Local Attention and Global Representation Collaborating for Fine-grained Classification

He Zhang[1], Yunming Bai[2], Hui Zhang[1]*, Jing Liu[1], Xingguang Li[1], Zhaofeng He[3]
1 Beijing IrisKing Co., Ltd.
2 Forenisc Science Center of Guangdong Provincial Public Security Bureau
3 Beijing University of Posts and Telecommunications

1 INTRODUCTION

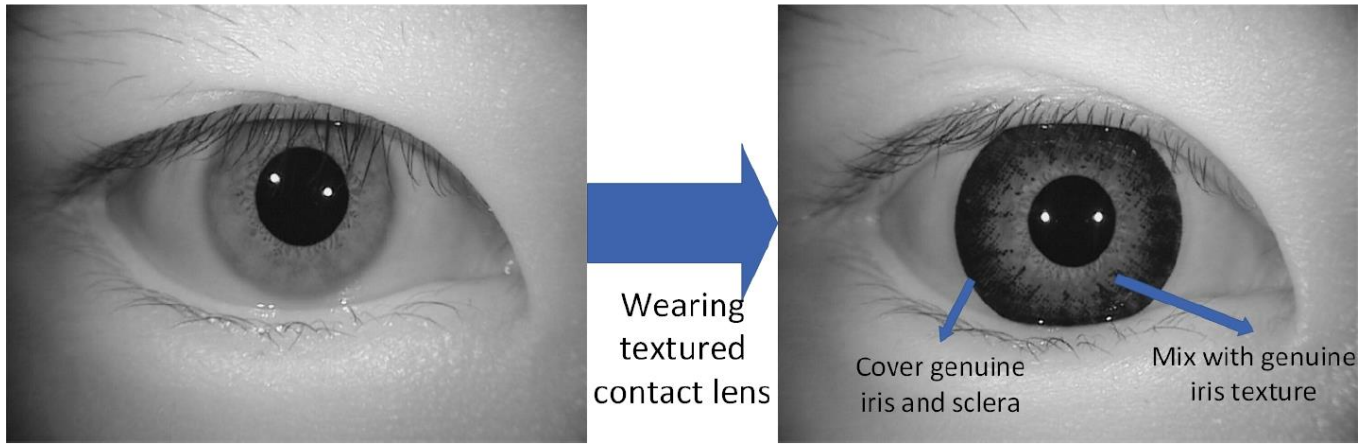
1.1 Iris recognition

- One of the most promising biometric modalities.
- Widely applied in authentication security, which is important to citizens, organization and country.

1 INTRODUCTION

1.2 Cosmetic lenses: easy-to-use iris presentation attack means.

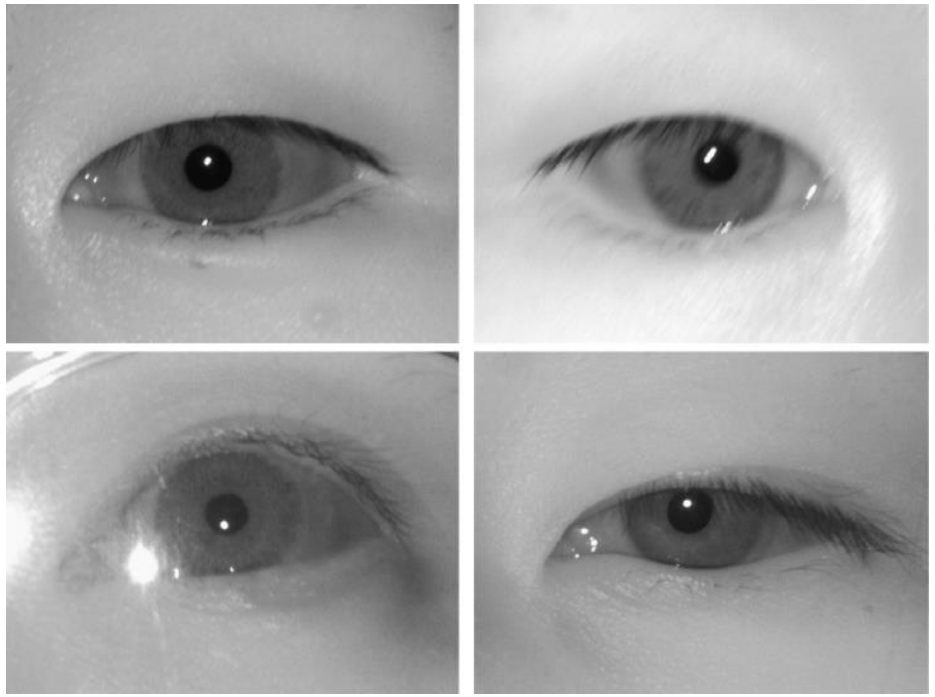
- Iris textural pattern is the foundation for iris recognition.
- The cosmetic contact lenses over an iris may change original iris textural pattern.



