MultiResUNet : Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation

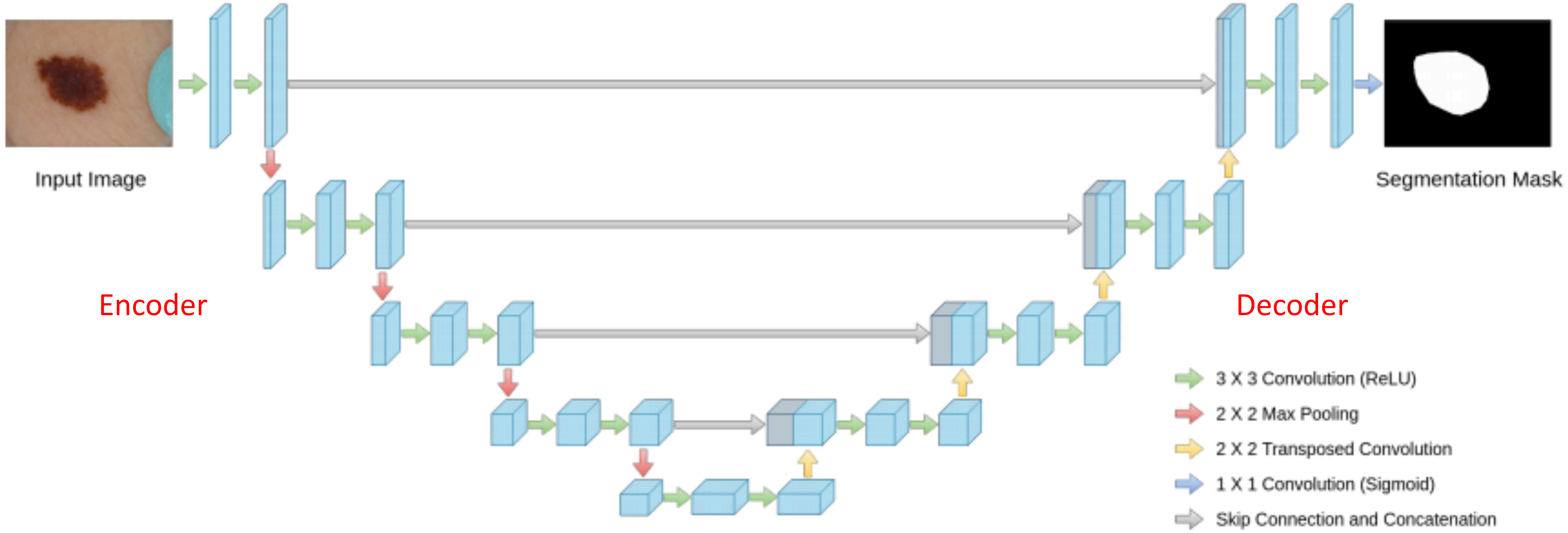
Author: Nabil Ibtehaz and M. Sohel Rahman

