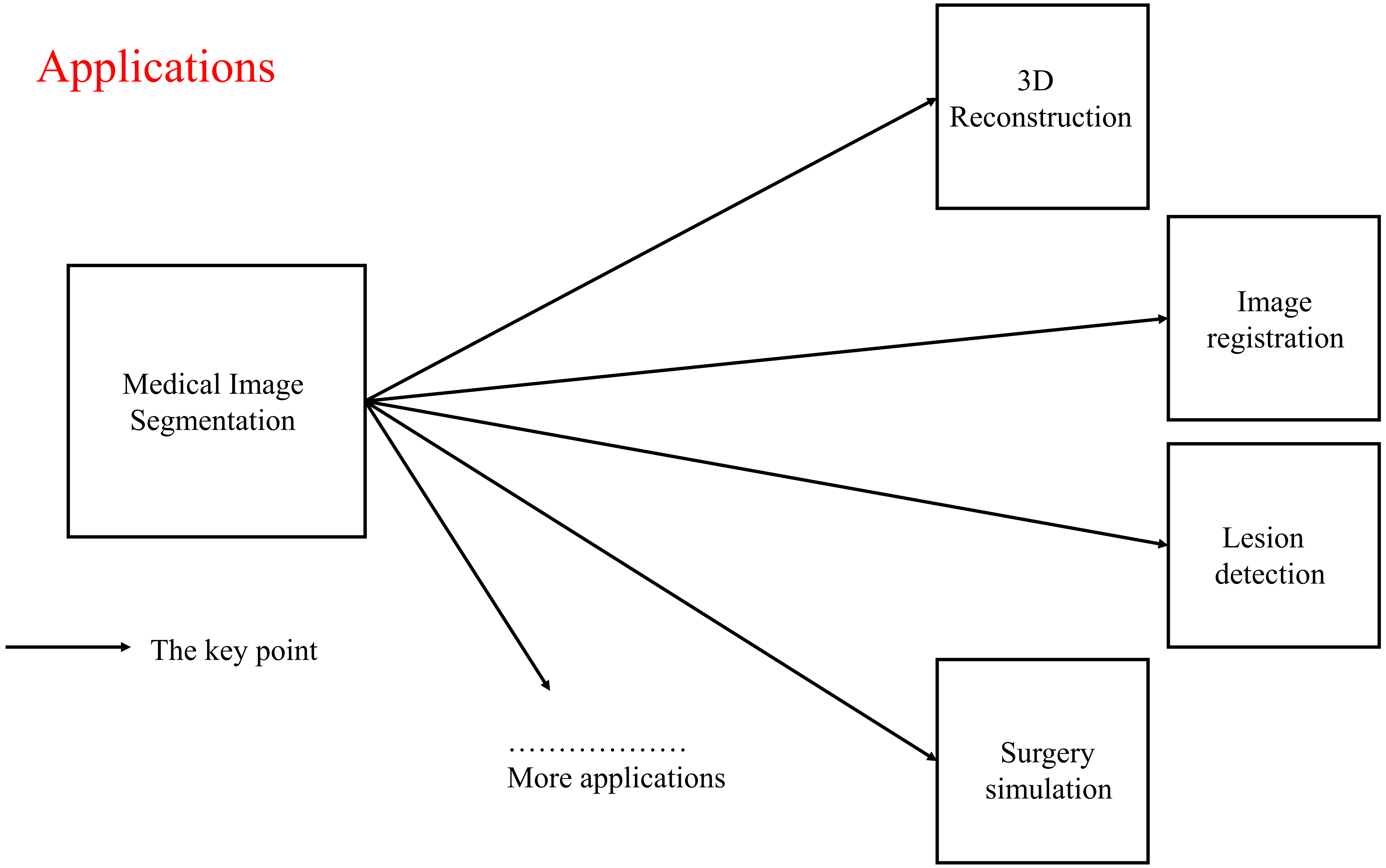
MIRD-Net for Medical Image Segmentation

Yongfeng Huang, Xueyang Li, Cairong Yan, Lihao Liu, Hao Dai

Presenter: Xueyang Li

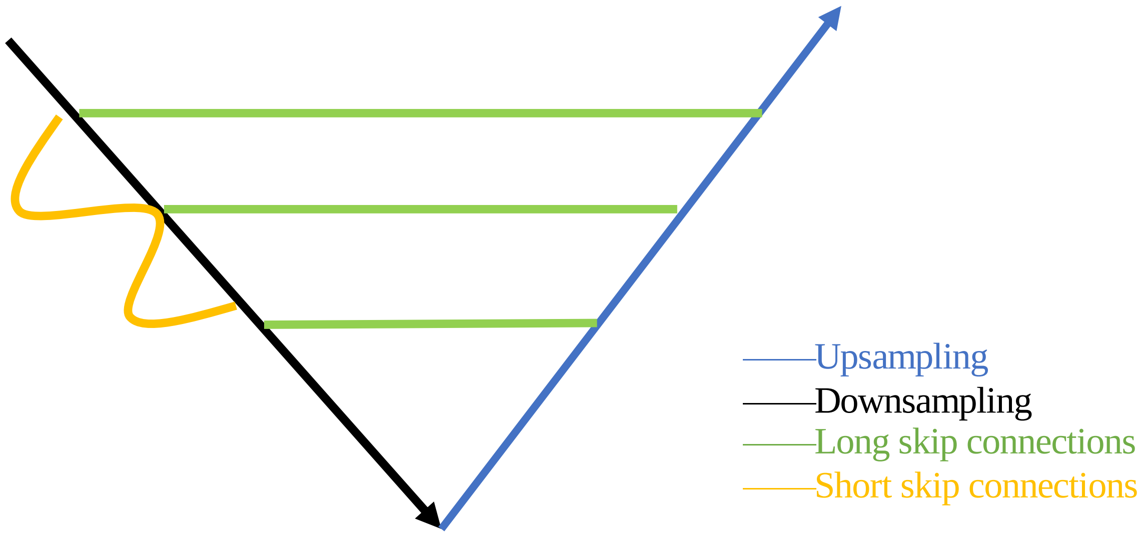
(1) Introduction

Applications



Challenges

1. Adjacent tissues in boundary has approximate pixel values.
2. the features between pixels are non-linear.
3. Distinguishing subtle features between different categories after pooling layers is still a difficult task, which affects the segmentation accuracy.



