Context-Aware Cascade Network for Semantic Labeling in VHR Image

Yongcheng Liu, Bin Fan, Lingfeng Wang, Jun Bai

Shiming Xiang, Chunhong Pan

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences

Outline

1 Introduction

2 Related work

3 CAC-NET

4 Future work

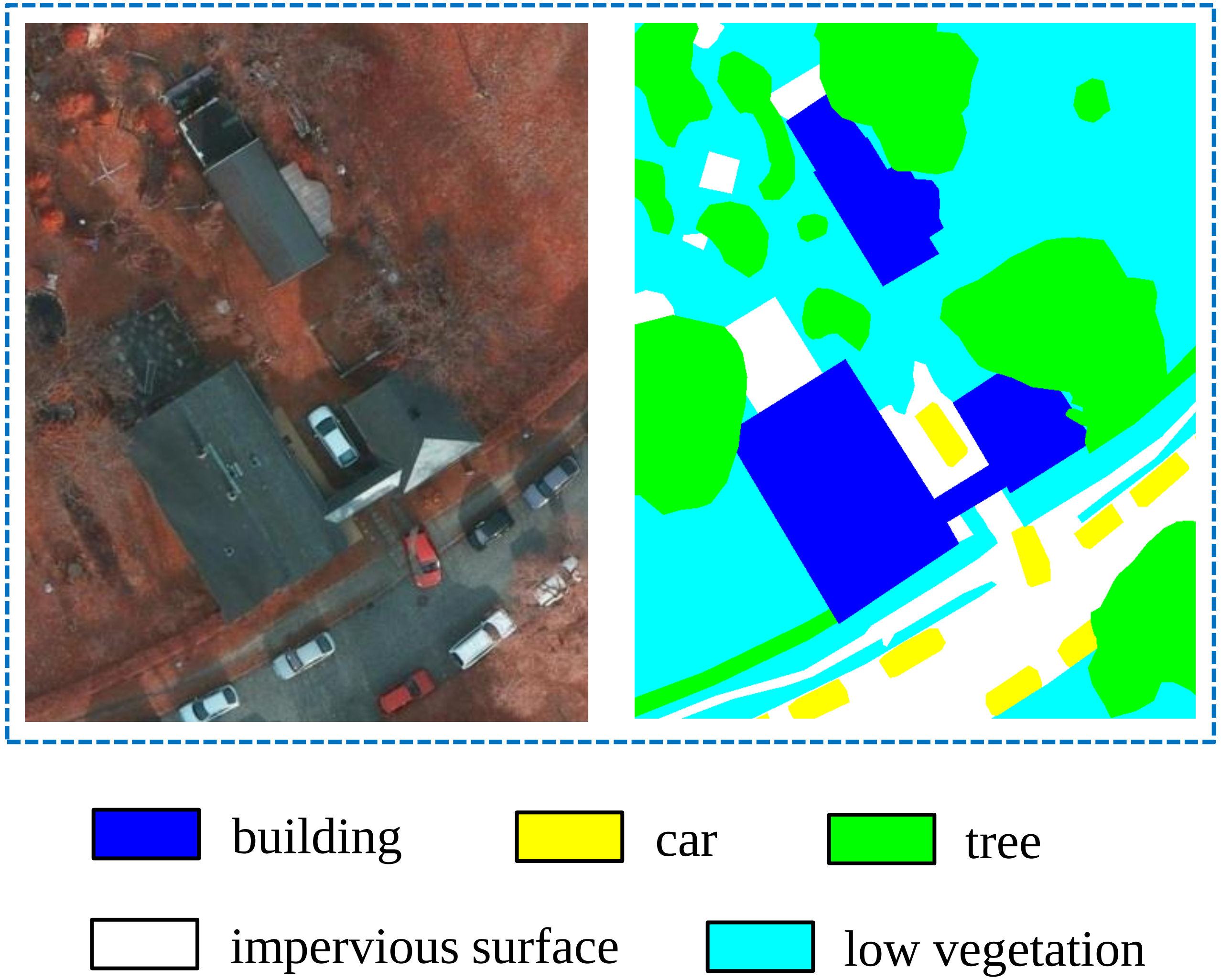
Introduction

Semantic labeling :

Assign each pixel in a given image to a semantic object class

Important application :

- Infrastructure planning
- Urban change detection
- Disaster exploration ……



Introduction

Challenges :