Presenter: Luu Minh Sao Khue

Introduction

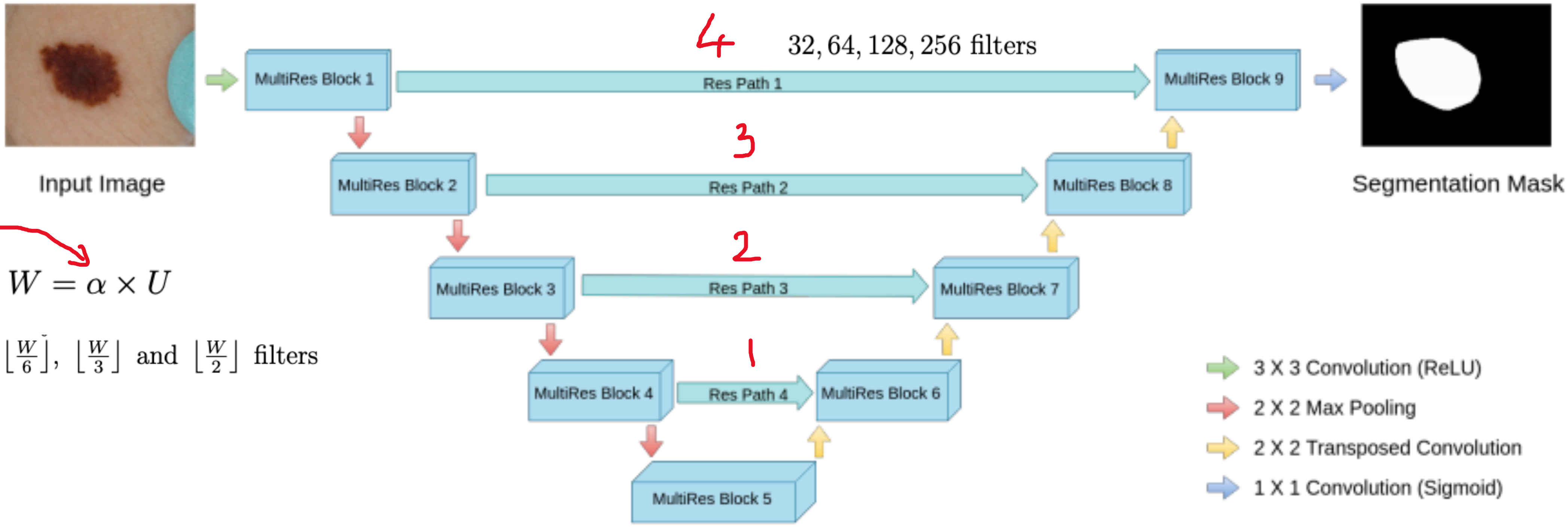
- What have been done:
  - Add modifications to improve upon the U-Net model
  - Compare MultiResUNet with the classical U-Net 5 datasets of various medical images
- Result:
  - Ideal images: slight improvements
  - Challenging images: remarkable gain in performance
  - Relative improvements respectively: 10.15%, 5.07%, 2.63%, 1.41%, and 0.62%

Overview of the U-Net Architecture



Proposed Architecture

- Replace the sequence of two convolutional layers with the proposed MultiRes block

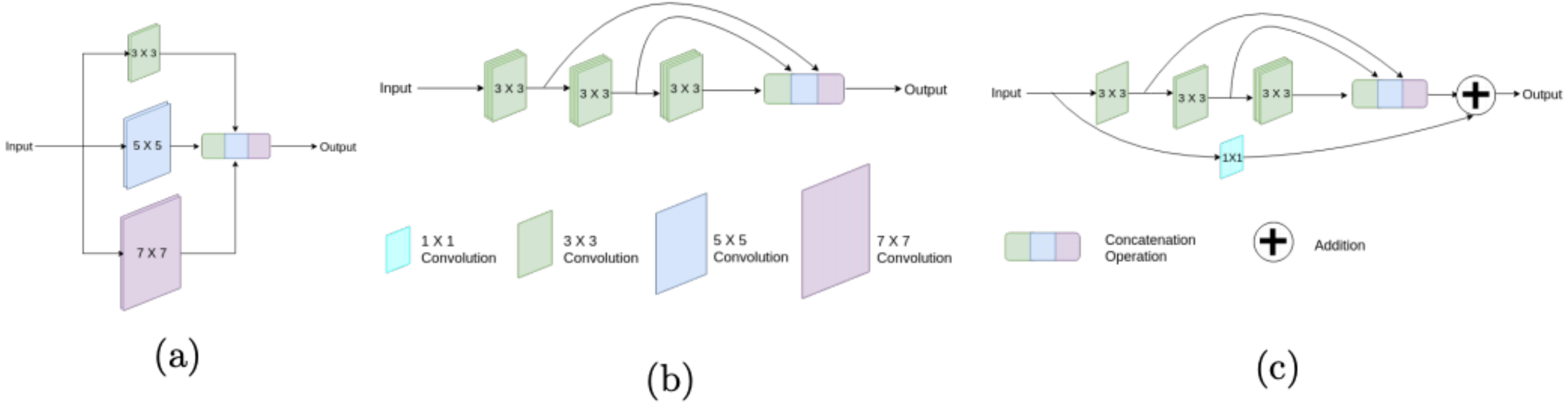


1.67

$$W = \alpha \times U$$

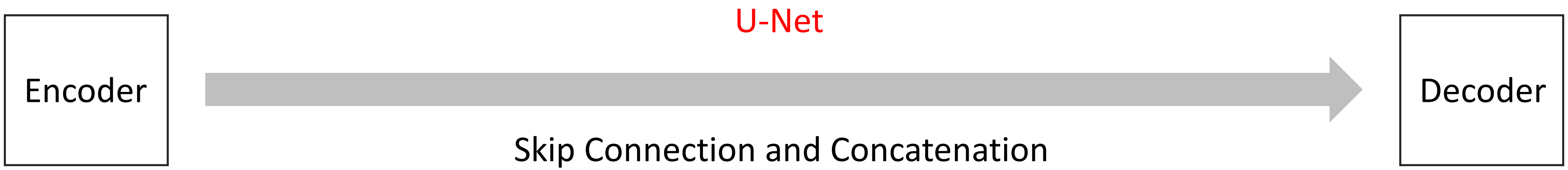
$\lfloor \frac{W}{6} \rfloor$, $\lfloor \frac{W}{3} \rfloor$ and $\lfloor \frac{W}{2} \rfloor$ filters