Accuracy ↑

Parameter ↓

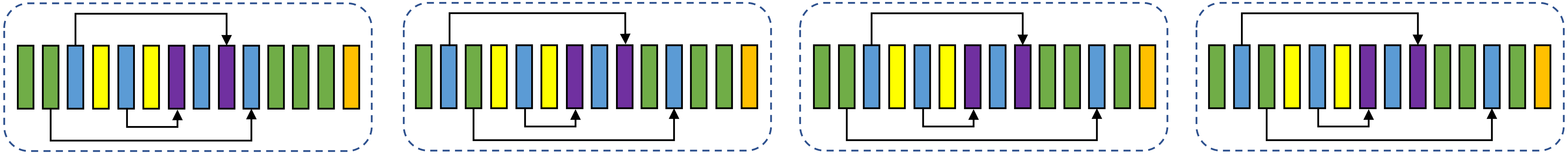
Ronneberger, O., Fisher, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer Assisted Intervention (MICCAI). pp. 234–241 (October 2015)

(2) MIRD-Net

Exploration steps

Find a shallower U-Net
(ten layers as a backbone of the network)

Add four RBs into up-down sampling path and find the best position of RBs



(a) RSU-Net(3,5,8,10) (b) RSU-Net(2,5,8,11) (c) RSU-Net(3,5,8,12) (d) RSU-Net(2,5,8,12)

