Neural Computing and Applications (2022) 34:6571–6574
https://doi.org/10.1007/s00521-021-06407-7

CORRECTION

Correction to: Playing to distraction: towards a robust training of CNN classifiers through visual explanation techniques

David Morales[1,2] · **Estefania Talavera**[3] · **Beatriz Remeseiro**[4]

Received: 14 August 2021 / Accepted: 14 August 2021 / Published online: 16 September 2021

Correction to: Neural Computing and Applications
https://doi.org/10.1007/s00521-021-06282-2

Unfortunately, the article was published with some errors in Tables 1, 3, 5, and 7 and in the captions of Figures 1, 2, and 3 in the online version of the article.

The correct tables and figures are given (Tables 1, 3, 5, and 7 and Figs. 1, 2, and 3).

The original article can be found online at https://doi.org/10.1007/s00521-021-06282-2.

✉ David Morales
david.morales@hattec.de

Estefania Talavera
e.talavera.martinez@rug.nl

Beatriz Remeseiro
bremeseiro@uniovi.es

1 University of Granada, 18071 Granada, Spain

2 HAT.tec GmbH. c/o Universität der Bundeswehr, Werner-Heisenberg-Weg 39, 85579 Neubiberg, Germany

3 Department of Computer Science, University of Groningen, Nijenborgh 9, 9747 AG Groningen, The Netherlands

4 Department of Computer Science, Universidad de Oviedo, Campus de Gijón s/n, 33203 Gijón, Spain

Table 1 Classification performance, averaged across five runs, of the different approaches on the Stanford cars [11] and FGVC-Aircraft [12] datasets

	FT-ResNet50	0-occlusion	R-occlusion	1-occlusion
Stanford cars				
Accuracy	0.849 ± 0.009	**0.871 ± 0.007**	0.860 ± 0.009	0.869 ± 0.008
Precision	0.855 ± 0.007	**0.876 ± 0.007**	0.866 ± 0.008	0.873 ± 0.008
Recall	0.849 ± 0.009	**0.870 ± 0.008**	0.860 ± 0.009	0.868 ± 0.009
F1	0.848 ± 0.009	**0.870 ± 0.008**	0.859 ± 0.009	0.867 ± 0.009
FGVC-Aircraft				
Accuracy	0.731 ± 0.013	**0.749 ± 0.005**	0.739 ± 0.012	0.743 ± 0.005
Precision	0.746 ± 0.011	**0.762 ± 0.005**	0.755 ± 0.010	0.759 ± 0.004
Recall	0.731 ± 0.013	**0.749 ± 0.005**	0.739 ± 0.012	0.743 ± 0.005
F1	0.731 ± 0.014	**0.748 ± 0.005**	0.739 ± 0.012	0.743 ± 0.005

Best results are in bold

Table 3 Classification performance, averaged across five runs, making use of different backbones on the Stanford cars [11] and FGVC-Aircraft [12] datasets

	FT-InceptionV3	0-occlusion-InceptionV3	FT-DenseNet	0-occlusion-DenseNet
Stanford cars				
Accuracy	0.778 ± 0.023	**0.791 ± 0.020**	0.883 ± 0.010	**0.894 ± 0.011**
Precision	0.788 ± 0.021	**0.798 ± 0.020**	0.888 ± 0.009	**0.898 ± 0.011**
Recall	0.777 ± 0.023	**0.791 ± 0.020**	0.882 ± 0.010	**0.893 ± 0.012**
F1	0.776 ± 0.023	**0.790 ± 0.021**	0.882 ± 0.010	**0.893 ± 0.012**
FGVC-Aircraft				
Accuracy	0.618 ± 0.029	**0.633 ± 0.026**	0.767 ± 0.026	**0.780 ± 0.025**
Precision	0.630 ± 0.030	**0.641 ± 0.029**	0.786 ± 0.024	**0.794 ± 0.023**
Recall	0.618 ± 0.028	**0.633 ± 0.026**	0.767 ± 0.026	**0.780 ± 0.025**
F1	0.616 ± 0.029	**0.630 ± 0.026**	0.768 ± 0.026	**0.780 ± 0.025**

Best results per backbone architecture are in bold

Table 5 Classification performance, averaged across five runs, of the different approaches on the EgoFoodPlaces dataset [15]

