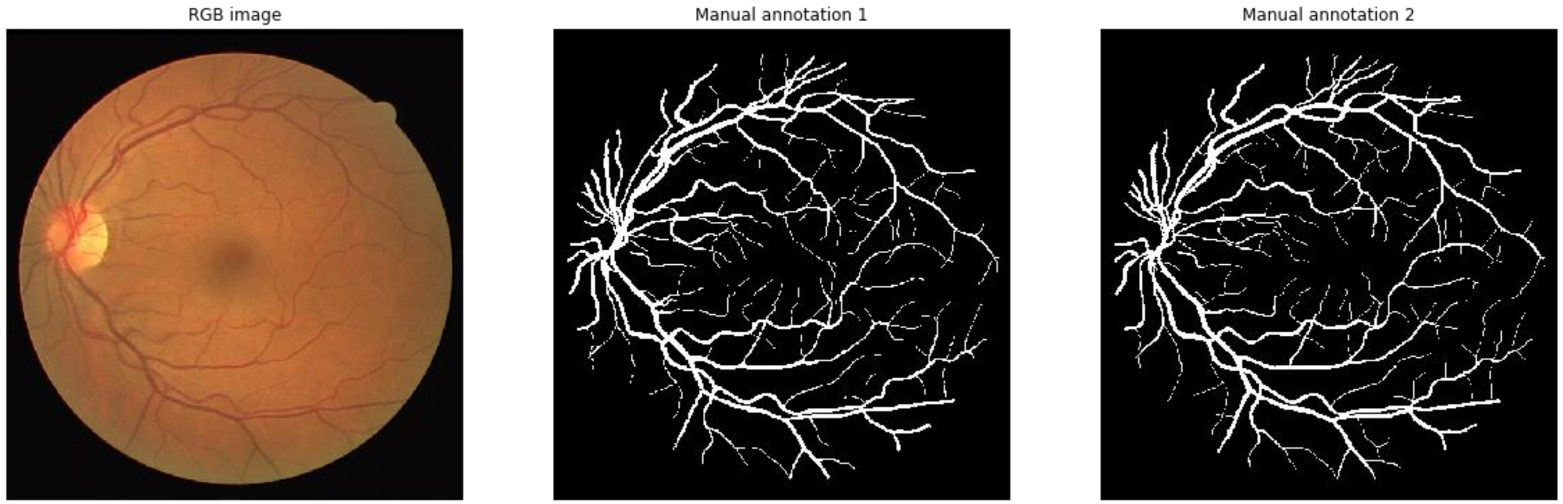
Group One

Qingru Liu (z5449255)
Zheng Lu(z5536953)
Yiyu Chen (z5581750)
Tan Peilin(z5530348)
Ash Peng(z5473493)

UNSW
SYDNEY

Motivation

- Semantic Segmentation: Identifying which image pixels belong to the class
- One of the core tasks in computer vision
- Application: Segmentation results aid in diagnostics, treatment planning, and monitoring, improving clinical efficiency and patient care.



Problem Statement for Retinal Vessel Segmentation Using Deep Learning

➢Objective: Develop a deep learning-based algorithm for the automatic semantic segmentation of retinal vessels from images.

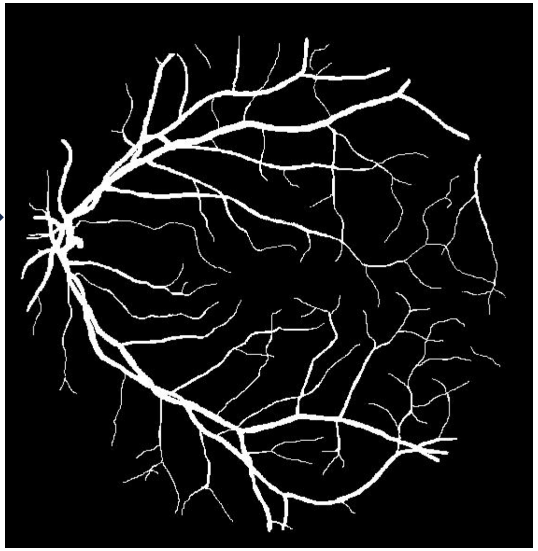
Assign every pixel's classes of the input image to extract the blood vessels of the retinal

Two classes:
1. Retinal blood vessels as foreground
2. Background

Input

Semantic Segmentation

Output



Literature Review

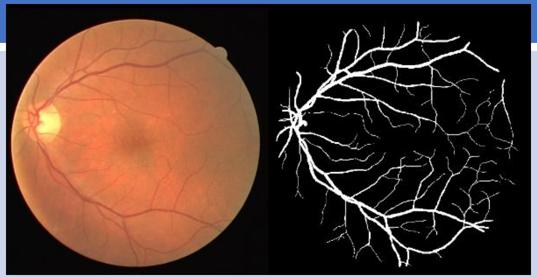
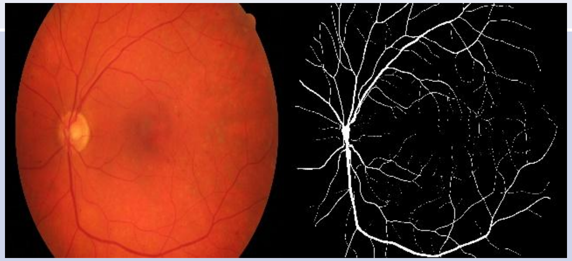
- **1.Deep Learning-based Bio-Medical Image Segmentation using UNet Architecture and Transfer Learning**
  - **Authors**: Abouzar Ghavami,Nima Hassanpour (2023)
- **Methodology**:
  - Architecture Implementation: The researchers adapted the architecture of Unet to be dynamically adjustable to the input image size.
  - Transfer Learning: Utilized transfer learning by initializing the UNet model with pretrained weights from models trained on large image datasets.
- **Evaluation**:
  - Evaluated on DRIVE databases.
  - the model achieved visually good results but had a low average Dice coefficient of 34% because of the thickness of predicted vessels.
  - Metrics: accuracy, sensitivity, specificity, AUC.
- **Significance**:
  - Addresses class imbalance and variability in image conditions.
  - Enhances U-Net for medical image segmentation.
  - A foundation for developing more robust and accurate biomedical image segmentation models.

Literature Review

- 2. **Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumors**
  - **Authors**: Dhiraj Maji,Prarthana Sigedar,Munendra Singh (2022)
- **Methodology**:
  - The study introduces a novel deep learning architecture called Attention Res-UNet with Guided Decoder (ARU-GD)
  - This methodology integrates attention mechanisms and guided decoding processes to enhance the performance of each decoder layer.
- **Evaluation**:
  - Use metrics: dice coefficient, sensitivity, and specificity to evaluate the performance.
  - providing a robust comparison between the predicted segmentation and the ground truth.
  - The performance was benchmarked against existing segmentation methods to show improvements in accuracy and efficiency.
- **Significance**:
  - Contribution to medical imaging field.
  - particularly in improving the accuracy of segmentation.
  - Provide better diagnostic tools and treatment planning in clinical settings.

