上海科技大学
ShanghaiTech University

Multi-Cell Multi-Task Convolutional Neural Networks for Diabetic Retinopathy Grading

Kang Zhou, Zaiwang Gu, Wen Liu, Weixin Luo,
Jun Cheng, Shenghua Gao, Jiang Liu

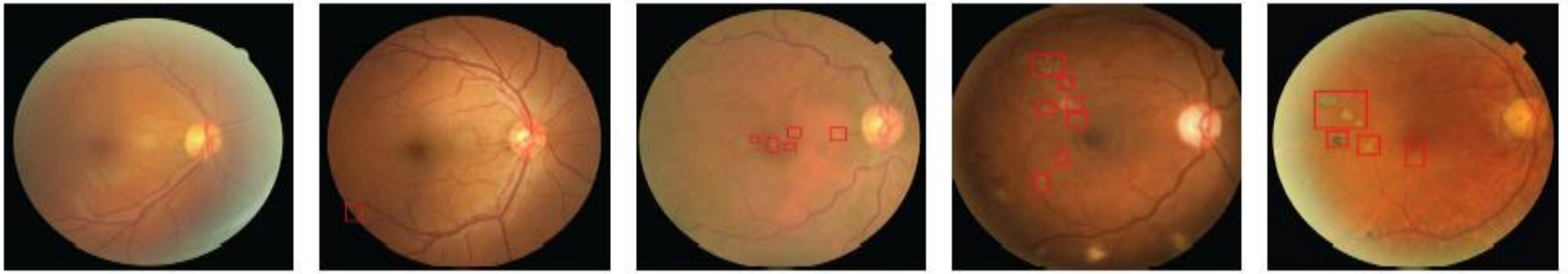
EMBC 2018, Honolulu, USA
Jul. 20, 2018

Contents

Background

Background

Diabetic Retinopathy Grading:

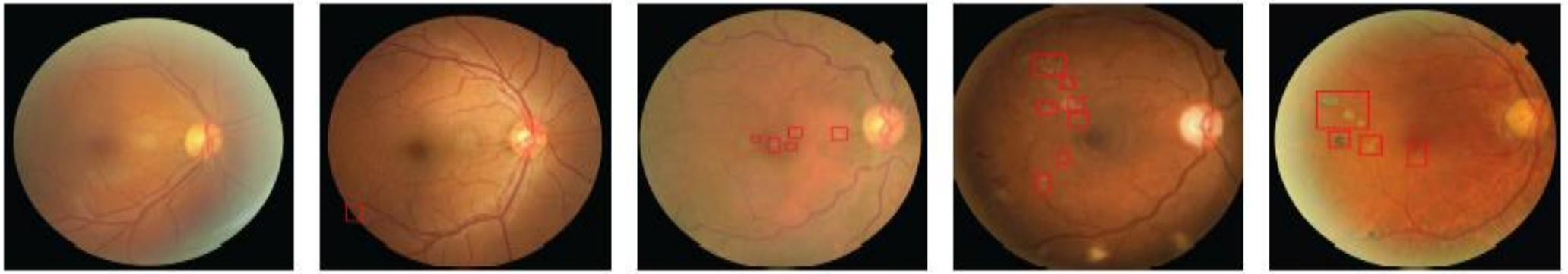


(a) Grade 0 (b) Grade 1 (c) Grade 2 (d) Grade 3 (e) Grade 4

- **Problem:**
  - Label: 0, 1, 2, 3, 4
  - Larger number means the severity of the disease becomes more significant
- **Task:**
  - Input: Image
  - ***Output***: Its grade