1 INTRODUCTION

1.3 Low iris image quality

- Changes of shooting angle.
- Different textures of lenses.
- Illumination and other factors.



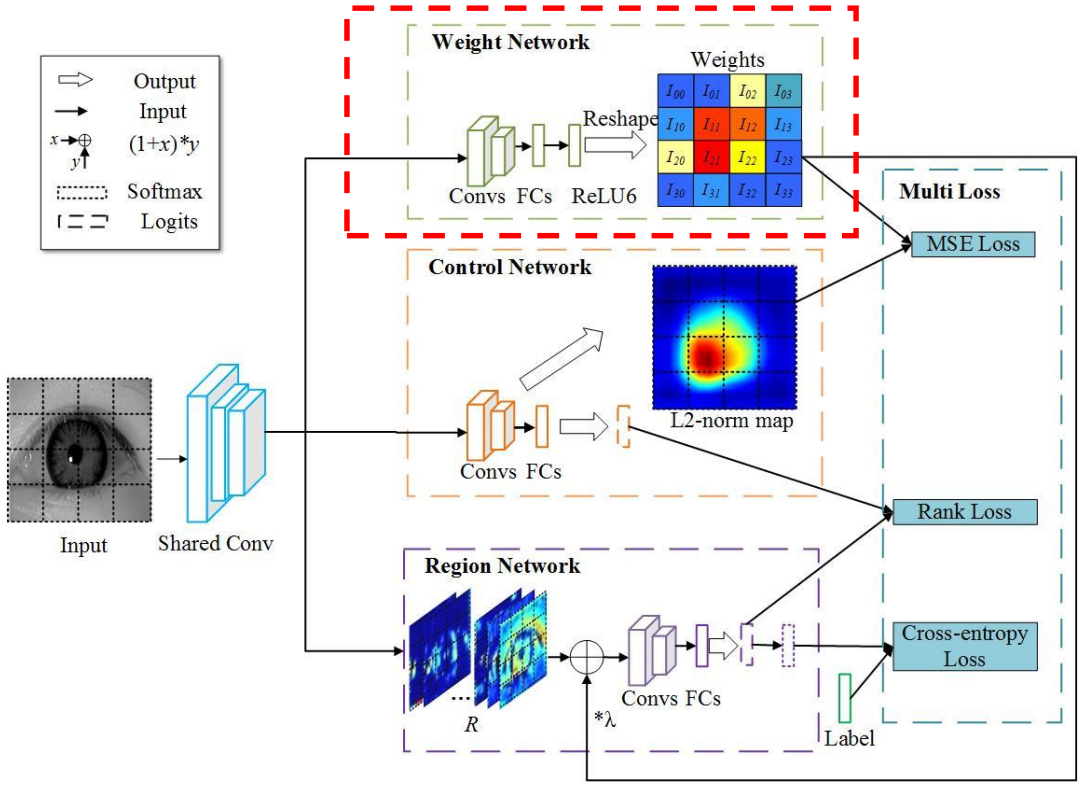
1.4 Drawbacks of existing methods:

- Accurate detection, location or segmentation requirements.
- Poor performance under complex environment.