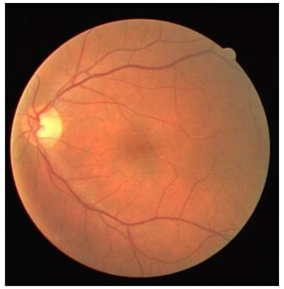
Dataset(s):

85 samples overall

Dataset Name	Dataset Source	Number of Retinal Image	Number of Segmented Image	Example
Drive	A diabetic retinopathy screening program in The Netherlands	20	20	
Stare	The Shiley Eye Center at the University of California and the Veterans Administration Medical Center	20	20	
HRF	Pattern Recognition Lab at Friedrich-Alexander-Universität	45	45	

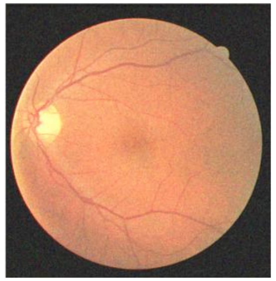
Data Analysis

- Datasets limitation
  - The current dataset comprises 85 retinal images and manually segmented images, which is **insufficient** for deep learning. This may result in **overfitting** and poor model generalization.
  - **Therefore, data augmentation is necessary to increase the amount of data.**
- Biomedical semantic feature
  - Given the minute dimensions of numerous retinal blood vessels, it is imperative that the model's network captures **intricate details with precision.**
- Data Augmentation Methods



Original

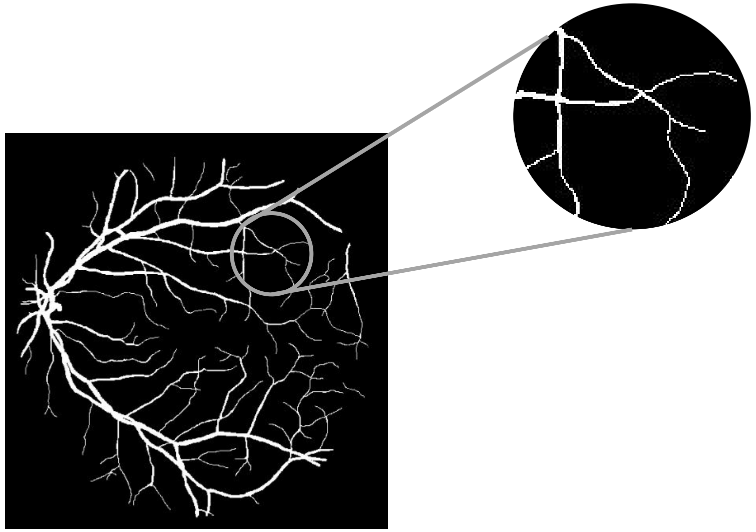
rotated



noised

blurred

- Need to captures **intricate details with precision**

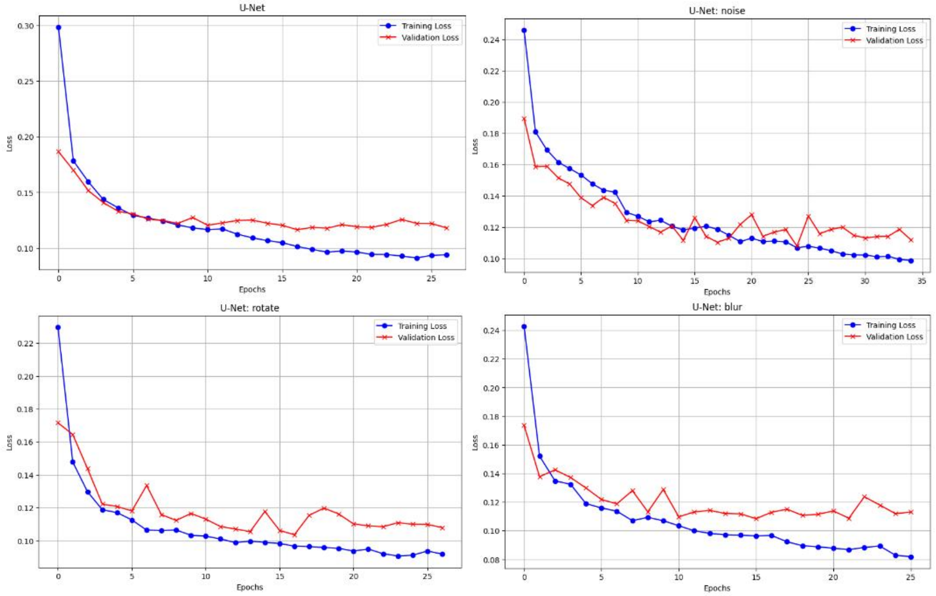


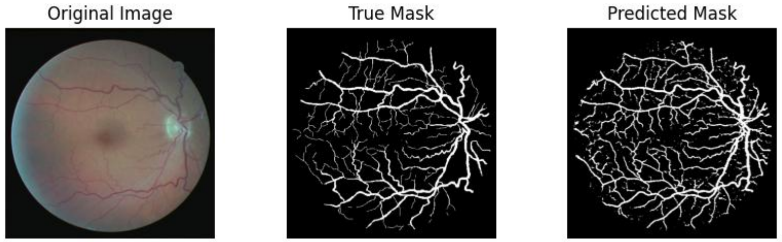
Results of DA Experiments

➢Data Augmentation Expieriments

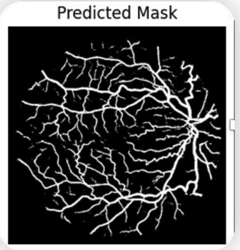
	Unet		
Data Type	DICE	IOU	ACC
Original data	0.7063	0.5493	0.9391
Rotate	0.7605	0.6142	0.9513
Add noise	0.7487	0.5991	0.9499
Gaussian blur	0.7193	0.563	0.9411

- ➢Based on the small size of our datasets, we did a lot work on the data augmentation to address the issue of data limitation based on deep learning.
- ➢First we conduct different DA methods on the original training datasets to train on the same model separately.
- ➢After comparing 3 different DA methods' performance on the testing datasets, then selected the best one DA method as the further experiments' setup.