Background

Diabetic Retinopathy Grading：



(a) Grade 0　(b) Grade 1　(c) Grade 2　(d) Grade 3　(e) Grade 4

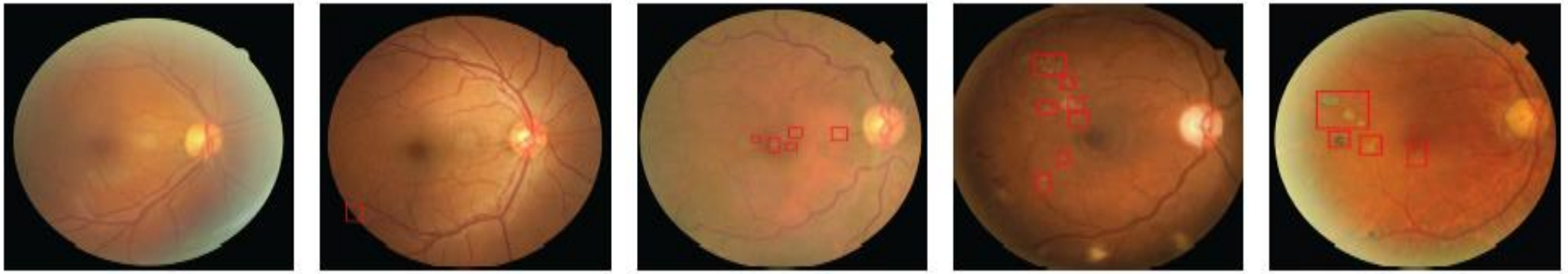
- **Challenge**

  (DR grading ≠ general image classification)：

  - The classes in DR grading are ***correlative*** while in general image classification are not
  - The image resolution of DR images is significantly higher than that of general images

Background

Diabetic Retinopathy Grading：



(a) Grade 0　(b) Grade 1　(c) Grade 2　(d) Grade 3　(e) Grade 4

- **Challenge**

  (DR grading ≠ general image classification)：

  - The classes in DR grading are correlative while in general image classification are not
  - The image resolution of DR images is ***significantly higher*** than that of general images

Background

Diabetic Retinopathy Grading：

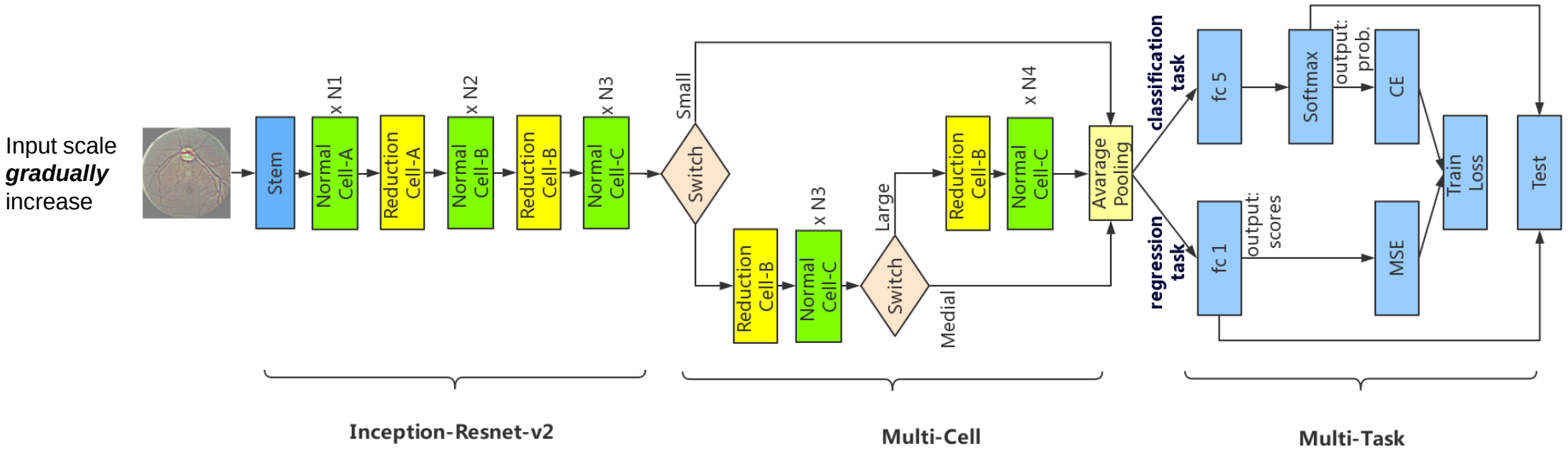
- **Contribution ：**
  - We propose a **Multi-Task Learning** strategy to simultaneously improves the classification accuracy and discrepancy between ground-truth and predicted label.
  - We propose a **Multi-Cell CNN architecture** which not only accelerates the training procedure, but also improves the classification accuracy.
  - Experimental results validate the effectiveness of our method. Further, our solution can be readily integrated with many other existing CNN based DR image diagnosis and other disease diagnosis.