2 APPROACH

2.1 Weight Network

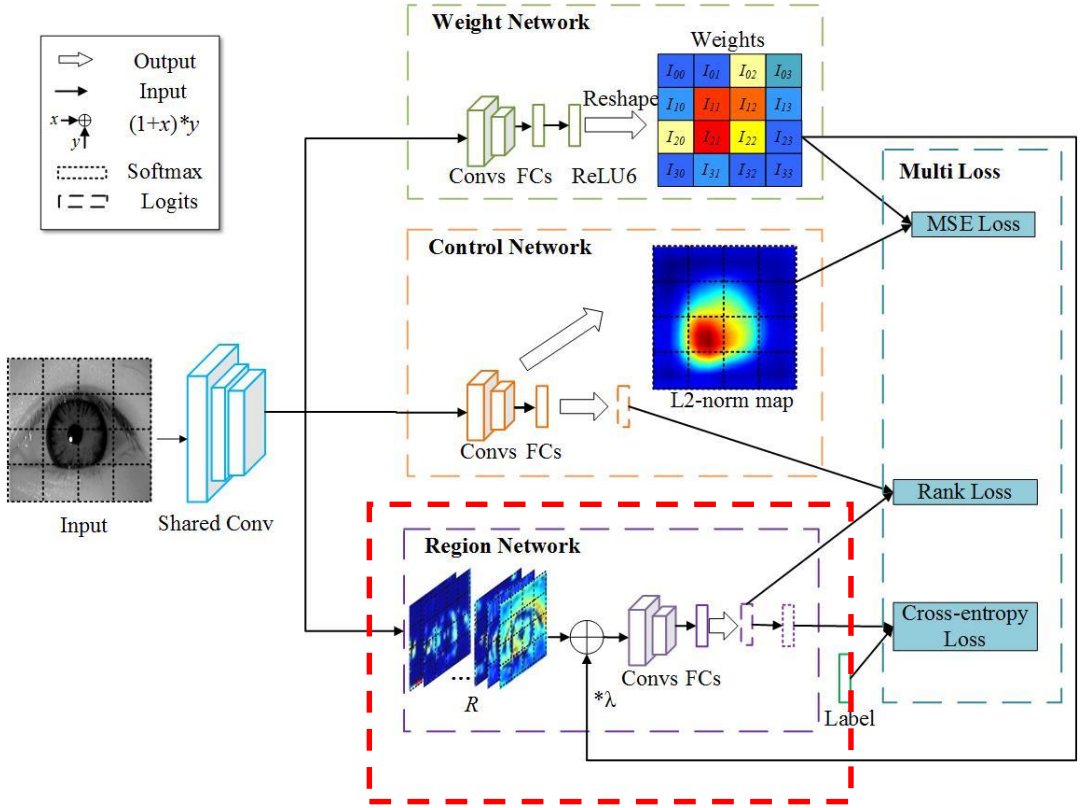
- To evaluate the effect of different regions to classification.
- Can be regarded as the attention distribution model.



2 APPROACH

2.2 Region Network

- Responsible for the whole classification task.
- Taking the fusion original image and attention distribution as input.



2 APPROACH

2.3 Discriminative Localization

- Higher L2-norm value indicates higher probability that objects or discriminative parts exist.
- Take L2-norm map as an initial ground truth to train the Weight Network.

$$LN_{ij} = \sqrt{\sum_{k=1}^{c} F_{ijk}^2}$$

2 APPROACH

2.4 Loss

Rank loss

$$L_r = \max\{0, mg - (g_w^{(j)} - g_p^{(j)})\}$$

Mean Squared Error (MSE) loss

$$LN'_{ij} \leftarrow \frac{LN'_{ij} - \min(LN')}{\max(LN') - \min(LN')} \cdot s$$

$$L_{mse} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (I_{ij} - LN'_{ij})^2$$

Cross-entropy loss

$$p_i = \frac{e^{g_w^{(i)}}}{\sum_{j=1}^{c} e^{g_w^{(j)}}} \qquad L = \frac{1}{m} \sum_{i=1}^{m} \log p_{y_i}$$