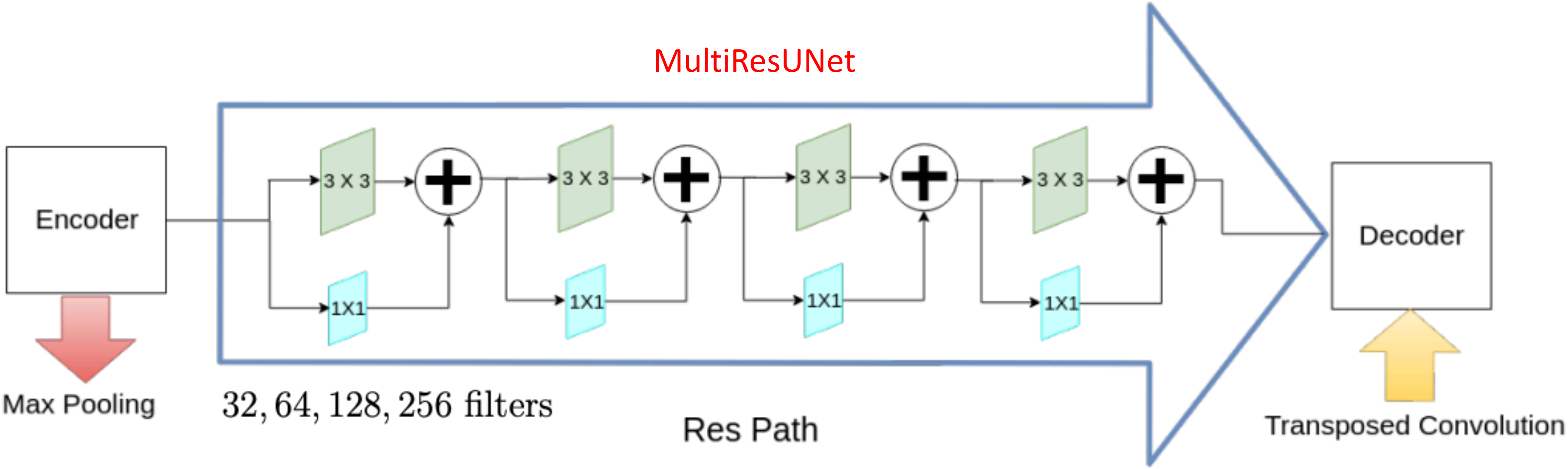
MultiRes block – capture features in different scales



- Replacing the convolutional layers with Inception-like blocks should facilitate the U-Net architecture to reconcile the features learnt from the image at different scales
- Factorize the bigger, more demanding 5 × 5 and 7 × 7 convolutional layers, using a sequence of smaller and lightweight 3 × 3 convolutional blocks
- Gradually increase the filters in those (from 1 to 3), to prevent the memory requirement of the earlier layers from exceedingly propagating to the deeper part of the network

Res path – reduce semantic gap





Datasets

- Fluorescence Microscopy Image
- Electron Microscopy Image
- Dermoscopy Image
- Endoscopy Image
- Magnetic Resonance Image

Table 2: Overview of the Datasets.