2

Proposed Method: M²CNN

Proposed Method: M²CNN

Multi-Cell Multi-Task Convolutional Neural Networks:



- **Overall :**
  - The overall network architecture of our M²CNN
  - Inception-Resnet-v2 is proposed in [1]

[1] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In AAAI, 2017.

Proposed Method: M²CNN

Multi-Cell Multi-Task Convolutional Neural Networks:

- **Multi-Task Learning :**
  - Softmax loss doesn't consider the relationships of DR images with different stages:

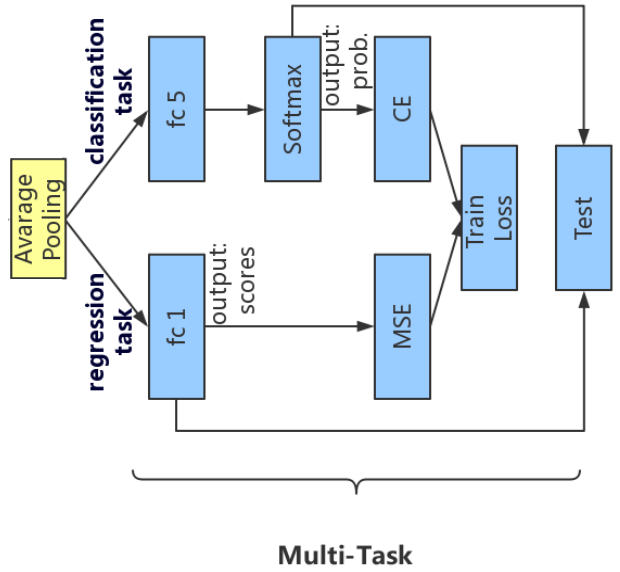
$$L_1 = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k} 1\{y^{(i)} = j\}\log(\text{Prob}_{ij})\right]$$

  - Mean Square Error (MSE) loss is not robust for classification task:

$$L_2 = \frac{1}{m}\sum_{i=1}^{m}(y - y^{(i)})^2$$

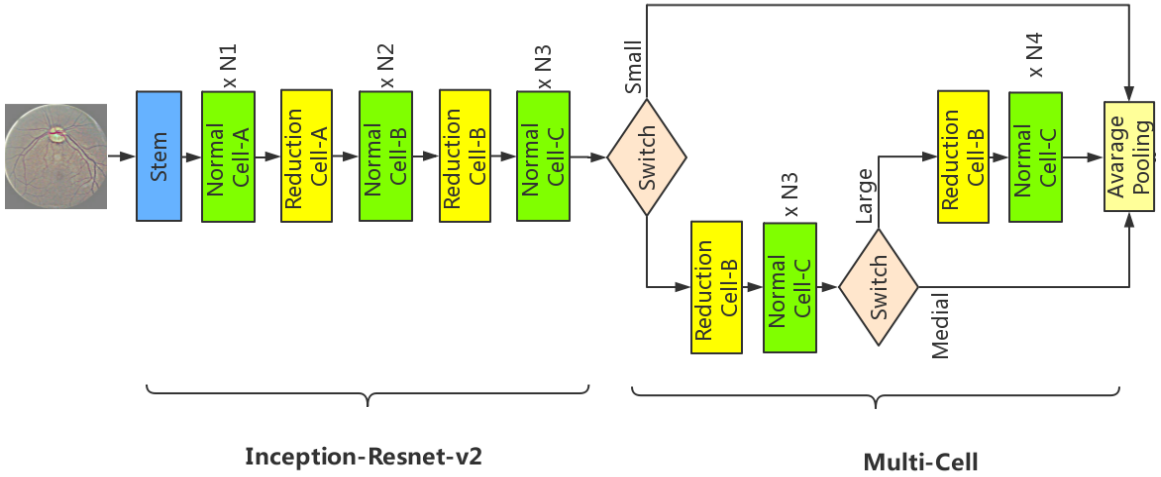
  - Multi-task loss:

$$L = L_1 + L_2$$



Proposed Method: M²CNN

Multi-Cell Multi-Task Convolutional Neural Networks：



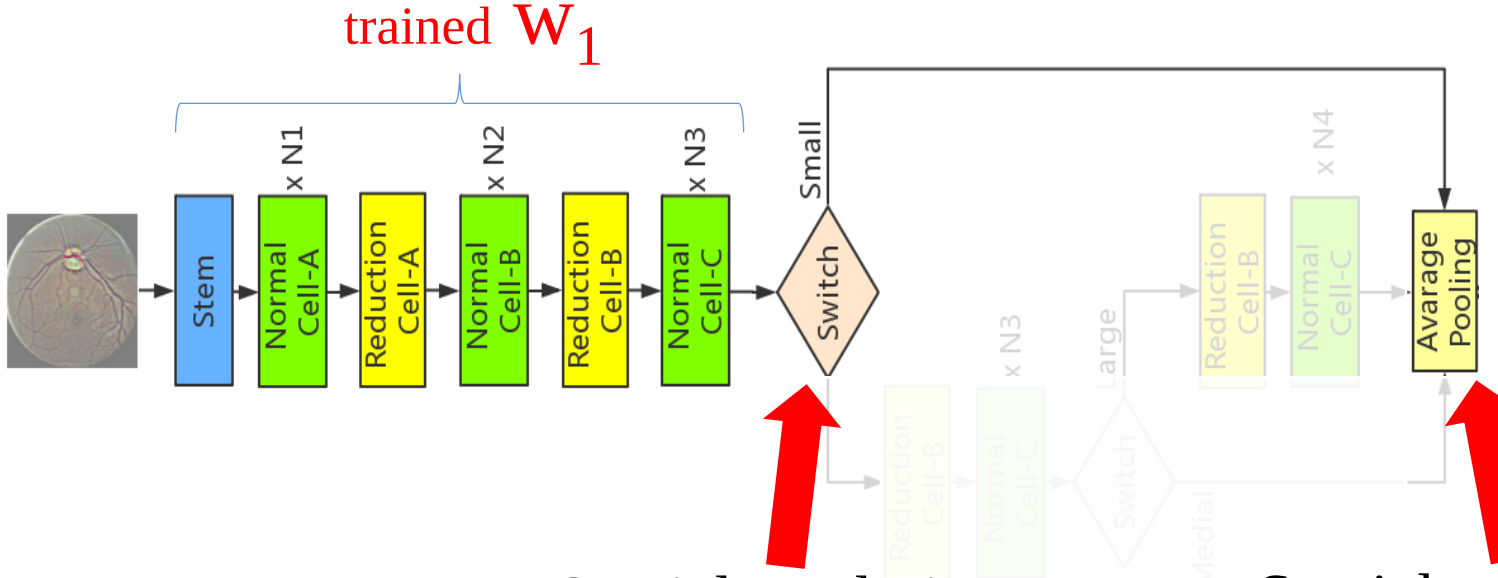
input image	224×224	448×448	720×720
before switch	5×5	12×12	21×21
after multi-cell	5×5	5×5	4×4

The spatial resolution of input image and some feature map

- **Multi-Cell Architecture ：**
  - ***Small resolution*** image often leads to information loss especially when the lesion is small
  - ***Large resolution*** image will introduce more computational costs and lead to the gradient vanishing/exploding problem in optimization
  - **Note:** Multi-Cell Architecture ***gradually increase*** the depth of network architecture and the resolution of images
  - **Note:** The architecture of ***Normal Cell-C*** and ***Reduction Cell-B*** in Multi-Cell and Inception-Resnet-v2 are same.