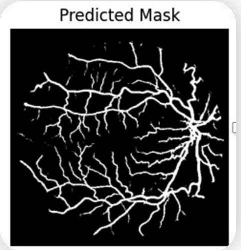




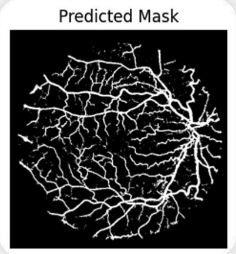
Original Data



Rotate



Noise



Gaussian Blur

Method

Baseline

Based on the benchmark

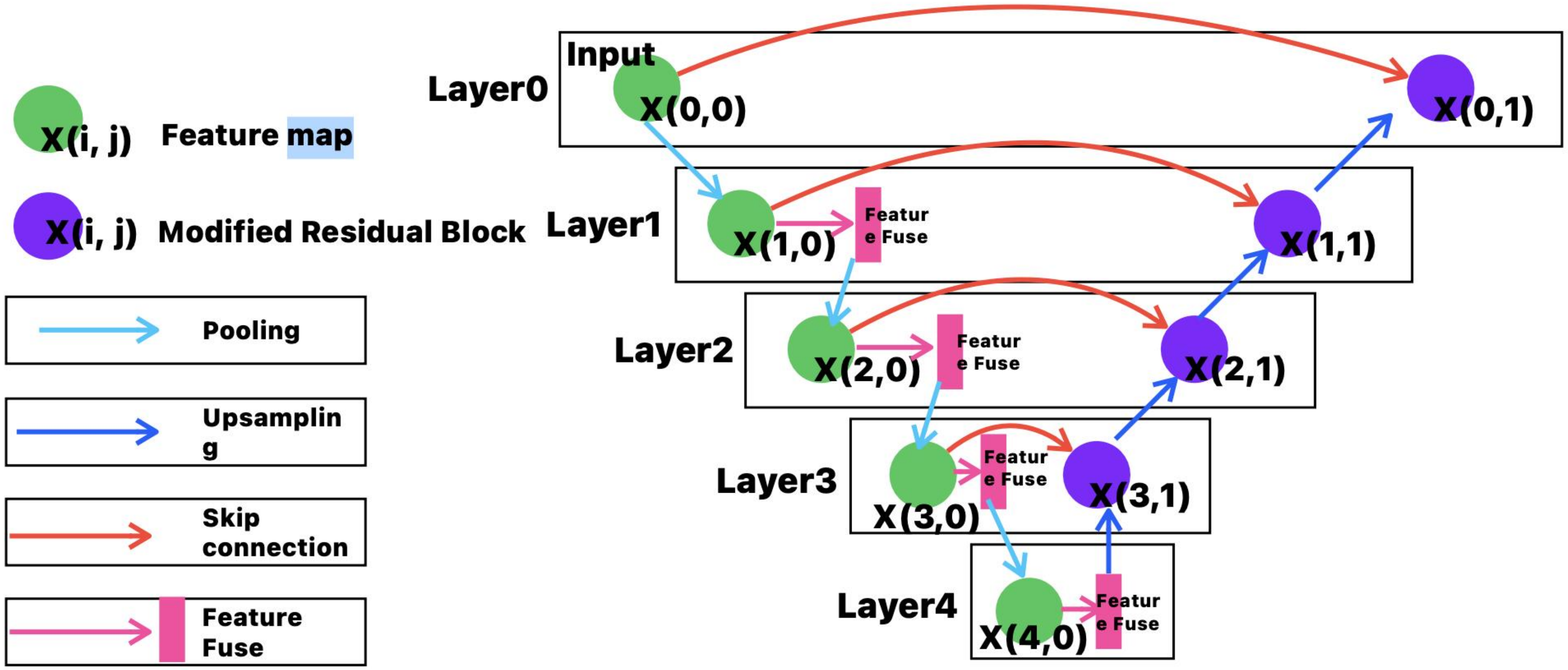
Baseline Model	Average Dice Coefficient	Average IoU	Average Accuracy
ResU-Net	0.7437	0.5949	0.9511
DeepLabv3	0.4424	0.2854	0.7848
FCN	0.6132	0.4434	0.9025
Attention UNet	**0.7543**	**0.6081**	**0.9539**

Attention UNet achives the best performance on the retinal vessels segmentation,
ResU-Net ranked the 2nd
DeepLabv3 performs poor on this datasets