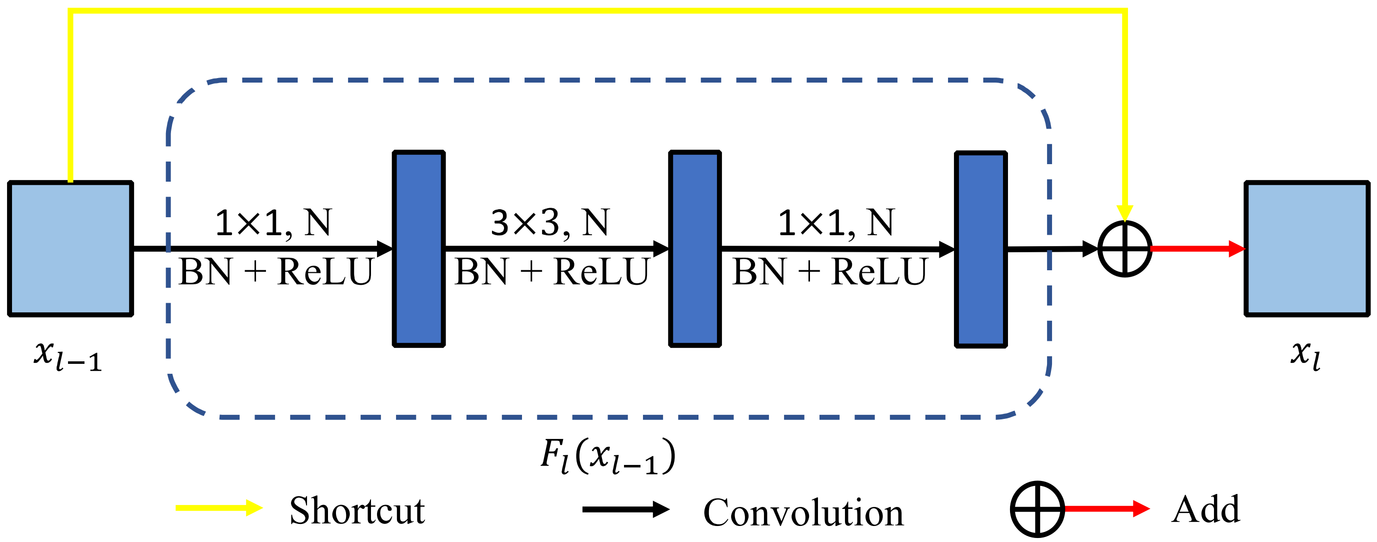
Concatenation Conv(1,1)+Sigmoid Conv(3,3)+ReLU Maxpooling(2,2) Residual Block Upsampling(2,2)

Use MIRD Blocks to replace the RBs in downsampling path
(MIRD Block is composed of DBs, RBs, and Inception)

Remove the pooling layers

MIRD Block

Residual Block It can reuse the feature from the previous layer and ease the training of deeper networks.



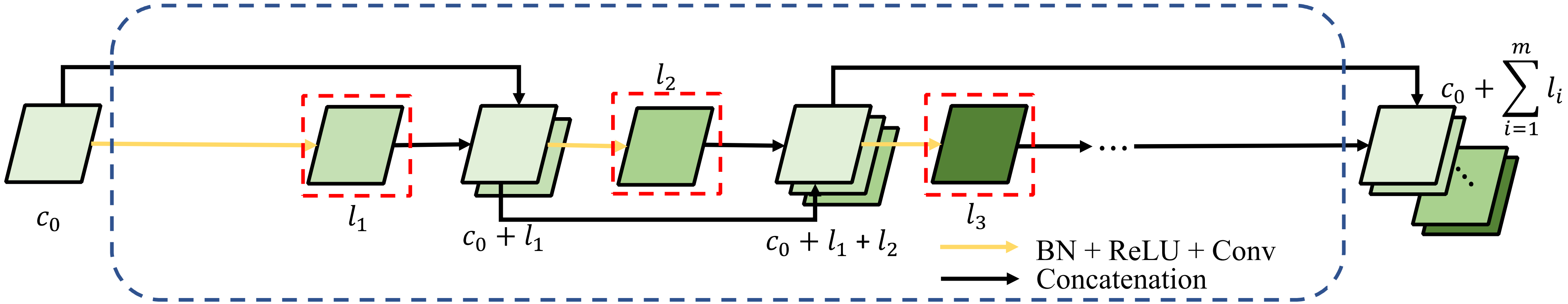
$$x_l = F_l(x_{l-1}) + x_{l-1}$$

x_l : the output of the current layer
x_{l-1}: the output of the previous layer
$F_l(\cdot)$: the non-linear calculation

He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (2016)

MIRD Block

Dense Block Each layer is connected to all previous layers through concatenation.