Proposed Method: M²CNN

Process of Multi-Cell Architecture : **1-st training stage**



Input scale:
224 x 224

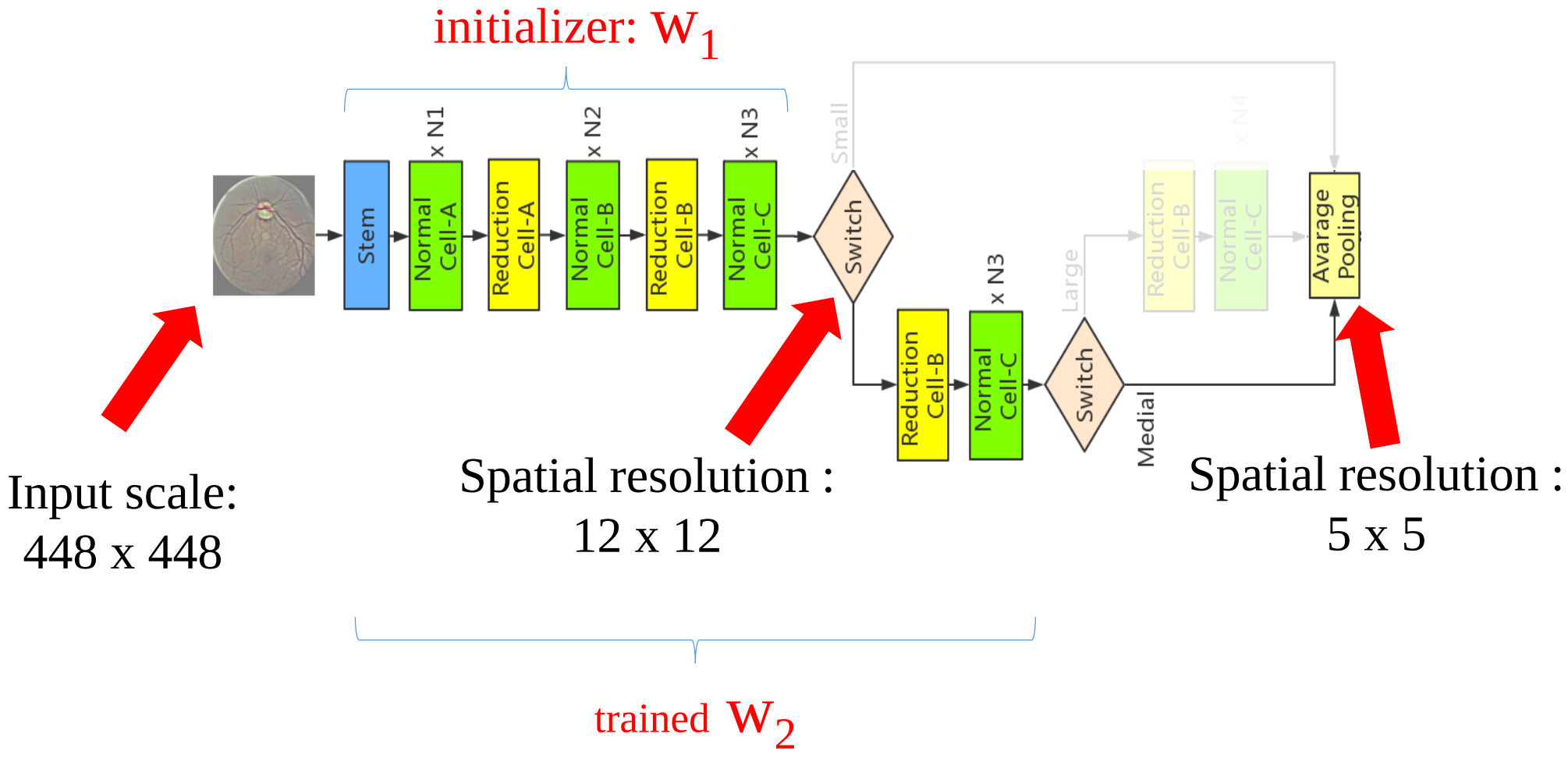
Spatial resolution :
5 x 5

Spatial resolution :
5 x 5

Depth of network architecture and the ***scale*** of images are ***gradually increased.***