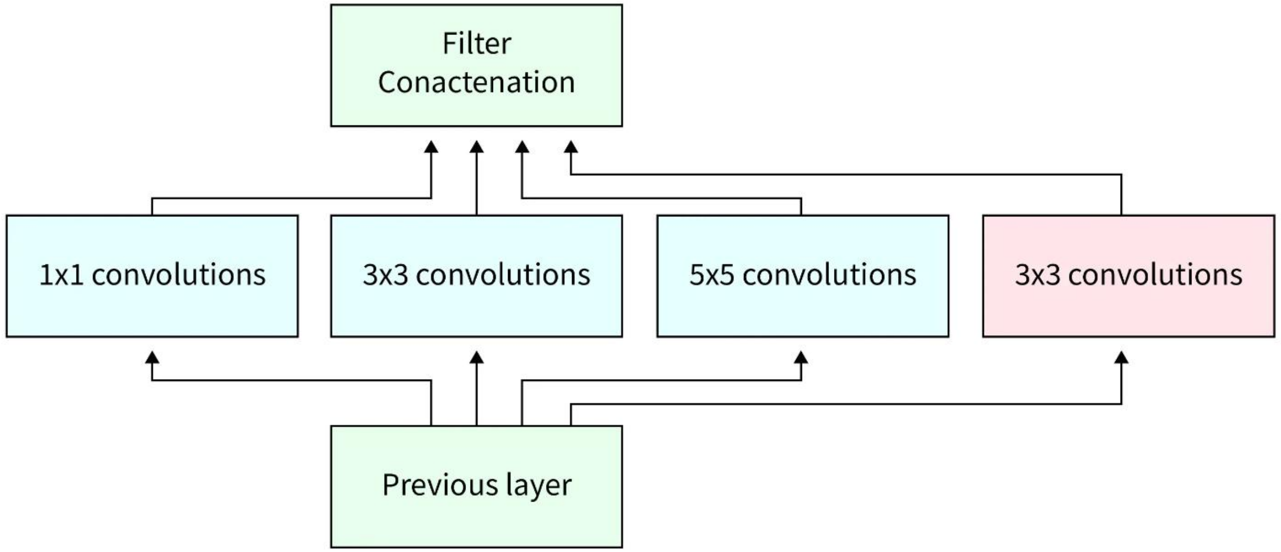
Method

1. Modified Residual Module in ResUNet
2. Feature Aggregation Block
3. Attention Block

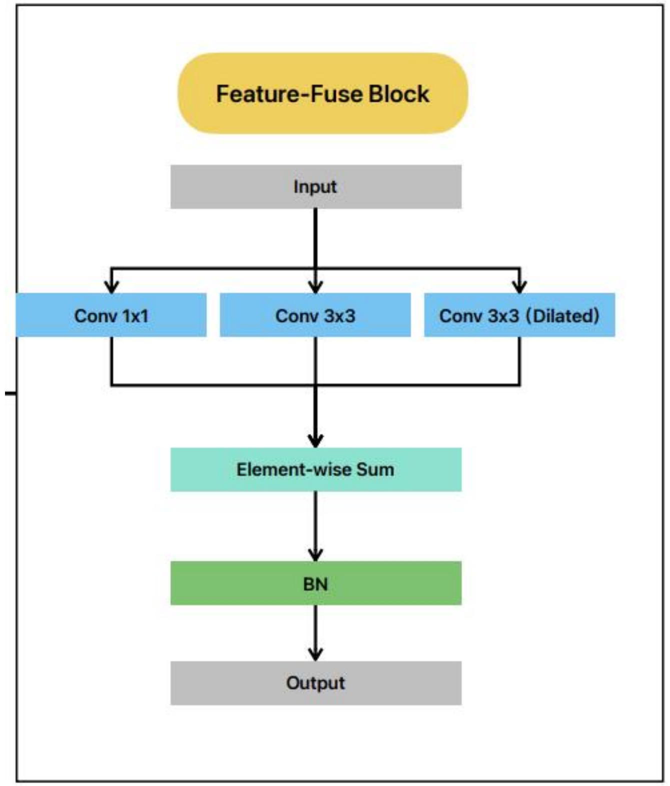
Aggregation UNet



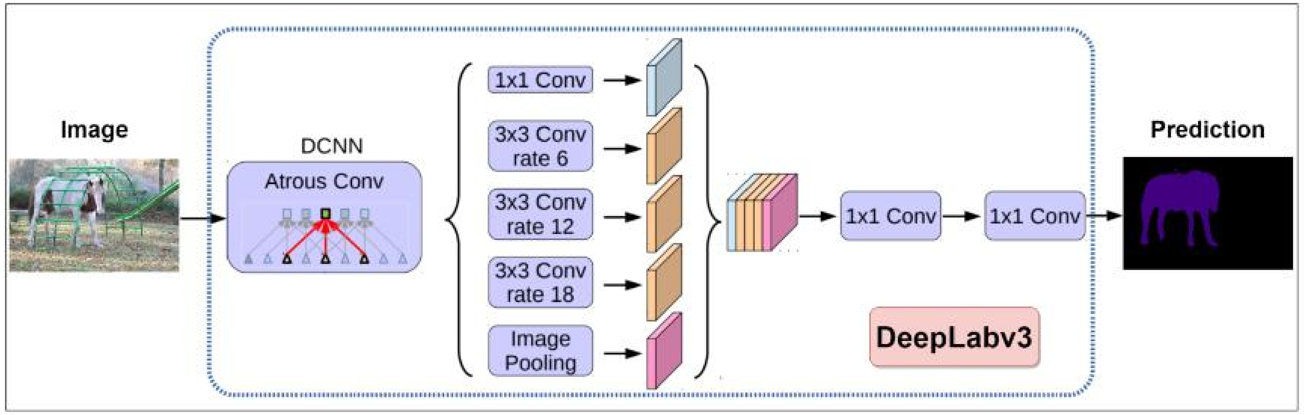
Feature Fuse Module



Inception Module, Naive Version



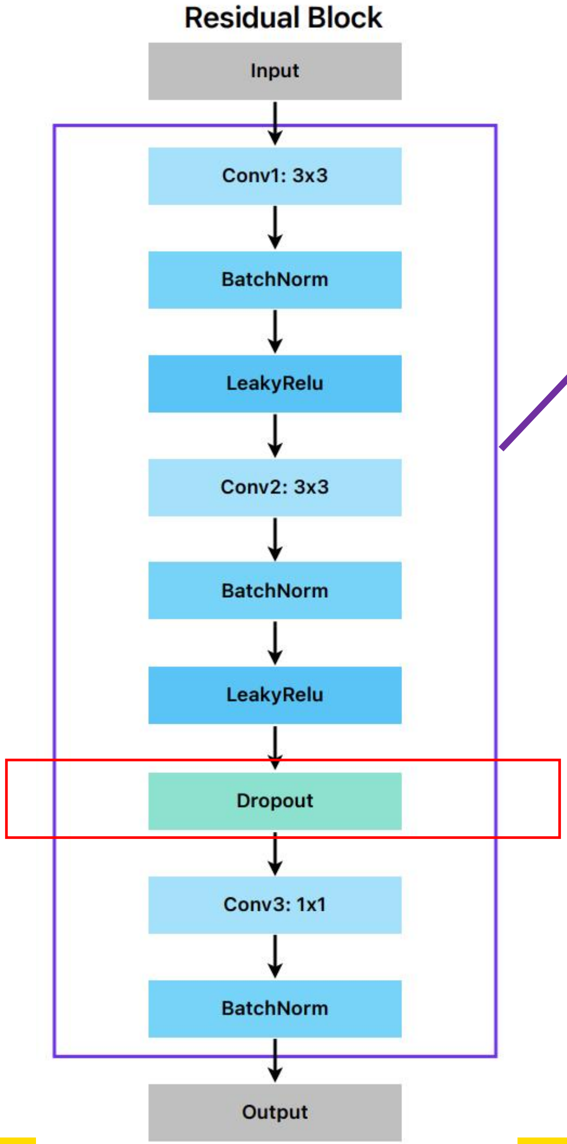
deeplabv3

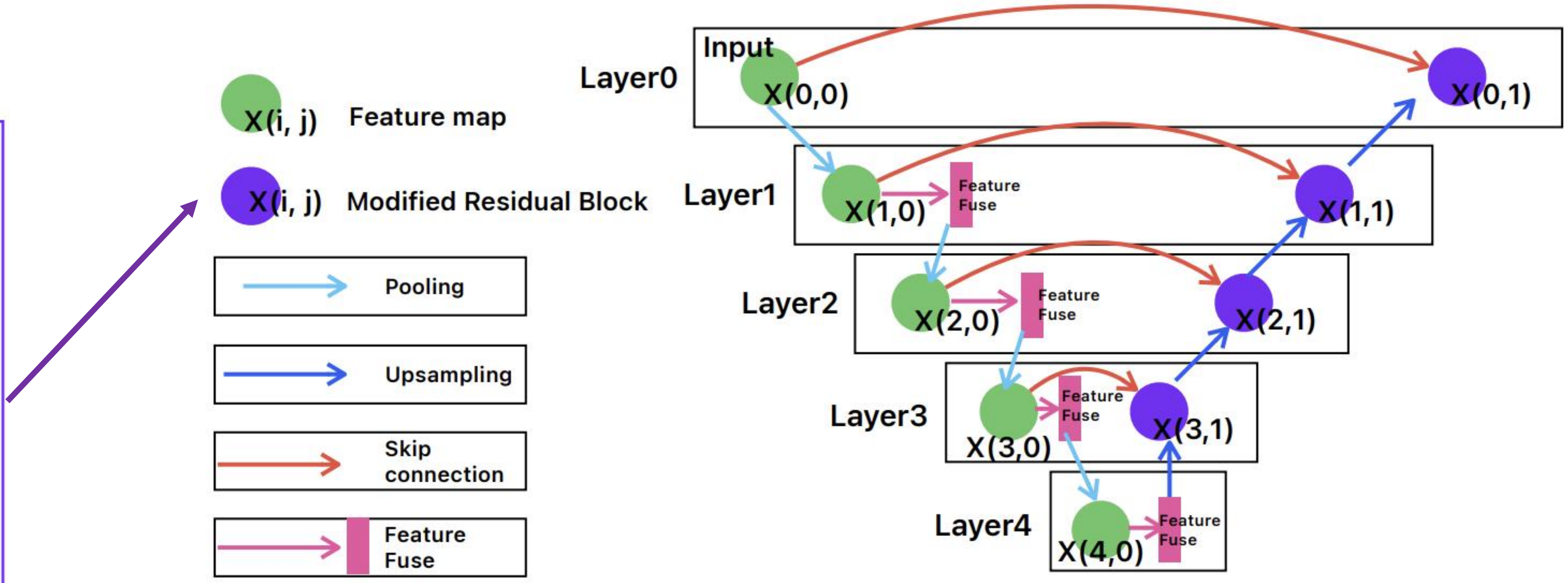


Advantage

1. a better understanding of the complex structures
2. increase the receptive field
3. Handling high-resolution images