Huang, G., Liu, Z., v. d. Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2261–2269 (July 2017)

$$x_l = f[*(x_{l-1}, x_{l-2}, \cdots, x_0)]$$

x_l : the output of the current layer

$x_{l-1}, x_{l-2}, \cdots, x_0$: the output of all previous layers connected to x_l

$*(\cdot)$: the concatenation operation

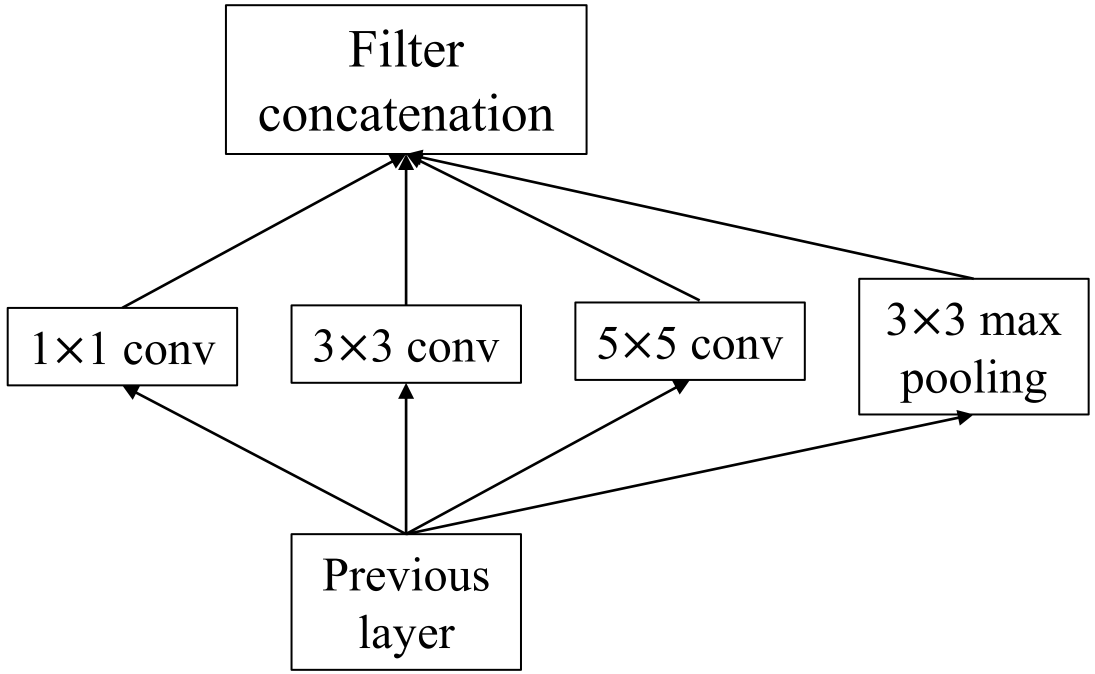
$f[\cdot]$: the non-linear calculation

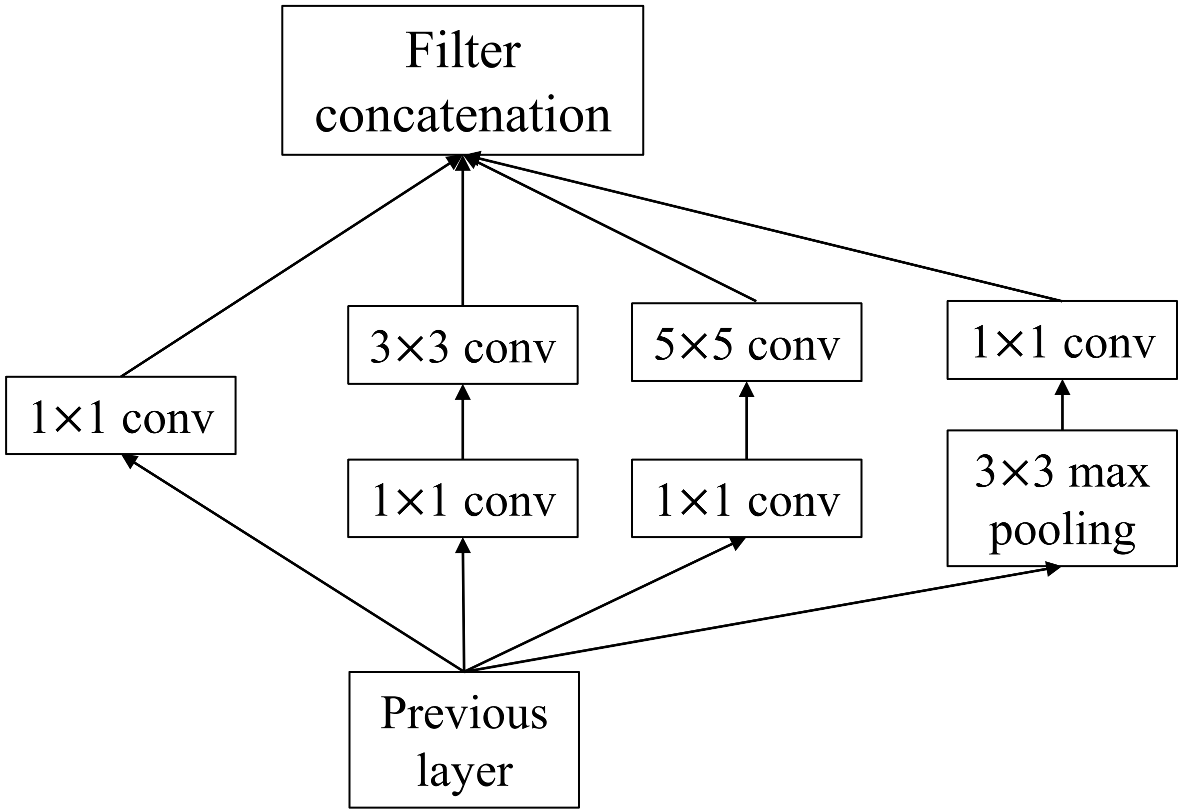
c_0 : the channel of input image