(1) Complex man-made objects

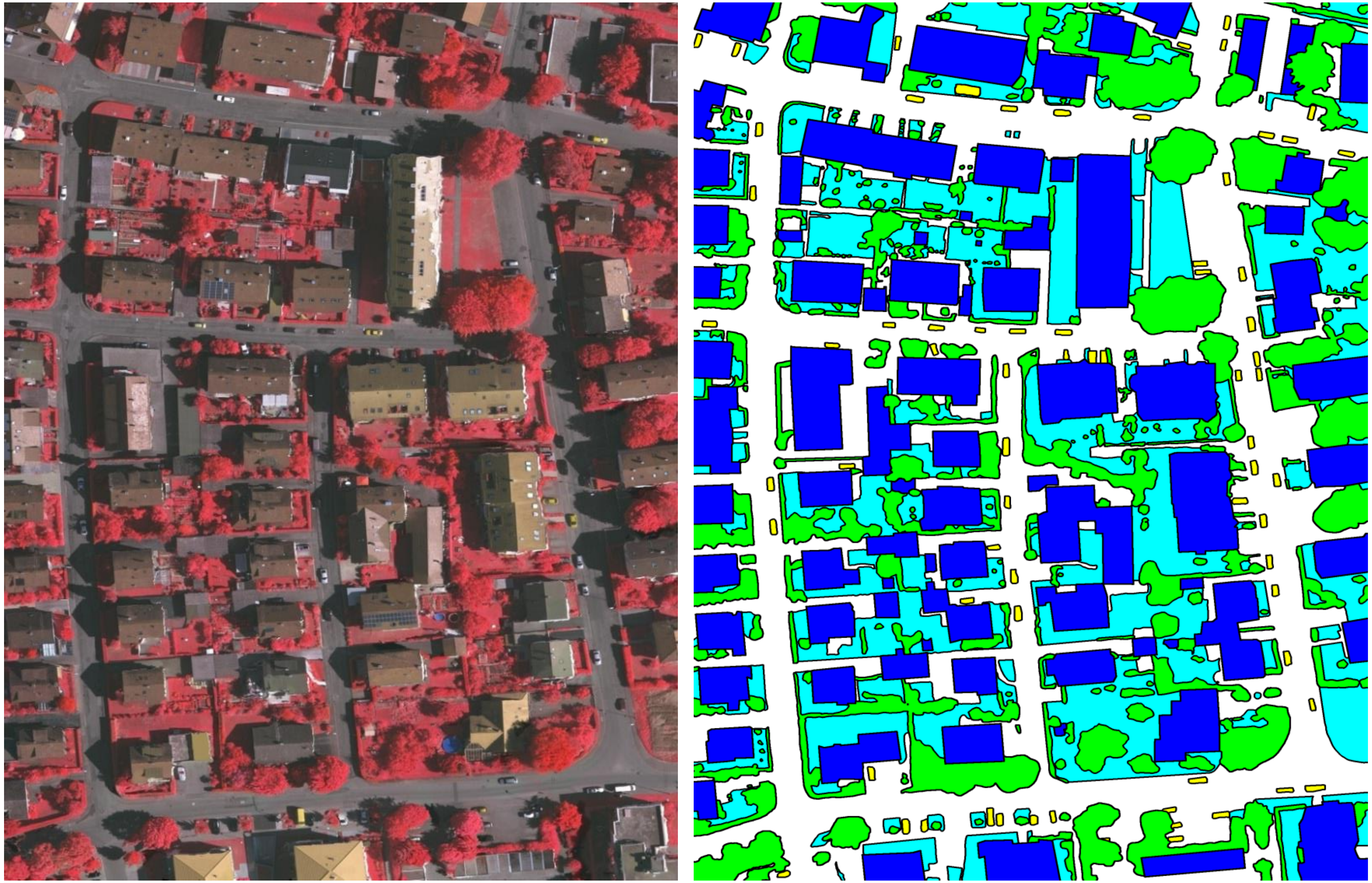
- high intra-class variance
- low inter-class variance