Improvement Residual Block

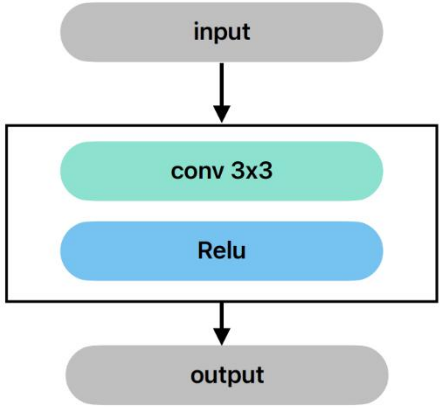




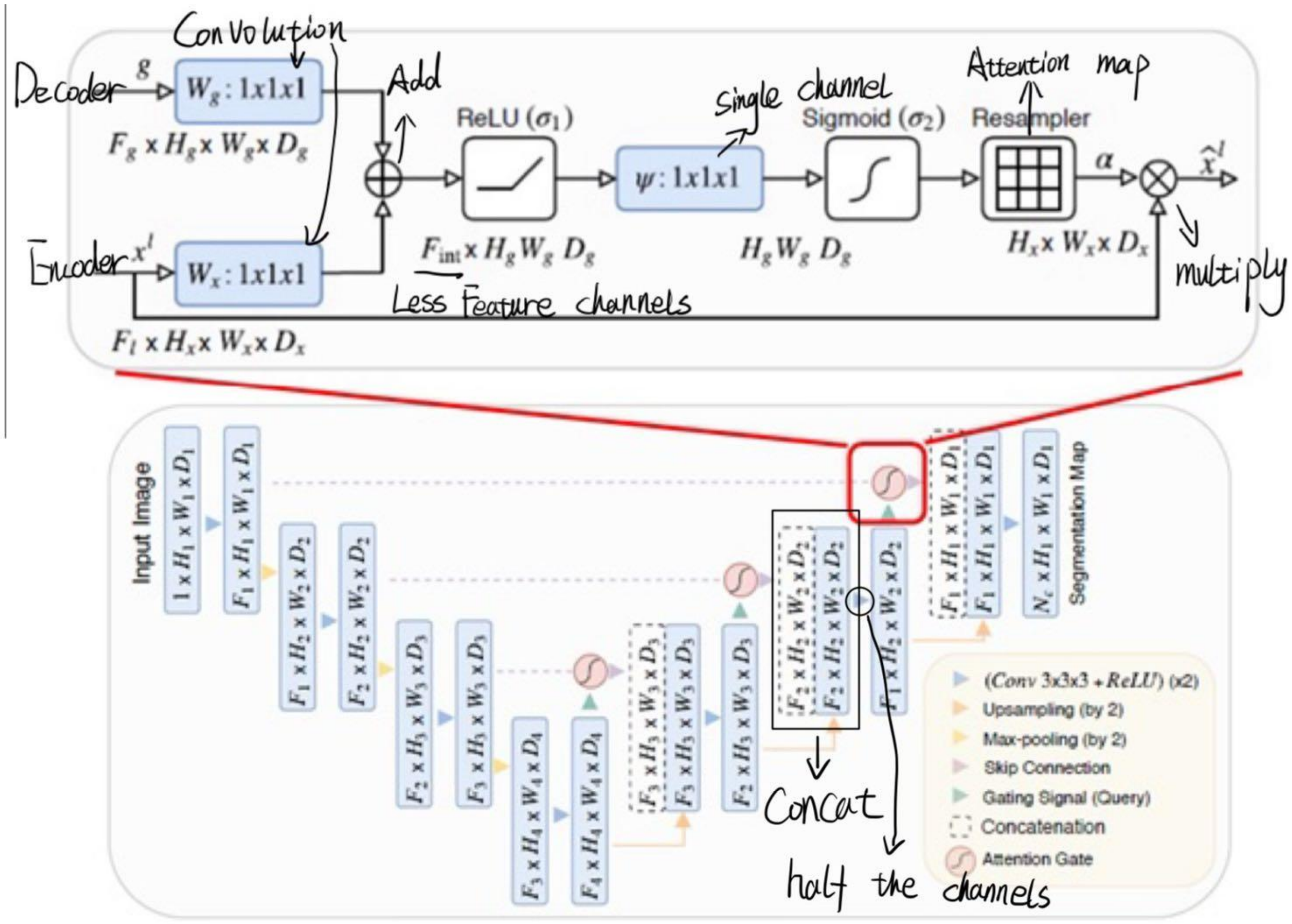
Summury

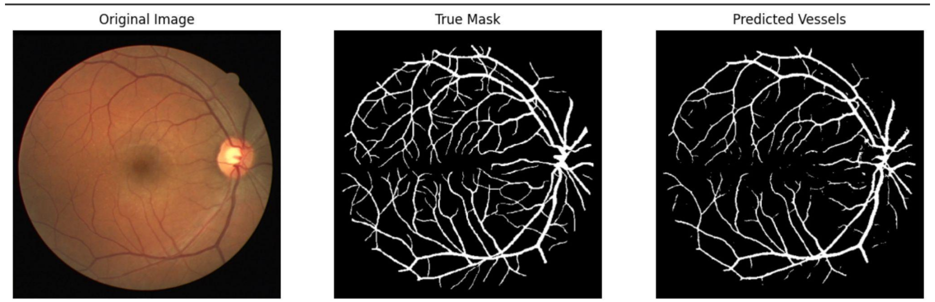
1. Complex Structure
2. LeakyRelu
3. Skip connection

Original Block



Attention Blocks

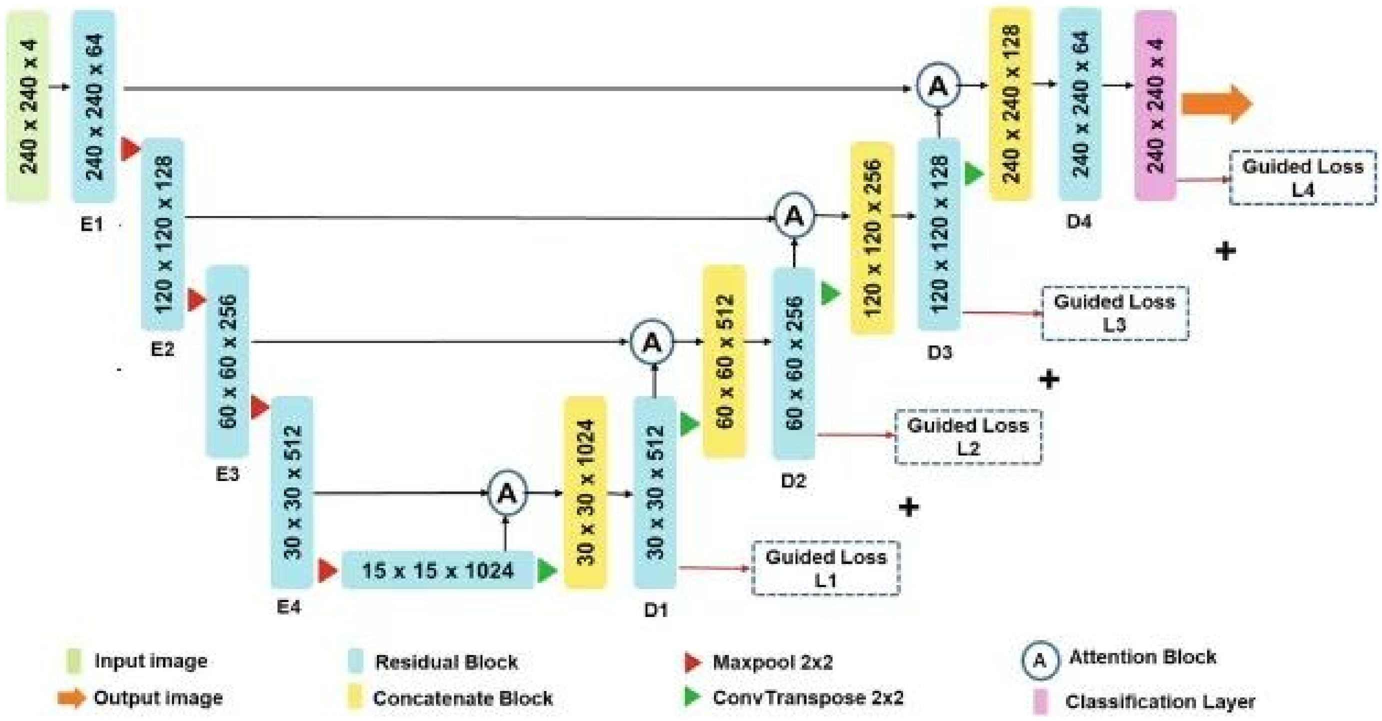




the Attention U-Net can use the feature maps from both the decoder and encoder. And use attention block for dynamic weighting, it enhances the feature fusion process, and improving the performance in segmentation.