Modality	Dataset	No. of images	Original Resolution	Input Resolution
Fluorescence Microscopy	Murphy Lab	97	Variable	256×256
Electron Microscopy	ISBI-2012	30	512×512	256×256
Dermoscopy	ISIC-2018	2594	Variable	256×192
Endoscopy	CVC-ClinicDB	612	384×288	256×192
MRI	BraTS17	210 HGG + 75 LGG	$240 \times 240 \times 155$	$80 \times 80 \times 48$

Experiments

- Baseline: original U-Net with five-layer deep encoder and decoder, with filter numbers of 32, 64, 128, 256, 512.
- 3D version: substituting the 2D with the 3D coutnerparts without any further alterations
- Pre-processing: resize and convert to range [0..1]
- No post-processing
- Sigmoid on the last layer with threshold = 0.5

2D		3D	
Model	Parameters	Model	Parameters
U-Net (baseline)	7,759,521	3D U-Net (baseline)	19,078,593
MultiResUNet (proposed)	7,262,750	MultiResUNet 3D (proposed)	18,657,689

Experiments

- 5-fold cross validation over 150
- Optimizer: Adam
- Loss function:

$$Cross\ Entropy(X, Y, \hat{Y}) = \sum_{px \in X} -(y_{px} \log(\hat{y}_{px}) + (1 - y_{px}) \log(1 - \hat{y}_{px}))$$

$$J = \frac{1}{n} \sum_{i=1}^{n} Cross\ Entropy(X_i, Y_i, \hat{Y}_i)$$

- Evaluation metric:

$$Jaccard\ Index = \frac{Intersection}{Union} = \frac{A \cap B}{A \cup B}$$

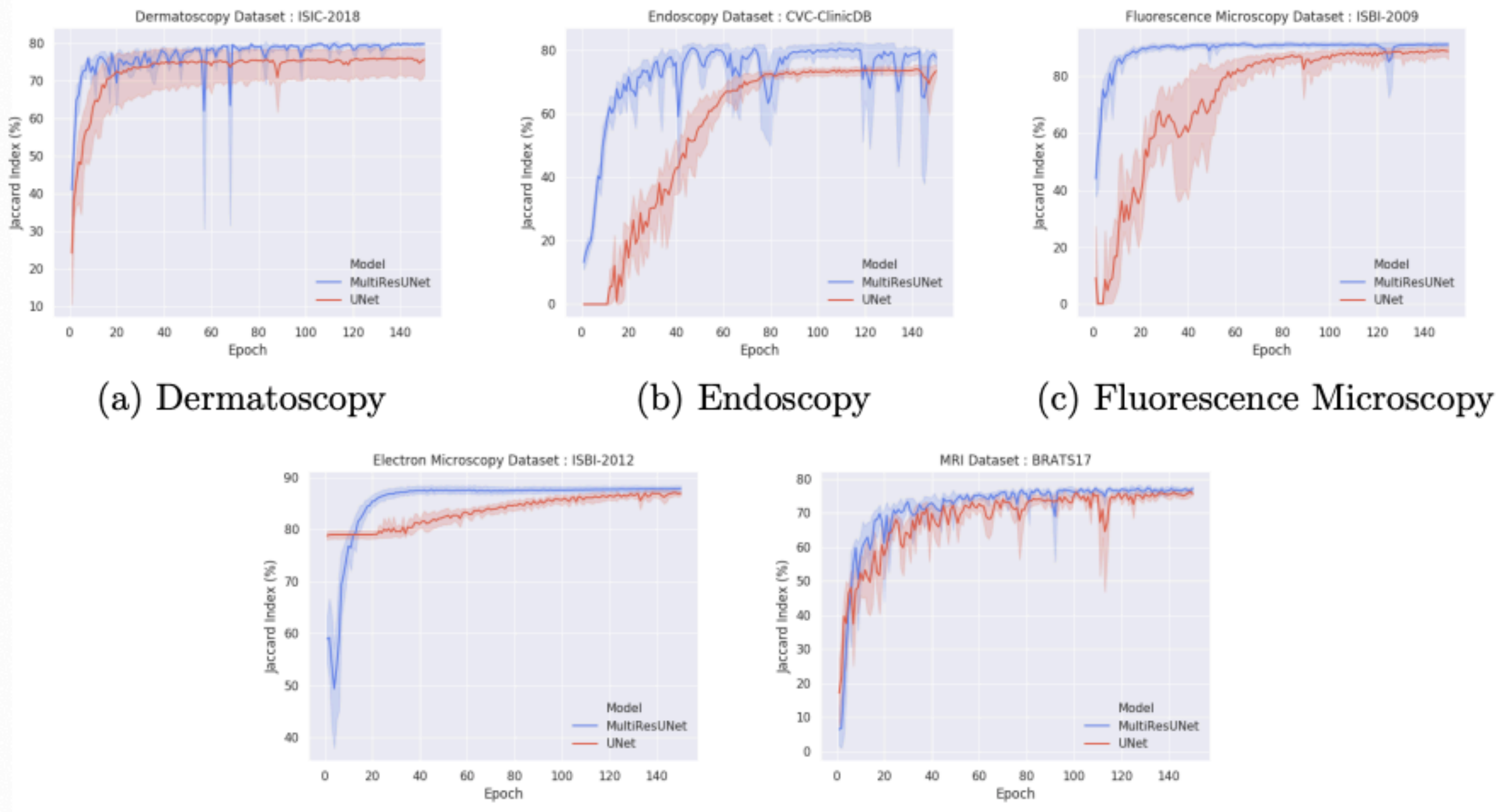
Results

- MultiResUNet Consistently Outperforms U-Net
  - On all different types of medical images, remarkable improvements for Dermoscopy and Endoscopy images (less uniform images)

Modality	MultiResUNet (%)	U-Net (%)	Relative Improvement (%)
Dermoscopy	80.2988 ± 0.3717	76.4277 ± 4.5183	5.065 ★
Endoscopy	82.0574 ± 1.5953	74.4984 ± 1.4704	10.1465 ★
Fluorescence Microscopy	91.6537 ± 0.9563	89.3027 ± 2.1950	2.6326
Electron Microscopy	87.9477 ± 0.7741	87.4092 ± 0.7071	0.6161
MRI	78.1936 ± 0.7868	77.1061 ± 0.7768	1.4104

Results

- MultiResUNet can Obtain Better Results in Less Number of Epochs

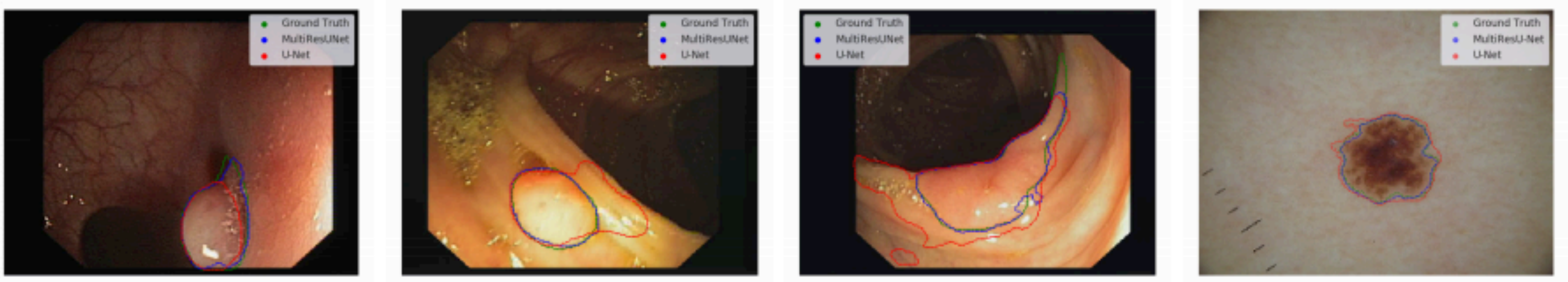


(a) Dermatoscopy (b) Endoscopy (c) Fluorescence Microscopy

(d) Electron Microscopy (e) MRI

Results

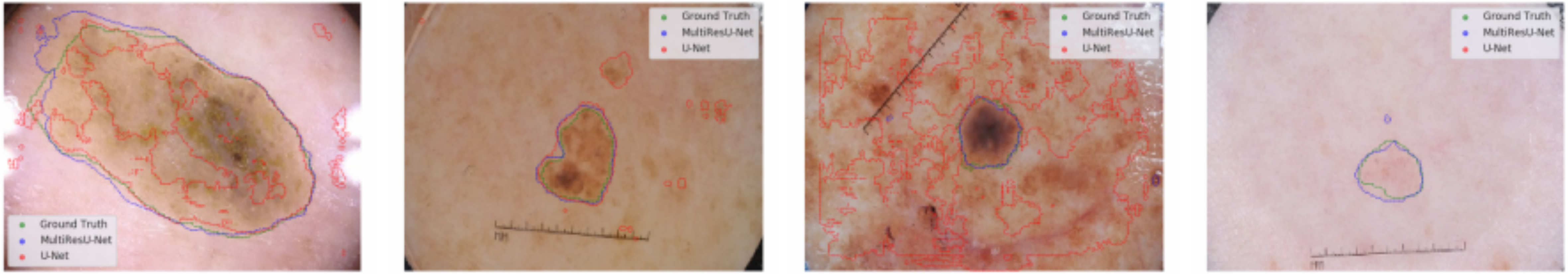
- MultiResUNet Delineates Faint Boundaries Better
  - For more challenging images, especially with not so much conspicuous boundaries, U-Net seems to be struggling a bit