(2) Fine-structured objects

- small or threadlike
- locate together
- occlusions and cast shadows

(3) Additional challenge

- different solutions

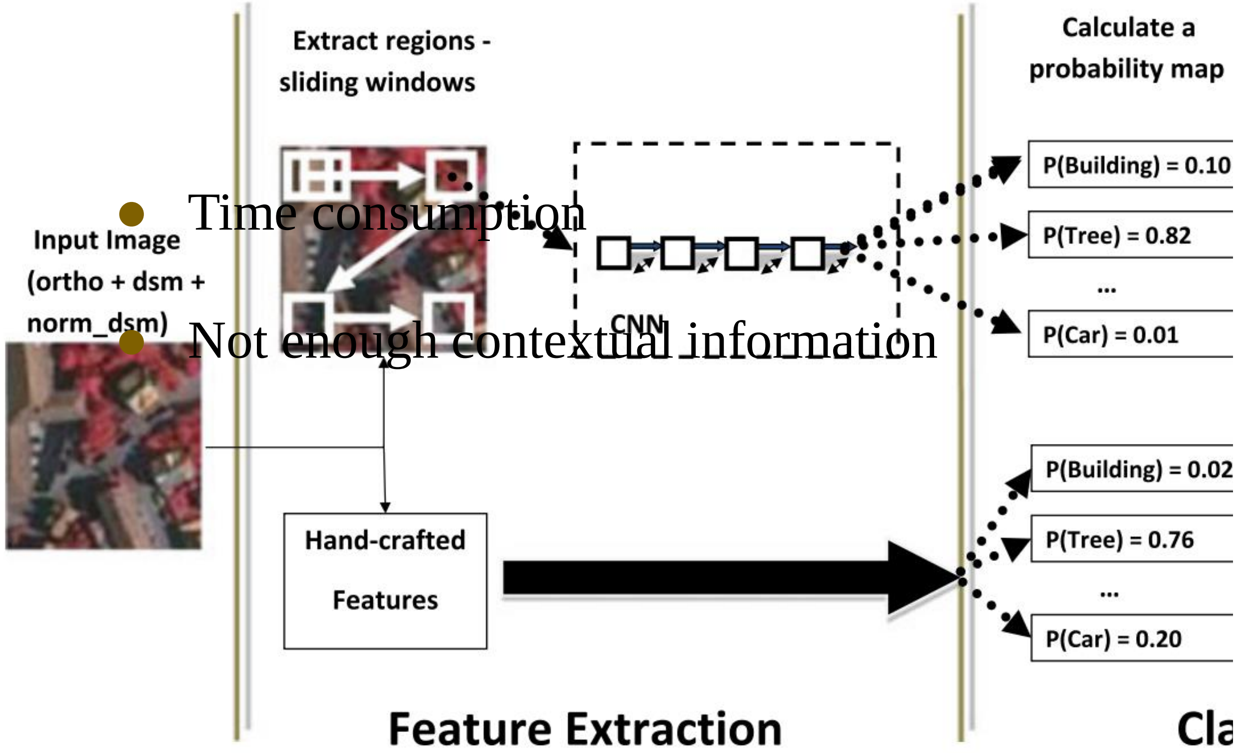


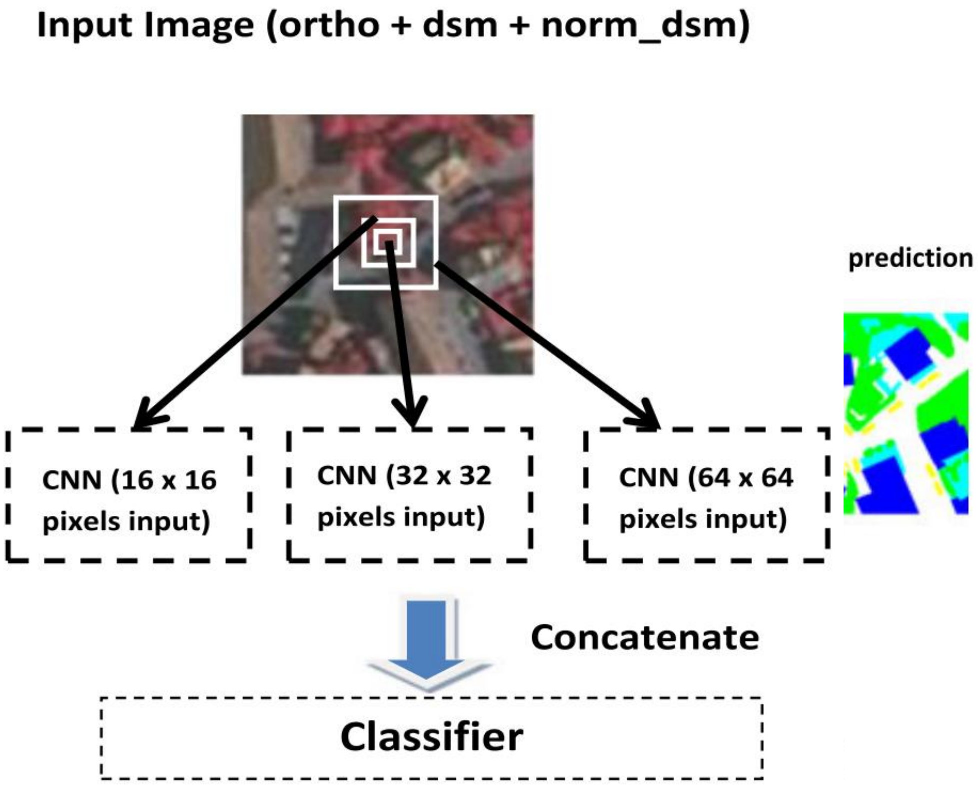
Related work

1) Patch-based methods

2016, Paisitkriangkrai et al.

