

ERRATUM

Open Access



# Erratum to: Refining deep convolutional features for improving fine-grained image recognition

Weixia Zhang<sup>1</sup>, Jia Yan<sup>1</sup>, Wenxuan Shi<sup>2</sup>, Tianpeng Feng<sup>1</sup> and Dexiang Deng<sup>1\*</sup>

## Erratum

Upon publication of the original article [1], it was noticed that there were several blanks in the Table 5 and the footnote of the Table 5, 'The 'n/a' entries in the table means that bounding box or part annotation is not used.' was incorrectly given as 'The 'n/a' entries in the table means that the results are not available.' This has now been acknowledged and corrected in this erratum. This has now been incorporated in the new Table 5 shown below.

## Author details

<sup>1</sup>School of Electronic Information, Wuhan University, Wuhan, China. <sup>2</sup>School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China.

Received: 3 May 2017 Accepted: 3 May 2017

Published online: 08 May 2017

## Reference

1. W Zhang, J Yan, W Shi, T Feng, D Deng, Refining deep convolutional features for improving fine-grained image recognition. *EURASIP Journal on Image and Video Processing* 2017(1), 27 (2017)

\* Correspondence: wythia1989@gmail.com

<sup>1</sup>School of Electronic Information, Wuhan University, Wuhan, China

**Table 5** Comparison of performance of our methods with some recent state-of-the-arts methods in cub. BBox, Parts denote bounding-box and parts annotation respectively

Methods	Train phase	Test phase	Dim.	Model	Acc.	DPD
Dataset: <b>cub</b>						
Part-Stacked CNN [1]	BBox + Parts	BBox	4,096	Part-Stacked CNN	76.2%	1.484
Deep LAC [2]	BBox + Parts	BBox	12,288	Alex-Net	80.3%	0.521
PN-CNN [3]	BBox + Parts	n/a	13,512	Alex-Net	85.4%	0.506
PG-Alignment [4]	BBox	n/a	126,976	VGG-19	82.8%	0.052
Symbolic [5]	BBox	BBox	20,992	Shallow feature: SIFT	59.4%	0.226
Cross layer pooling[6]	BBox	BBox	4,096	Alex-Net	73.5%	1.436
Mask-CNN [12]	Parts	n/a	8,192	VGG-16 + FCN	85.4%	0.834
Spatial Transformer CNN [33]	n/a	n/a	4,096	ST-CNN	84.1%	1.643
Bilinear CNN [8]	n/a	n/a	262,144	VGG-16 + VGG-M	84.1%	0.026
Compact Bilinear CNN [25]	n/a	n/a	8,192	VGG-16	84.0%	0.820
PD + SWFV [14]	n/a	n/a	69,632	VGG-16	84.5%	0.097
SCDA [13]	n/a	n/a	4,096	VGG-16	80.5%	1.572
<b>Ours</b>	n/a	n/a	69,992	VGG-16	86.4%	0.099
<b>Ours (Compact vector)</b>	n/a	n/a	4,096	VGG-16	84.5%	1.650
Dataset: <b>air</b>						
Symbolic [5]	BBox	BBox	20,992	Shallow feature: SIFT	72.5%	0.276
Re-Fisher Vector [34]	n/a	n/a	655,360	Shallow feature: SIFT	81.5%	0.001
Bilinear CNN [8]	n/a	n/a	262,144	VGG-16 + VGG-M	83.9%	0.0256
<b>Ours (Full Vector + MI 2)</b>	n/a	n/a	69,992	VGG-16	87.7%	0.100
<b>Ours (Compact vector)</b>	n/a	n/a	<b>4,096</b>	VGG-16	82.5%	1.611
Dataset: <b>cars</b>						
Symbolic [5]	BBox	BBox	20,992	Shallow feature: SIFT	78.0%	0.297
PG-Alignment [4]	BBox	n/a	126,976	VGG-19	92.6%	0.058
Re-Fisher Vector [34]	n/a	n/a	655,360	Shallow feature: SIFT	82.7%	0.011
Bilinear CNN [8]	n/a	n/a	262,144	VGG-16 + VGG-M	91.3%	0.028
<b>Ours</b>	n/a	n/a	69,992	VGG-16	92.4%	0.106
<b>Ours (Compact vector)</b>	n/a	n/a	4,096	VGG-16	87.5%	1.709
Dataset: <b>dogs</b>						
Symbolic [5]	BBox	BBox	20,992	Shallow feature: SIFT	45.6%	0.174
Selective Pooling [35]	BBox	BBox	163,840	Shallow feature: SIFT	52.0%	0.025
Re-Fisher Vector [34]	n/a	n/a	327,680	Shallow feature: SIFT	52.9%	0.013
NAC[36]	n/a	n/a	4,096	Alex-Net	68.6%	1.340
PD + SWFV [14]	n/a	n/a	36,864	Alex-Net	71.9%	0.156
<b>Ours</b>	n/a	n/a	40,000	Alex-Net	72.6%	0.145
<b>Ours (Compact vector)</b>	n/a	n/a	4,096	Alex-Net	68.4%	1.335

The 'n/a' entries in the table means that bounding box or part annotation is not used