(a) (b) (c) (d)

Results

- MultiResUNet is More Immune to Perturbations
  - U-Net was unable to segment the forground as a continuous region
  - for images where the background is not uniform, the U-Net model seems to make some false predictions
  - More false on the rough background or even fail to make predictions

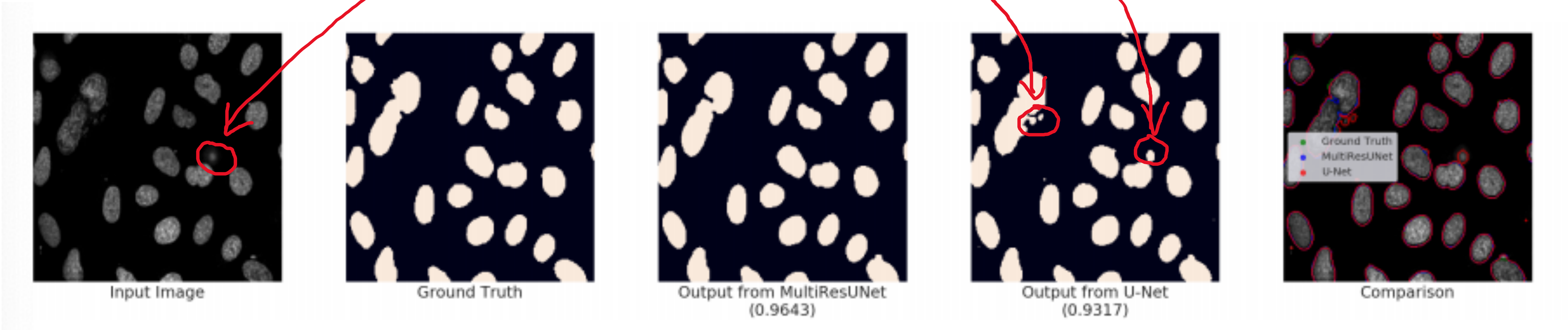


(a) (b) (c) (d)

Results

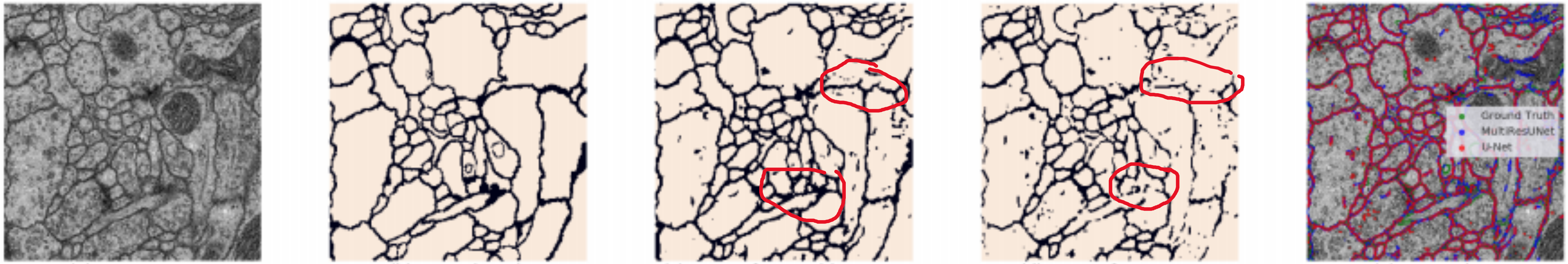
- MultiResUNet is More Reliable Against Outliers
  - MultiResUNet segmentation on outliers were consistently better than that of the U-Net.
  - fluorescence microscopy images

Not cell nuclei



Results

- MultiResUNet on Segmenting the Majority Class
  - Usually, ROI consists of a small portion, but in the Electron Microscopy dataset the ROI under consideration comprises the majority of the images



(a) Input Image (b) Ground Truth (c) Output from MultiResUNet (0.8914) (d) Output from U-Net (0.8841) (e) Comparison

References

- MultiResUNet : Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation (https://arxiv.org/pdf/1902.04049.pdf)
- U-Net: Convolutional Networks for Biomedical Image Segmentation (https://arxiv.org/pdf/1505.04597.pdf)