- Time consumption
- Not enough contextual information

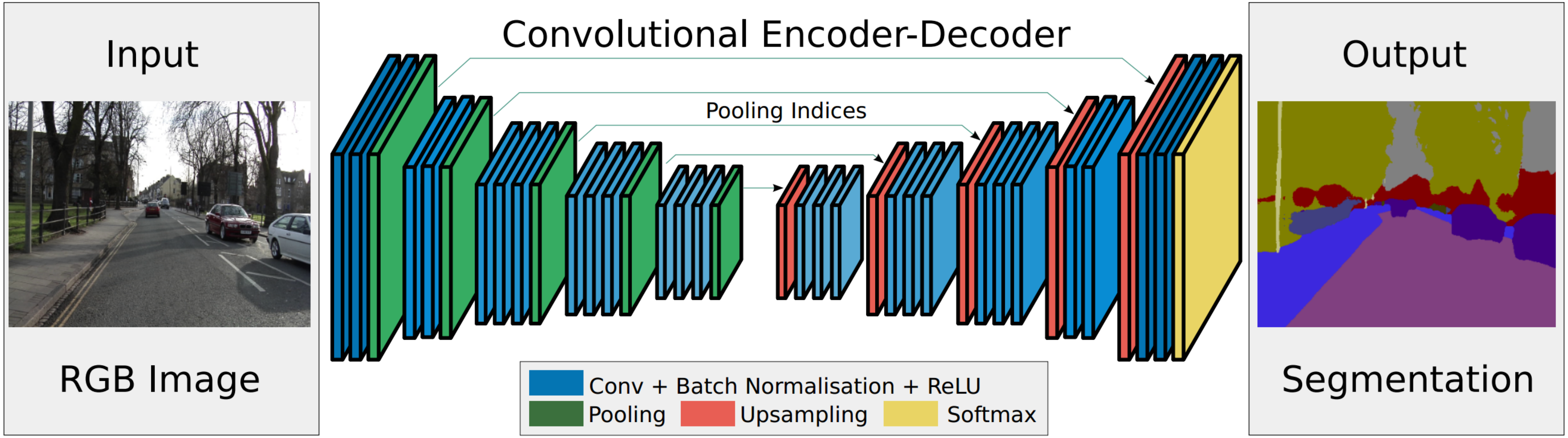




Related work

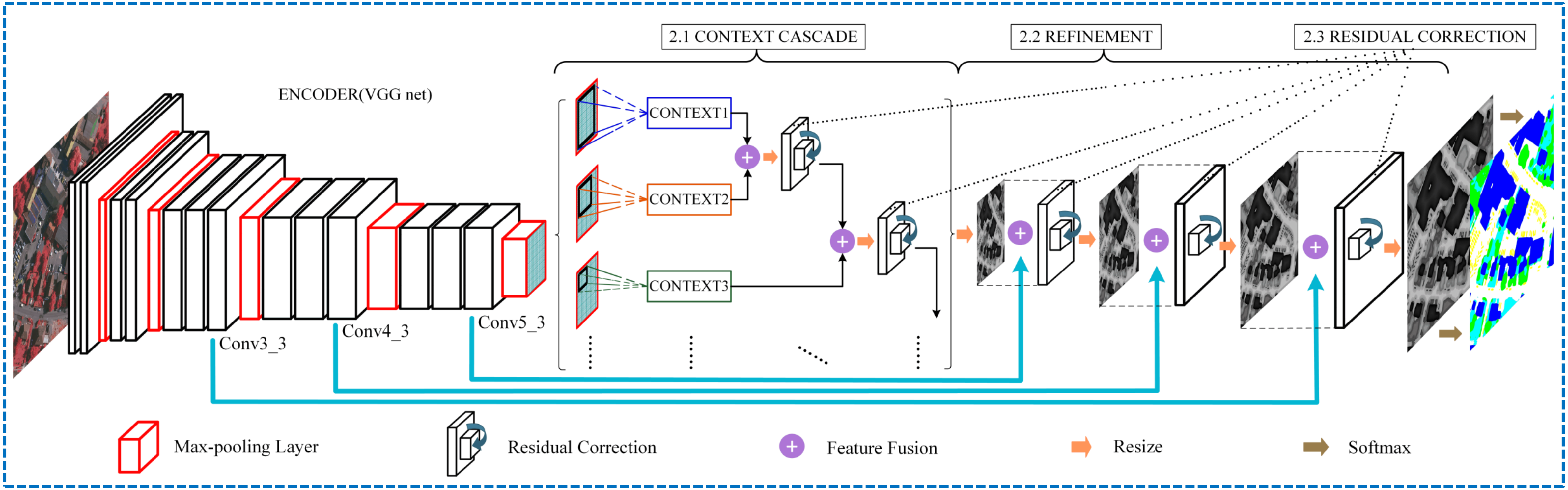
2) Fully convolutional methods

2015, Badrinarayanan et al. (Segnet)



CAC-NET

Context-Aware Cascade Network



- ✓ **Encoder** : extract features of different levels
- ✓ **Context Cascade**: capture contextual information for complex objects
- ✓ **Refinement**: refine the coarse labeling of fine-structured objects
- ✓ **Residual correction**: improve the fusion of different-level features