Proposed Method: M²CNN

Process of Multi-Cell Architecture : **2-ed training stage**

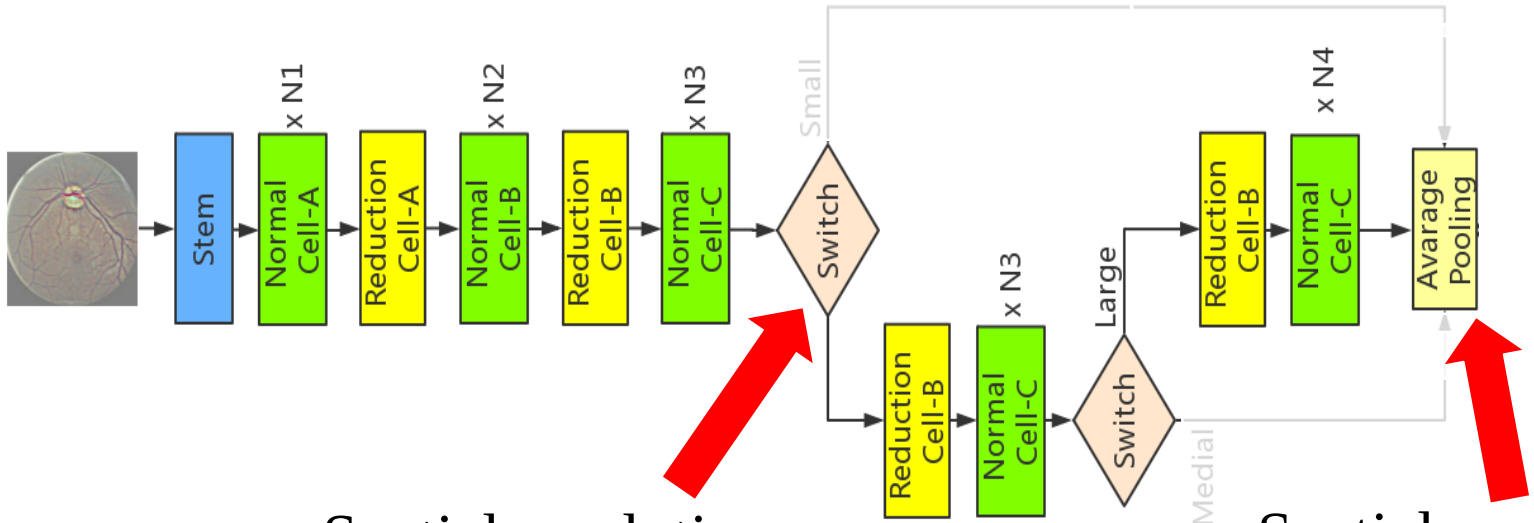


Depth of network architecture and the ***scale*** of images are ***gradually increased.***

Proposed Method: M²CNN

Process of Multi-Cell Architecture : **3-rd training stage**

trained: W_3 (***Training's finished***!!!)



Input scale:
720 x 720

Spatial resolution :
21 x 21

Spatial resolution :
4 x 4

initializer: W_2

Depth of network architecture and the ***scale*** of images are ***gradually increased.***

3

Experiment

Experiment

Experimental Setup

- **Dataset:**
  Kaggle organized a comprehensive competition in order to design an automated retinal image diagnosis system for DR screening in 2015 [2].
- **Evaluation Metric:**
  We use the ***quadratic weighted kappa*** to evaluate our proposed methods, which is used in Kaggle DR Challenge.

[2] Diabetic retinopathy detection. https://www.kaggle.com/c/diabetic-retinopathy-detection/data