From： https://code.likeagirl.io/u-net-vs-residual-u-net-vs-attention-u-net-vs-attention-residual-u-net-899b57c5698

Aggregation u-net with attention



By adding the attention blocks inside the model as the same structure as attention u-net, it would be able to dynamic adjust the weights during the up sampling, enhance the ability of catching features .

From： https://code.likeagirl.io/u-net-vs-residual-u-net-vs-attention-u-net-vs-attention-residual-u-net-899b57c5698

Results

Baseline

Baseline Model	Average Dice Coefficient	Average IoU	Average Accuracy
ResU-Net	0.7437	0.5949	0.9511
DeepLabv3	0.4424	0.2854	0.7848
FCN	0.6132	0.4434	0.9025
Attention UNet	**0.7543**	**0.6081**	**0.9539**

Modifed model comparison with baseline

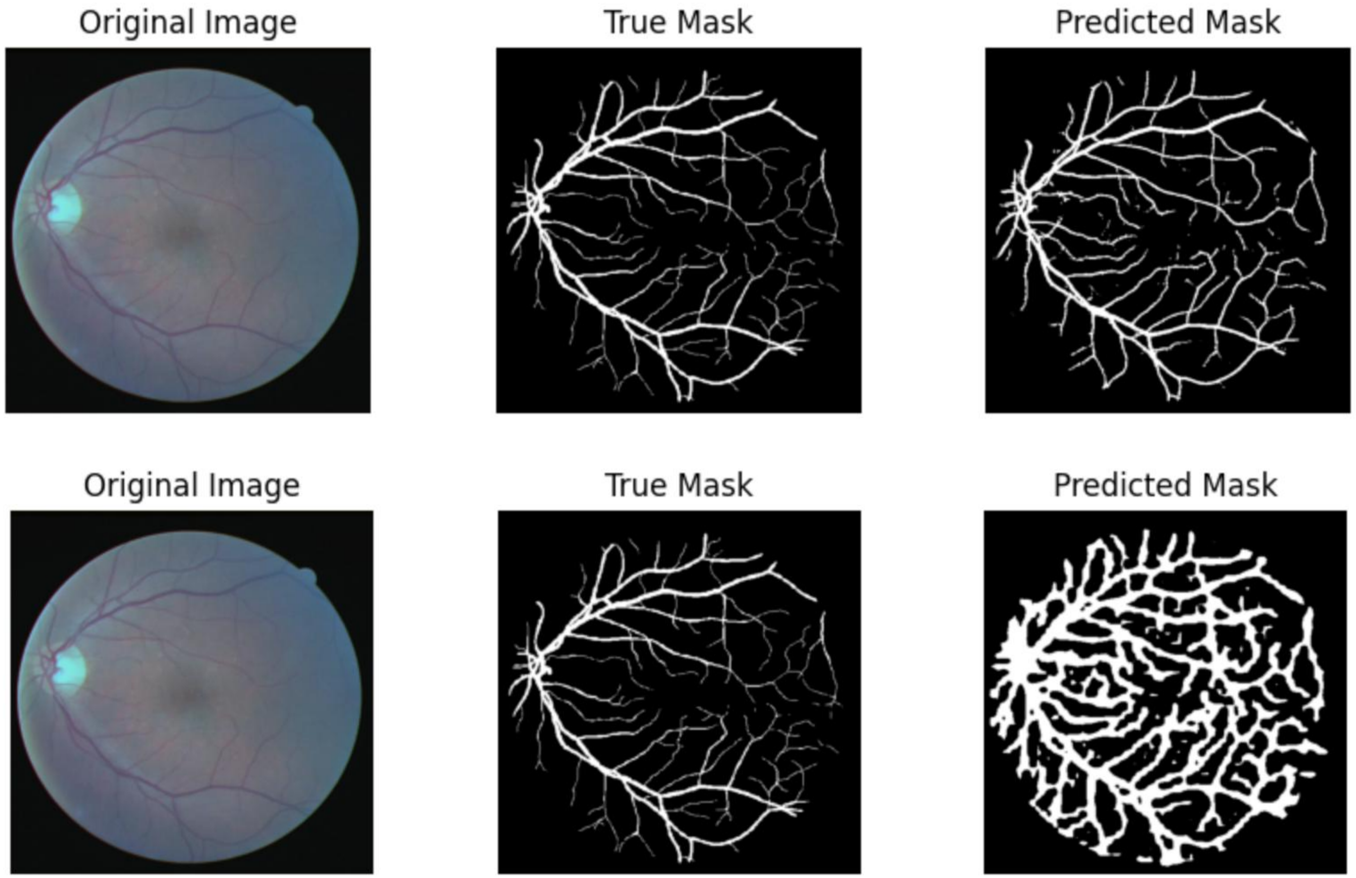
Model	Average Dice Coefficient	Average IoU	Average Accuracy
DeepLabv3	0.4424	0.2854	0.7848
FCN	0.6132	0.4434	0.9025
Attention UNet	0.7543	0.6081	0.9539
ResU-Net	0.7437	0.5949	0.9511
Aggregation UNet	**0.7784**	**0.6383**	**0.9619**
Aggregation UNet with Attention	0.7651	0.6218	0.9550

Compared with UNet, our model increased 3.47% from 74. 37 to 77.84 on IoU

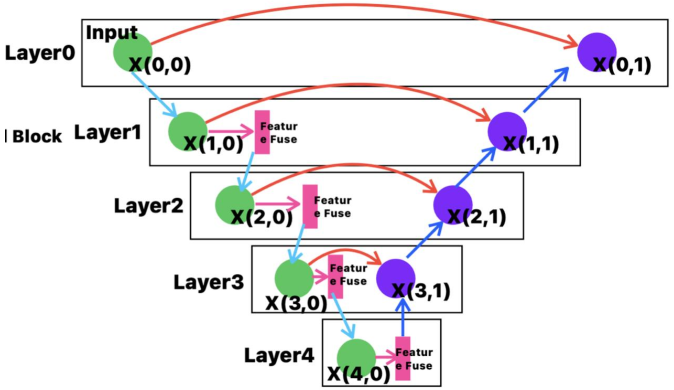
- **Experiment Setup**
  - **Data Split**:
    - 80% Train
    - 10% Validation
    - 10% Test
    - Total: 85 datasets
  - **Data Augmentation**:
    - Random Rotation
  - **Loss Function**:
    - `nn.BCEWithLogitsLoss` for binary classification.
  - **Optimizer**:
    - `optim.AdamW` with `lr=0.001` and `weight_decay=0.01`.
  - **Epochs**:
    - Maximum of 100 epochs.
    - May stop earlier if there is no improvement on validation.
  - **Batch Size**: 1