CAC-NET: context cascade

Contextual information

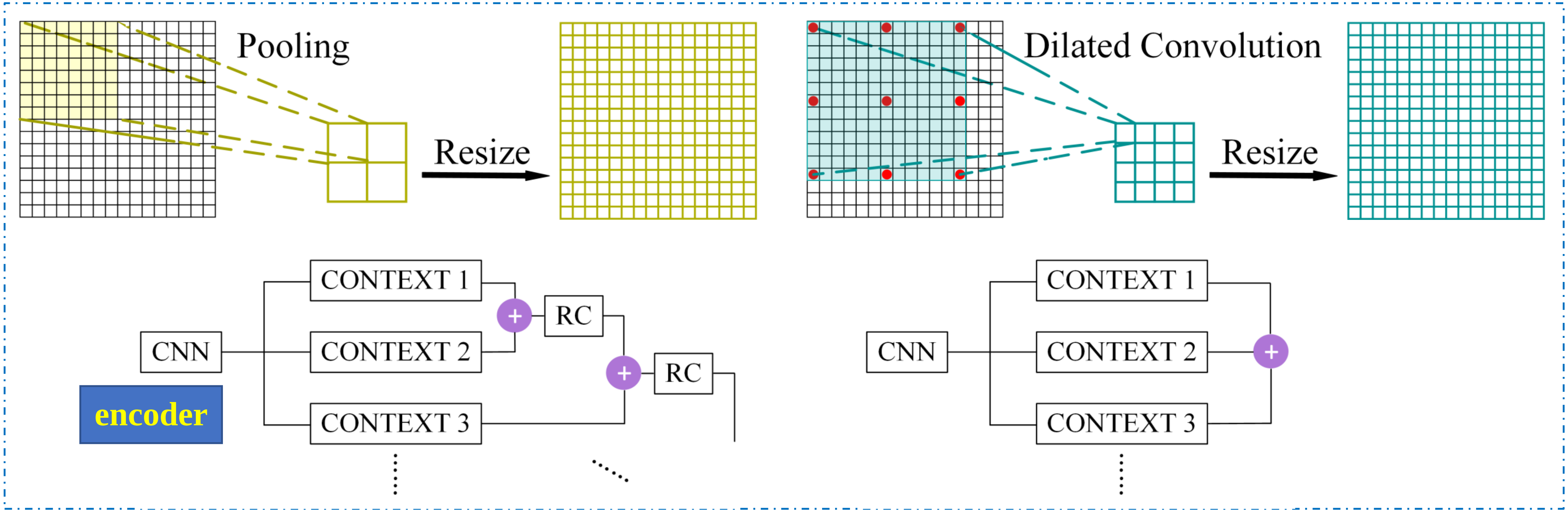
Latent dependencies between an object and its surroundings.

How did we do ?

- Multi-scale images
- Multi-size convolutional kernel

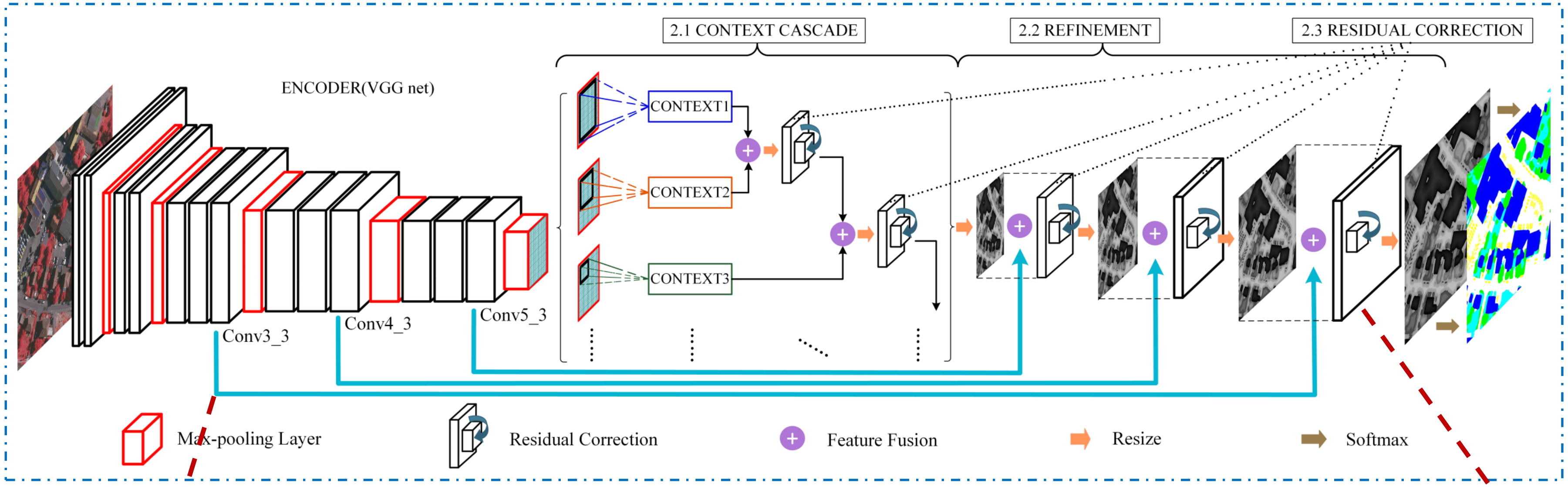
EFFECTIVE and EFFICIENT ?

CAC-NET: context cascade



- ✓ **Context capturing** : multi-kernel pooling and dilated convolution
- ✓ **Context aggregating**: **from global to local** in a sequentially cascaded manner
- ✓ **Residual correction**: improve the fusion of **different-level context**