l_i : the channel of the convolved image

MIRD Block

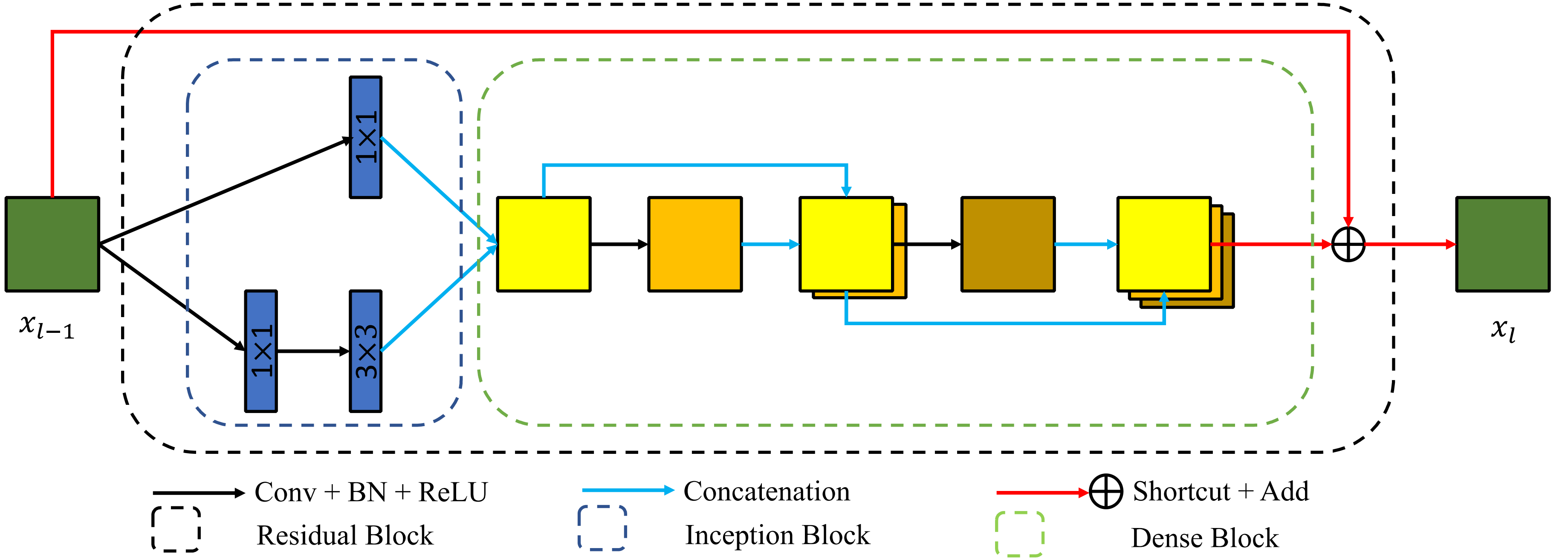
Inception The network can learn more represented features .





Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1–9 (2015)

MIRD Block Mini-Incepetion-Residual-Dense Block



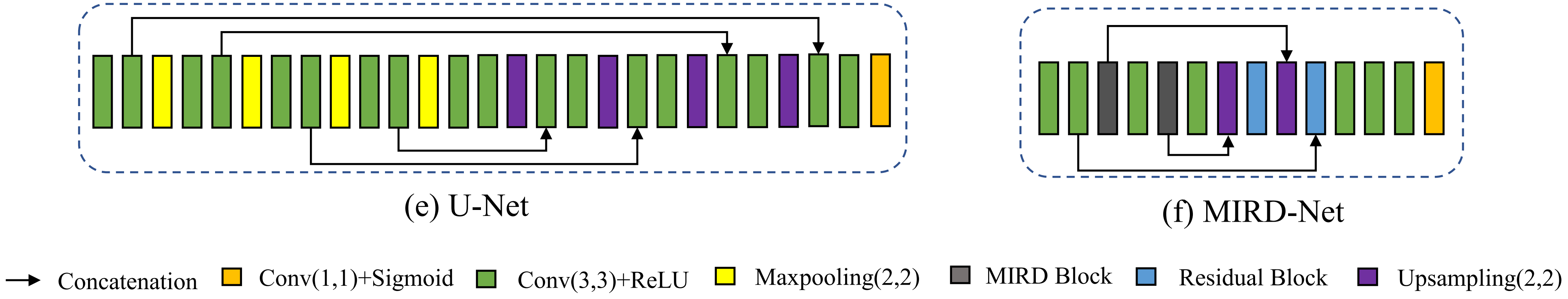
$$x_l = F\Big(H\big(G(x_{l-1})\big)\Big) + x_{l-1}$$

$G(\cdot)$: the function of Inception Block

$H(\cdot)$: the calculation in Dense Block

$F(\cdot)$: the calculation in Residual Block

Comparison



MIRD-Net has four different points from U-Net:

1. A shallower backbone is used in MIRD-Net, aiming to keep parameters low.
2. MIRD-Net has no pooling layers, such a design avoids loss of information during forward propagation.
3. MIRD-Net is designed with MIRD Blocks, which makes the network learn more represented features.
4. Two Residual Blocks (RB) are embedded in the upsampling path.