3 EXPERIMENTS

3.1 Datasets

Iris cosmetic contact lenses datasets

	Train		Test	
Datasets	Live	Cosmetic	Live	Cosmetic
ND-I	2000	1000	800	400
ND-II	400	200	200	100
IF-VE	20000	20000	5000	5000
CASIA-IF	4800	592	1200	148

CUB-200-2011

Black Footed Albatross

White throated Sparrow

3 EXPERIMENTS

3.2 Results

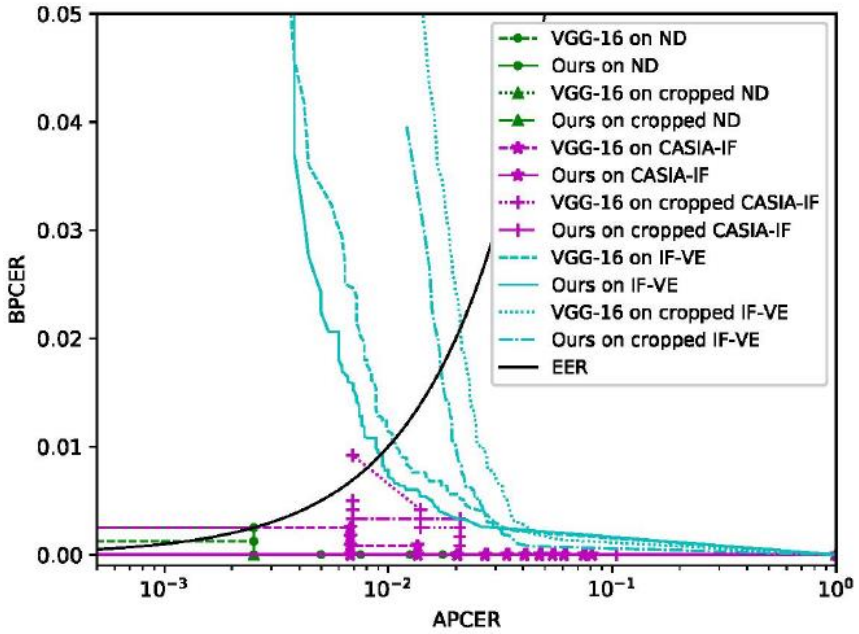
	ND-I	IF-VE	CASIA-IF
GHCLNet [7]	95.83	-	-
ContlensNet [6]	96.63	-	-
2-Channel [13]	-	-	97.07
HVC [5]	**100.00**	-	99.32
FT VGG-16	99.92	98.22	99.63
Ours	**100.00**	**98.58**	**99.70**

With iris region localizing

	M-ND-I	M-IF-VE	M-CASIA-IF
FT VGG-16	99.25	81.86	95.04
Ours	**99.33**	**89.91**	**95.26**

Randomly moved

	C-ND-I	C-IF-VE	C-CASIA-IF
FT VGG-16	99.83	98.92	99.85
Ours	**99.92**	**99.11**	**100.00**



Without iris region localizing

3 EXPERIMENTS

3.2 Results

Methods	w/o. BBox	w/. G-BBox
FT VGG-16	74.3	78.9
Ours	75.6	80.4

Experiments on CUB-200-2011

4 CONCLUSIONS

- The Weighted Region Network (WRN) is proposed for cosmetic contact lenses detection.
- With the inherent attention mechanism, WRN is able to automatically find the most discriminative regions, showing advantages in analyzing the low-quality iris images with cosmetic contact lenses.
- It can also be used for fine-grained image classification.

Local Attention and Global Representation Collaborating for Fine-grained Classification

He Zhang[1], Yunming Bai[2], Hui Zhang[1]*, Jing Liu[1], Xingguang Li[1], Zhaofeng He[3]
1 Beijing IrisKing Co., Ltd.
2 Forenisc Science Center of Guangdong Provincial Public Security Bureau
3 Beijing University of Posts and Telecommunications

Thanks!