CAC-NET: refinement



Local details

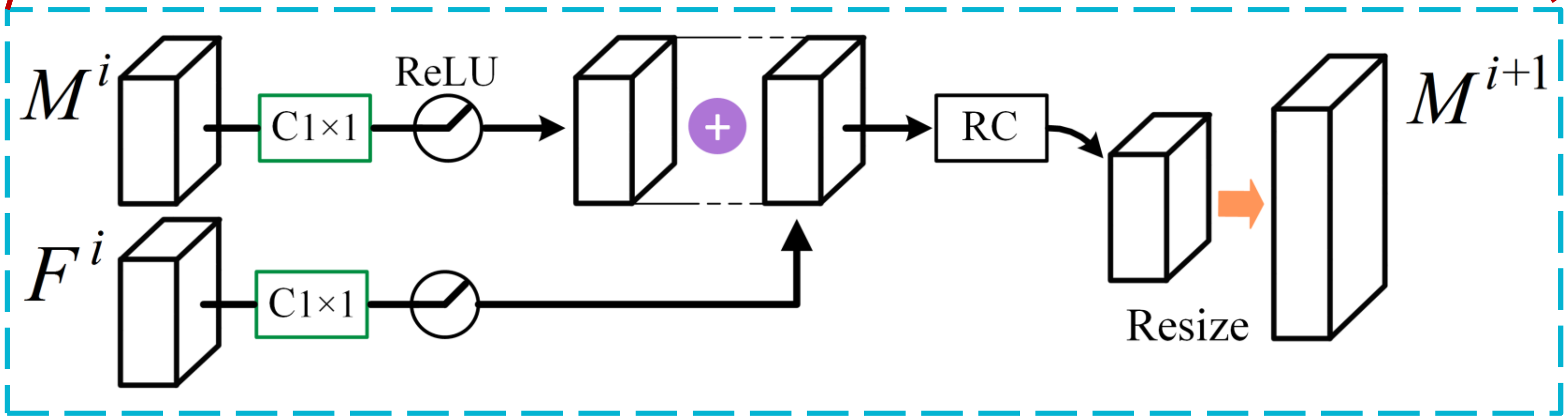
long-span connection

progressively introduced

Residual correction

improve different-level

features fusion



CAC-NET: residual correction

different-level context

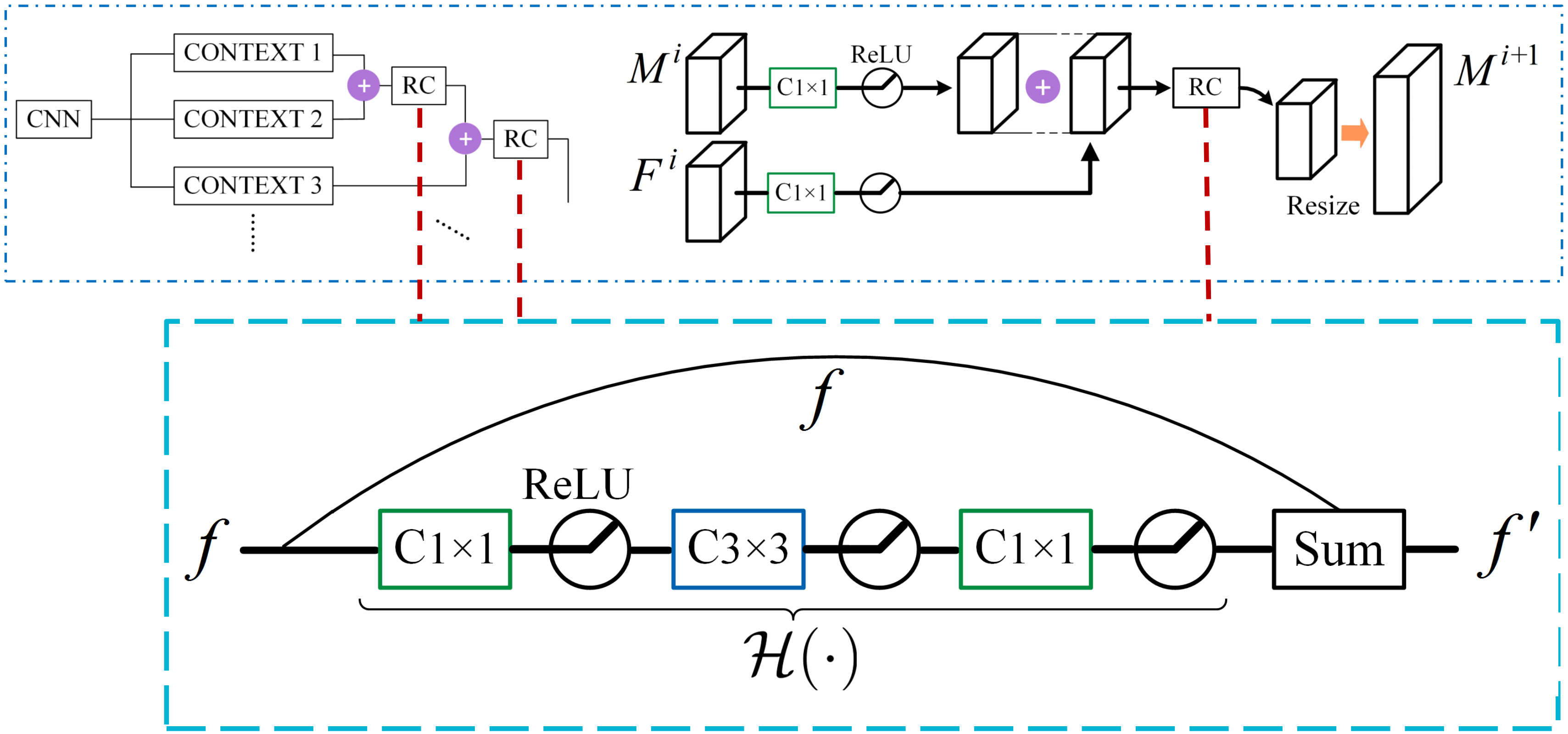
different-level features

semantic gap

sum fusion: information loss

remedy

residual correction scheme



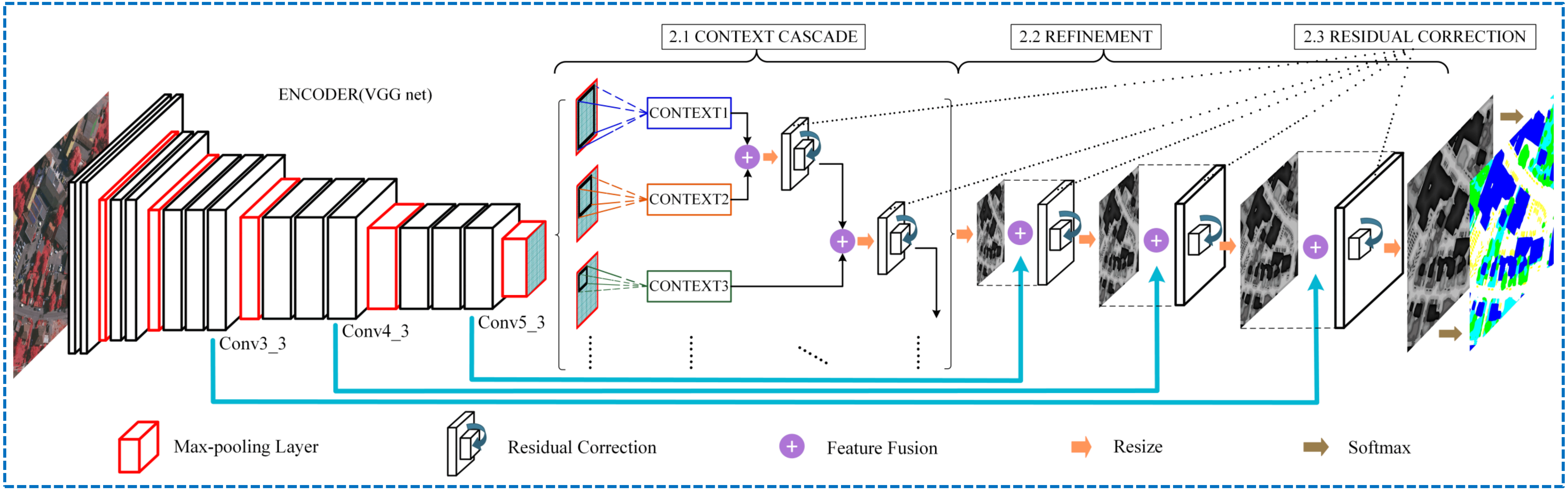
f : fused features

f' : underlying desired fusion

$$\mathcal{H}(\cdot) = f' - f \longrightarrow f' = f + \mathcal{H}(\cdot)$$

CAC-NET

Context-Aware Cascade Network



global to local

coarse to fine

end-to-end

CAC-NET: experiment

Dataset:

ISPRS Vaihingen 2D semantic labeling Challenge

Image: IRRG (infrared red green) ✓ ONLY

Elevation data: DSM (digital surface model)

NDSM (normalized ~)

Training: crop patches (400 * 400)

data augmentation