(3) Experiments & Results

Experimental Platform and Datasets

Platform

Windows 10(1801)
Intel(R) Core (TM) i7-7700 CPU @ 3.60GHz
Nvidia GeForce GTX 1080 Ti, 16GB RAM
Samsung SSD 850 EVO 500GB

Datasets

Data name	Source	Image Size	Modality
Retinal Vessel	grand-challenge.org	$512 \times 512 \times 3$	non-mydriatic camera
Cells	ISBI 2012	512×512	microscopy
Nuclei	Data Science bowl 2018	360×360	microscopy
Lung	Kaggle	512×512	CT
Cervical Cytology	grand-challenge.org	512×512	microscopy
Skin Lesion	ISIC 2017	$512 \times 512 \times 3$	dermoscopy

Evaluation Metrics

MDice

(Mean Dice coefficient)

$$MDice = \frac{1}{r \times m} \sum_{i=1}^{r} \sum_{j=1}^{m} \frac{2|A_{ij} \cap B_{ij}|}{|A_{ij}| + |B_{ij}|}$$

StdDice

(Std of Dice coefficient)

$$StdDice = \sqrt{\frac{\sum_{i=1}^{r}\left(\frac{1}{m}\sum_{j=1}^{m} \frac{2|A_{ij} \cap B_{ij}|}{|A_{ij}| + |B_{ij}|} - MDice\right)^2}{r-1}}$$

A_{ij}: the predicted image

B_{ij}: the ground truth corresponding to A_{ij}, 0-1 matrix

m : the number of images in one subset

r : the fold used in cross-validation

We split each dataset into five subsets (F1-F5) equally and run a 5-fold cross-validation.