Experiment

Evaluation of Different Modules

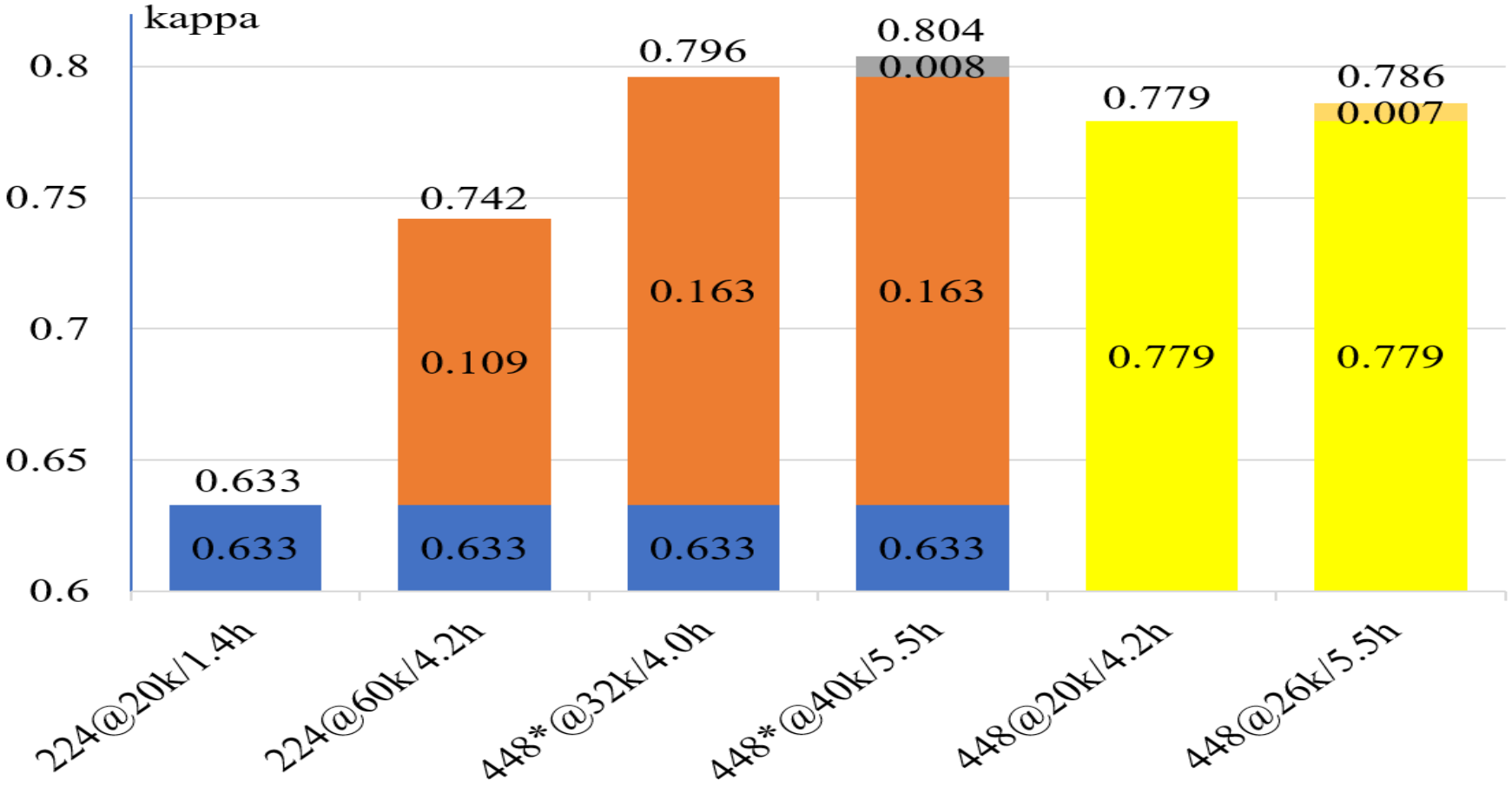
- **Multi-Task Learning Module**
- **Multi-Cell Architecture Module**

Train	MSE	CE	Multi-Task		M^2CNN	
Test	scores	prob.	scores	prob.	scores	prob.
224×224	0.720	0.725	0.742	0.718	-	-
448×448	0.790	0.772	0.812	0.782	**0.830**	0.812
720×720	0.835	0.751	**0.841**	0.826	**0.844**	0.842

Experiment

Evaluation of Different Modules

- Multi-Task Learning Module
- **Multi-Cell Architecture Module**



Experiment

Performance Comparison

Algorithm	val set	test set
Min-pooling	0.860	0.849
Zoom-in-Net [11]	0.857	0.849
o_O	0.854	0.844
Reformed Gamblers	0.851	0.839
M-Net+A-Net [11]	0.837	0.832
BaseNet	0.835	0.828
BaseNet+MT	0.841	0.838
M^2CNN	0.844	0.841

[11] Z. Wang, Y. Yin, J. Shi, W. Fang, H. Li, and X. Wang. Zoom-in-net: Deep mining lesions for diabetic etinopathy detection. InMICCAI, 2017

Thanks

Q & A