Source: http://www2.isprs.org/commissions/comm3/wg4/semantic-labeling.html

CAC-NET: experiment

Evaluation metric:

Intersection over Union (IoU)

$$IoU(P_m, P_{gt}) = \frac{|P_m \cap P_{gt}|}{|P_m \cup P_{gt}|}$$

P_{gt} : ground truth

P_m : prediction

Table1: comparison with excellent deep models

Table2: ablation experiment

Table 1: Comparison with the state-of-the-art models(%). surf: impervious surface (roads), veg: low vegtation.

Method	surf	roof	veg	tree	car	Mean
Segnet [5]	66.9	76.1	44.6	69.7	62.4	63.9
FCN-8s [1]	75.2	80.4	65.6	70.5	45.8	67.5
Deeplab-vgg [16]	80.0	87.9	70.0	75.4	36.1	69.9
Ours(vgg)	**81.3**	**89.3**	**70.3**	**75.5**	**66.4**	**76.6**
Deeplab-res101	81.6	90.7	71.4	**76.7**	58.9	75.9
Ours(res101)	**84.0**	**90.9**	**72.1**	76.6	**75.3**	**79.8**

Table 2: Ablation Experiment(%). MPD: multiple average pooling and dilation, MCC: multi-context cascade, RC: residual correction.