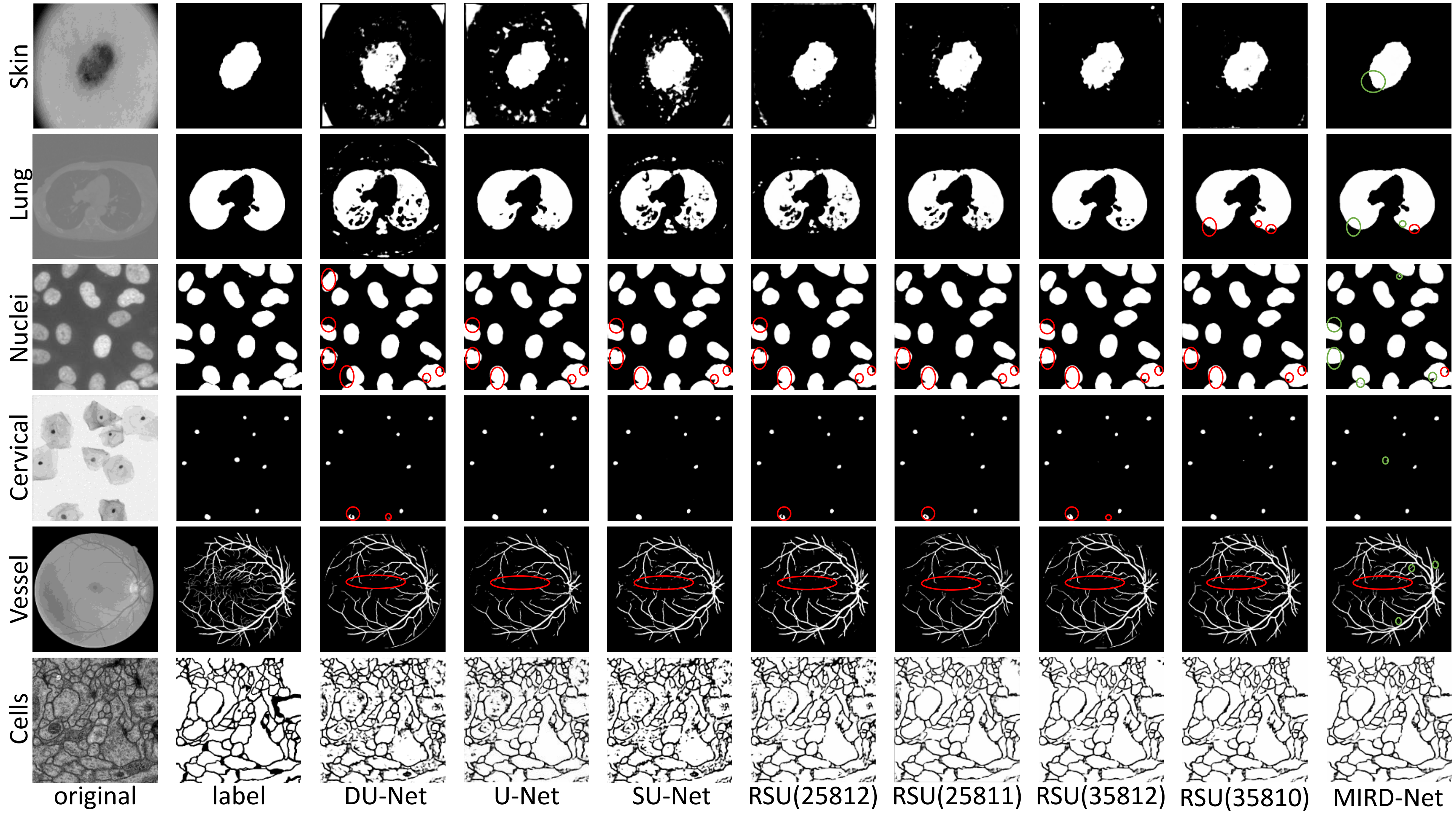
Results

The parameters of each network ($\times 10^6$)

DU-Net	U-Net	SU-Net	RSU(25812)	RSU(25811)	RSU(35812)	RSU(35810)	MIDR-Net
40.61	31.03	25.42	0.47	0.51	0.47	0.51	0.59

The parameters of MIRD-Net are only about 1/50 of U-net.

Results



Subtle differences:

: incomplete correct mask which is compared to the label

: the results of MIRD-Net which show the better performance

Results

Dice coefficient

Models	Skin		Lung		Nuclei		Cervical		Vessel		Cells	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
DU-Net	0.455	0.063	0.632	0.069	0.851	0.085	0.844	0.025	0.651	0.105	0.847	0.032
U-Net	0.514	0.044	0.748	0.087	0.897	0.102	0.886	0.077	0.698	0.087	0.869	0.115
SU-Net	0.469	0.085	0.667	0.041	0.902	0.068	0.882	0.058	0.721	0.098	0.866	0.076
RSU(25812)	0.603	0.095	0.709	0.077	0.896	0.064	0.855	0.133	0.715	0.105	0.892	0.034
RSU(25811)	0.617	0.117	0.731	0.124	0.909	0.125	0.864	0.058	0.714	0.094	0.887	0.072
RSU(35812)	0.624	0.047	0.772	0.076	0.911	0.118	0.897	0.029	0.730	0.098	0.895	0.046
RSU(35810)	0.656	0.126	0.794	0.089	0.928	0.097	0.902	0.047	0.742	0.086	0.902	0.021
MIRD-Net	0.742	0.076	0.810	0.067	0.954	0.075	0.925	0.017	0.765	0.047	0.919	0.029

1. SU-Net (shallower U-Net) has better performance on Nuclei and Vessel than U-Net and DU-Net (deeper U-Net).
2. The Residual Blocks in different positions of up-downsampling path can affect the performance of the network. Best: RSU(35810)
3. MIRD-Net achieves elegant results.

Reasons for improvement

- There are no pooling layers in MIRD-Net, such a design helps alleviate loss of information during forward propagation.

- The different kernels (1 × 1 and 3 × 3) used in MIRD-Block can make the network obtain large-structure information and tiny-structure information simultaneously.

- MIRD-Net not only use the standard skip connections used in U-Net but also reuse the feature from previous layer in MIRD-Block, which results in more represented features learned by the network.

- The connections used in MIRD Block can alleviate the gradient vanishing during the training period.

(4) Contributions

Contributions

- A shallower backbone is proposed to decrease the number of parameters.
- MIRD Block can be simply added to other backbones as a functional module.
- Simple and flexible implementation of our proposed network architecture.
- Great performance for challenging medical image segmentation tasks.

Thanks !!!