Discussion

U-Net VS. DeepLabv3



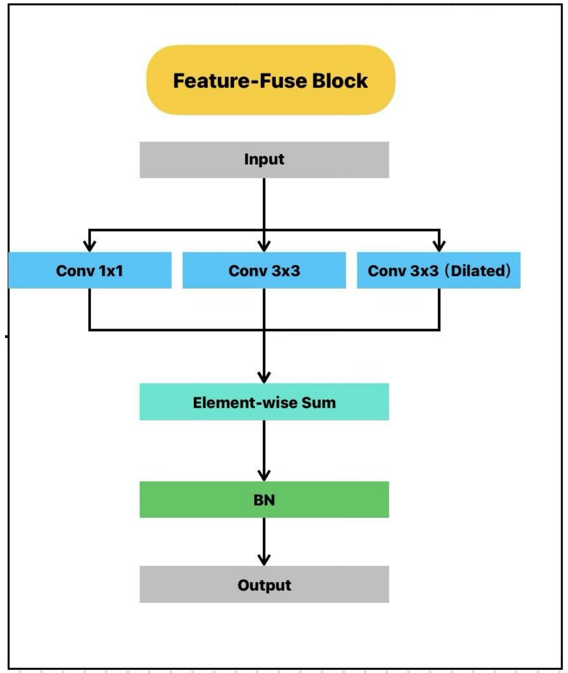
- Both are **Milestone Architectures in Semantic Segmentation**
- **Complexity & Parameters Count** :
  DeepLabv3 > U-Net
  - Dataset Size Requirement: **DeepLabv3 requires larger datasets**
  - **Performance on limited Small Datasets**: DeepLabv3 poor
- **Medical Imaging Performance**: DeepLabv3 poor on **fine-detail** segmentation (e.g., blood vessels)
  - High Level Resolution Feature Preservation:
    U-Net better due to **skip connections**

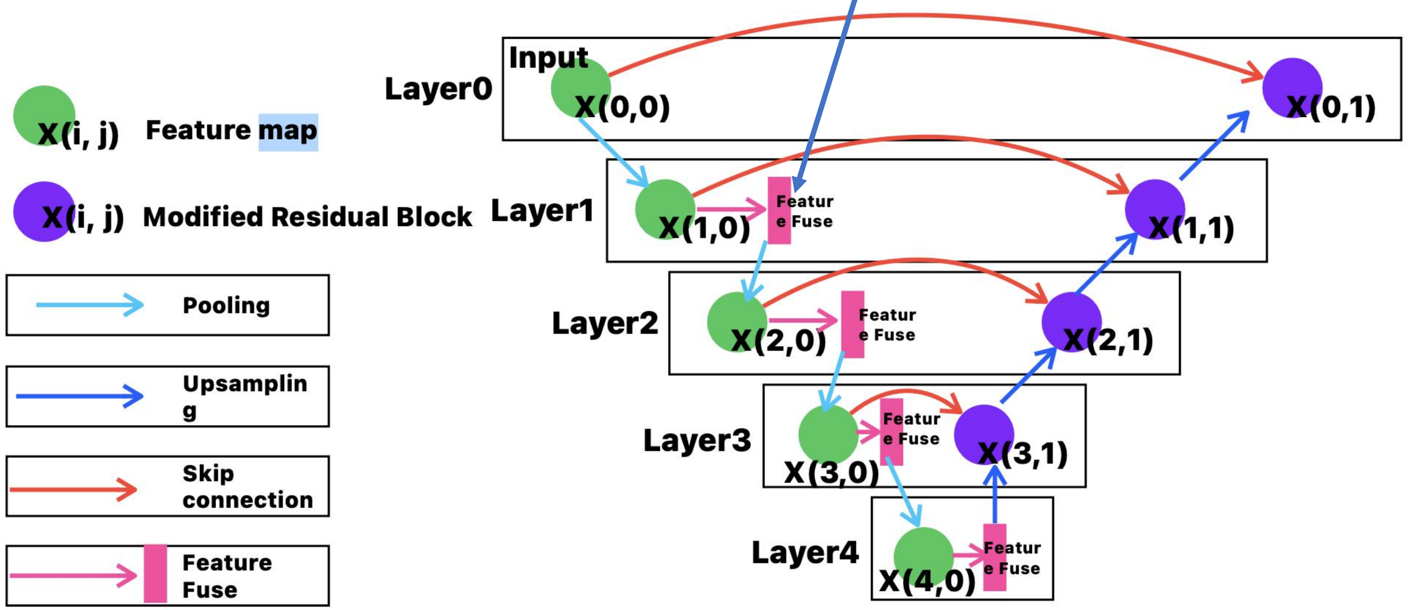


Discussion

However, does this mean DeepLabv3 does not work on our task at all?

- No. We also combine the ASPP Module in DeepLabv3 as the Feature Fuse Block into UNet

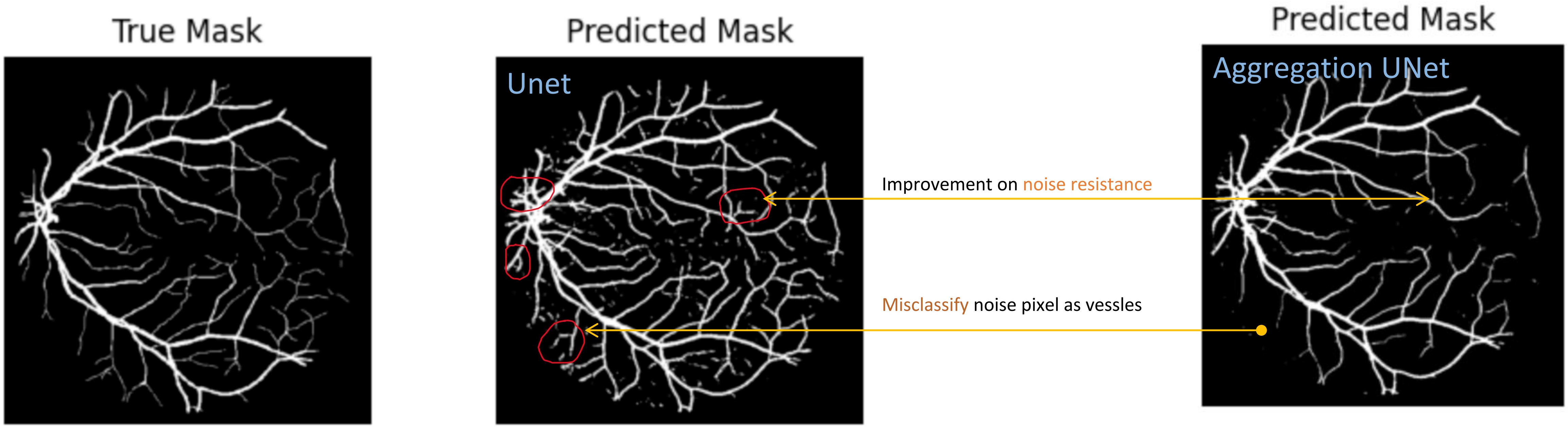




Discussion

- Our Aggregation UNet Vs. UNet

Compared with UNet, our model increased 3.47% from 74. 37 to 77.84 on IoU



Which is crucial for medical diagonal process

This is due to:

- Although Original U-Net still better at extracting fine features
- **Error in Predictions**: U-Net prone to misclassify noise as fine blood vessels

  Noise Resistance: Improved models perform better than U-Net

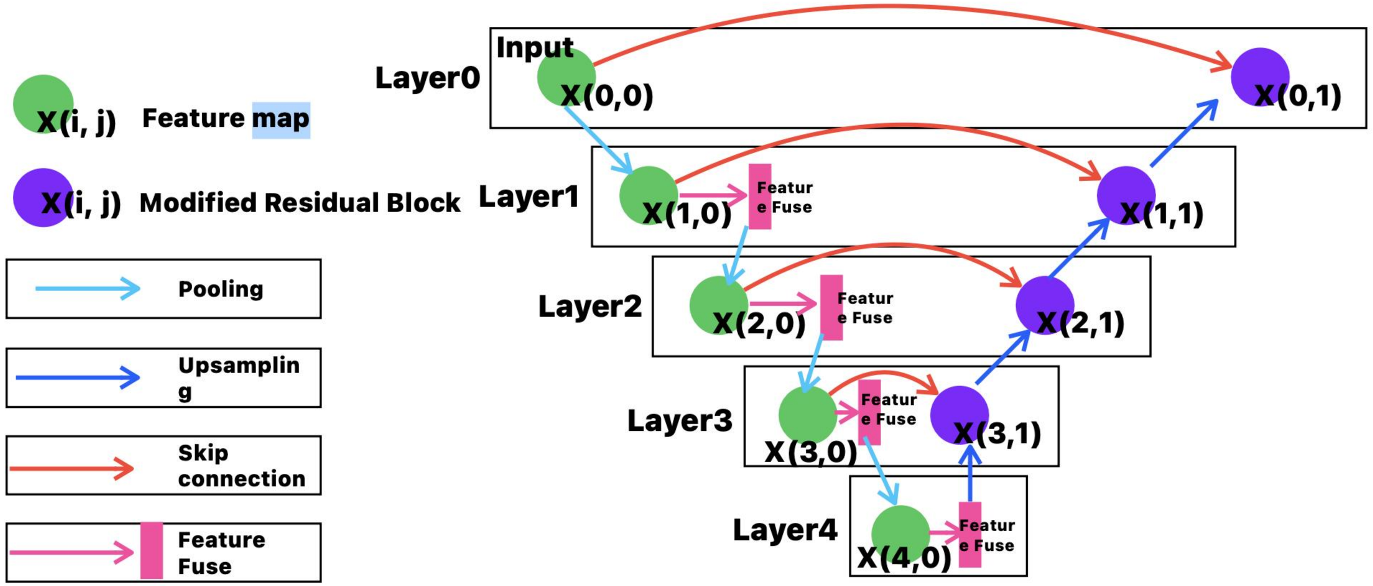
Presenter: Qingru Liu (z5449255)

UNSW SYDNEY

Conclusion

Summary of our modified model (Aggregation Unet)

- Retained **skip connections** in the UNet
  which is effective concatenation of low-level, high-resolution features.
- Incorporated **parallel convolution modules** from DeepLabv3
  to expand the receptive field, enhancing encoder feature extraction.
- Achieved improvement in the IoU metric.

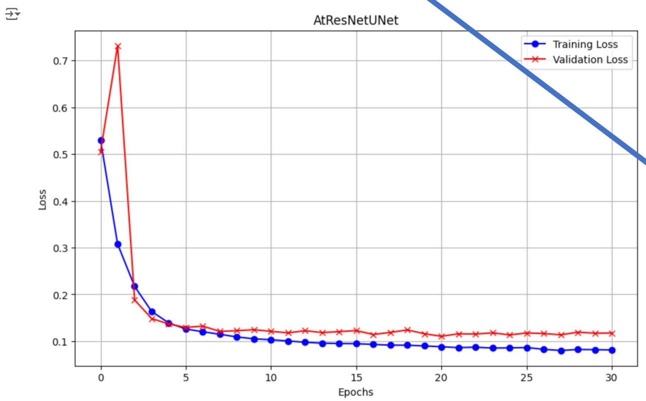


Conclusion

Summary of Work

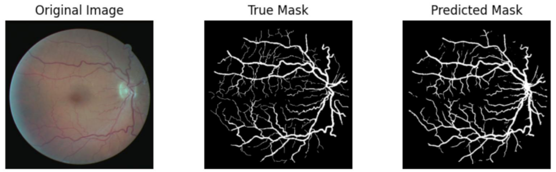
- **Address the Dataset Size Limitation:**
  - **Data augmentation** experiments
- **Baseline Model Comparisons on 5 models**
- **Experiments on Modified U-Net Models:**
  - Modify the residual block structures.
  - **Introduce feature fusion blocks** after down-sampling steps.
  - Implement residual-attention blocks after up-sampling steps.
- **Performance Improvement:**
  - The modified Aggregation U-Net **achieved a 3.47% improvement in performance**

- **All work & experiments has been organized into our Jupyter notebook**
- Environment requirements
- Experiment's settings, along with hyperparameters
- Training log & comparative results
- a **complete pipeline** capable of running all experiments from loading datasets to eva




```
[ ] model_path = os.path.join(dir_save, f'AtResNetUNet1.pth')
    model.load_state_dict(torch.load(model_path))
    evaluate_model(model, test_loader, device)
    evaluate_and_plot(model, test_loader, device, num_images_to_plot=3)
```

```
Average Dice Coefficient: 0.7720
Average IoU: 0.6300
Average Accuracy: 0.9571
```



- Datasets
- Models
  - ResUNet
  - Attention UNet
  - DeepLabv3
  - FCN
  - Aggregation UNet: Improved model based on U-Net
    - Aggregation UNet
    - Aggregation UNet with Residual Attention
- **Retinal Blood Vessel Semantic Segmentation Pipeline**
  - Environment Requirements
  - Load Datasets
  - Dataset preprocess
  - Split Train-val-test
  - Display the images & masks in dataloader
  - Trainnig model
    - plot the loss during training
  - Evaluate the model

- Experiments
  - Baseline models experiments
    - Batch Size Experiment
      - Batch Size = 8
      - Batch size = 1
    - Deeplabv3
    - ResUNet
    - FCN
    - Attention UNet
  - **Data Augmentation Expiremnets**
    - Implement the DA comparsion pipeline
    - Display of creating an augmentation dataset after random rotate
    - Training models with different DA method
    - Evaluation the performance of different DA method
    - DA Comparison Experiments Summary
    - Experiment on ResUNet after DA

1. **Average Dice Coefficient**:

$$\text{Dice Coefficient} = \frac{2 \times |X \cap Y|}{|X| + |Y|}$$

Where X is the predicted set of pixels and Y is the ground truth.

2. **Average IoU (Intersection over Union)**:

$$\text{IoU} = \frac{|X \cap Y|}{|X \cup Y|}$$

Where X is the predicted set of pixels and Y is the ground truth.

3. **Average Accuracy**:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

Presenter: Qingru Liu (z5449255)


Questions