Method	surf	roof	veg	tree	car	Mean
Ours(Deeplab_13)	76.7	82.3	67.8	72.6	40.7	68.0
+ MPD	79.7	86.5	68.3	74.6	47.2	71.3
+ Refinement	80.1	87.1	68.0	74.6	55.5	73.1
+ MCC	80.3	88.1	69.5	76.5	60.0	74.9
+ RC	81.3	89.3	70.3	75.5	66.4	76.6

CAC-NET: experiment

Online evaluation metric:

F1 score and *Overall Accuracy*

$$F1 = 2\frac{pre \times rec}{pre + rec} \text{ and } rec = \frac{tp}{C}, pre = \frac{tp}{P}$$

Table3: ISPRS 2D semantic labeling challenge results

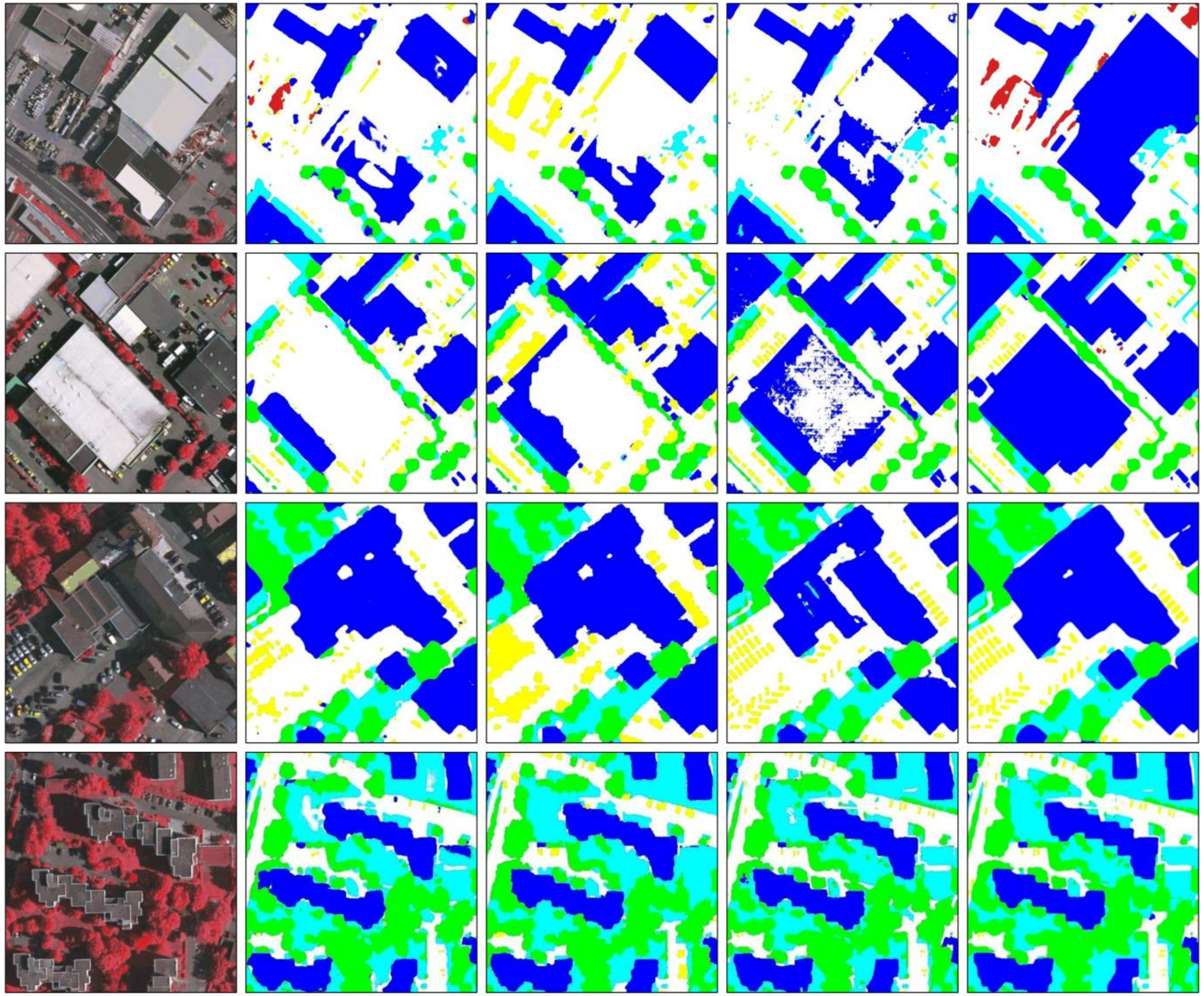
Table 3: *ISPRS 2D Semantic Labeling Challenge* results(%). OA: Overall Accuracy, DSM: Digital Surface Model

Method	surf	roof	veg	tree	car	OA
FCN+DSM('UZ_1')	89.2	92.5	81.6	86.9	57.3	87.3
CNN+RF+CRF+DSM [3]	89.5	93.2	82.3	88.2	63.3	88.0
FCN+RF+CRF [2]	90.5	93.7	83.4	89.2	72.6	89.1
FCN+Edge+DSM [10]	90.4	93.6	83.9	89.7	76.9	89.2
Segnet+DSM [19]	91.0	94.5	**84.4**	**89.9**	77.8	89.8
Ours(res101)	**92.7**	**95.3**	84.3	89.6	**80.8**	**90.6**

Source: http://www2.isprs.org/vaihingen-2d-semantic-labeling-contest.html

CAC-NET: experiment

Qualitative comparison



IRRG data | FCN+DSM 'UZ_1' | CNN+RF+ CRF+DSM [3] | Segnet+ DSM [17] | Ours

Source: *http://www2.isprs.org/vaihingen-2d-semantic-labeling-contest.html*

Future work

Instance labeling

Thank you for your attention !