	Hierarchical approach [15]	FT-ResNet50	0-occlusion
Macro Precision	0.56	**0.59 ± 0.03**	**0.59 ± 0.05**
Macro Recall	0.53	**0.55 ± 0.03**	0.54 ± 0.06
Macro F1	**0.53**	**0.53 ± 0.04**	**0.53 ± 0.06**
Weighted Precision	0.65	0.67 ± 0.02	**0.68 ± 0.03**
Weighted Recall	**0.68**	0.67 ± 0.03	**0.68 ± 0.04**
Weighted F1	0.65	0.64 ± 0.03	**0.66 ± 0.04**
Accuracy	**0.68**	0.67 ± 0.03	**0.68 ± 0.04**

Best results are in bold

Table 7 Classification performance, averaged across five runs, of the baseline method and the proposed training scheme when we randomly hid some regions on the test images

	FT-ResNet50	0-occlusion
Macro Precision	0.53 ± 0.01	**0.54 ± 0.02**
Macro Recall	0.47 ± 0.02	**0.48 ± 0.03**
Macro F1	0.47 ± 0.02	**0.48 ± 0.05**
Weighted Precision	**0.63 ± 0.02**	**0.63 ± 0.03**
Weighted Recall	0.59 ± 0.02	**0.65 ± 0.03**
Weighted F1	**0.59 ± 0.02**	**0.59 ± 0.02**
Accuracy	0.59 ± 0.02	**0.60 ± 0.02**

Best results are in bold

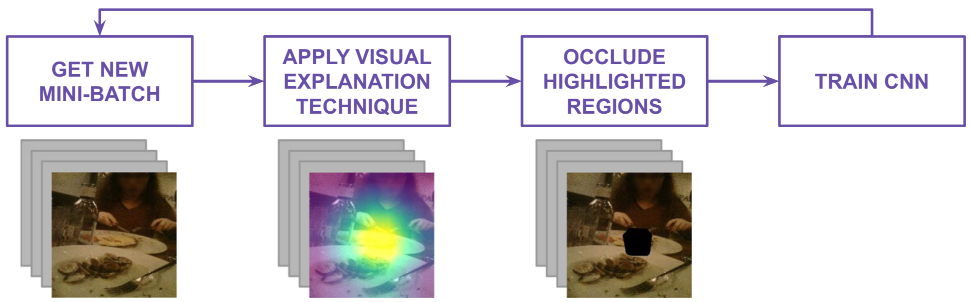

Fig. 1 Workflow of our alternative training scheme, which **1** gets a new mini-batch of input images, **2** applies a visual explanation technique to generate the heat maps, **3** occludes the regions highlighted in the previous step, and **4** trains the CNN classifier

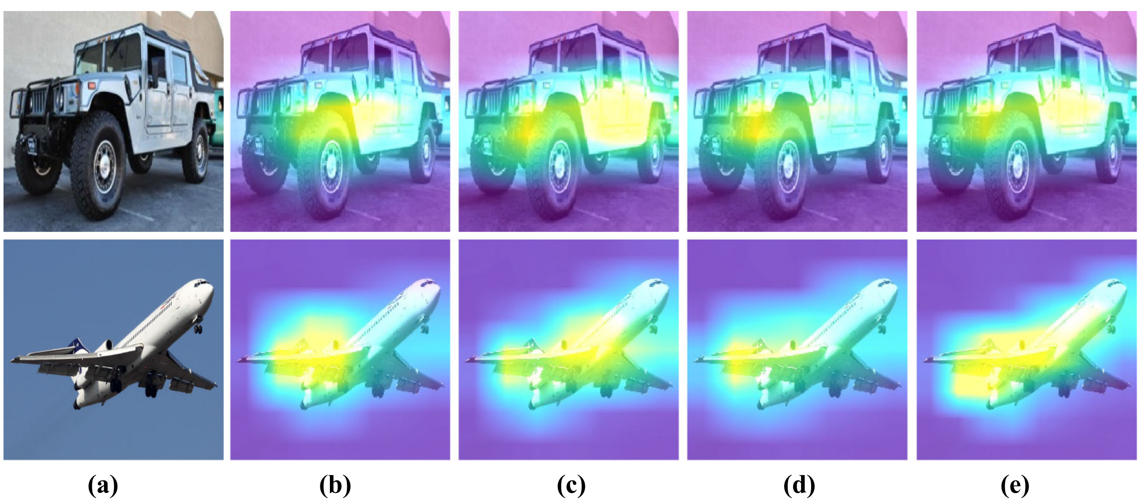

Fig. 2 **a** Input images from the Stanford cars (top) and FGVC-Aircraft (bottom) datasets, **b** heat maps generated by Grad-CAM for the baseline FT-ResNet50, and heat maps generated by Grad-CAM for the model trained with the proposed training scheme using **c** 0-occlusion, **d** R-occlusion, and **e** 1-occlusion

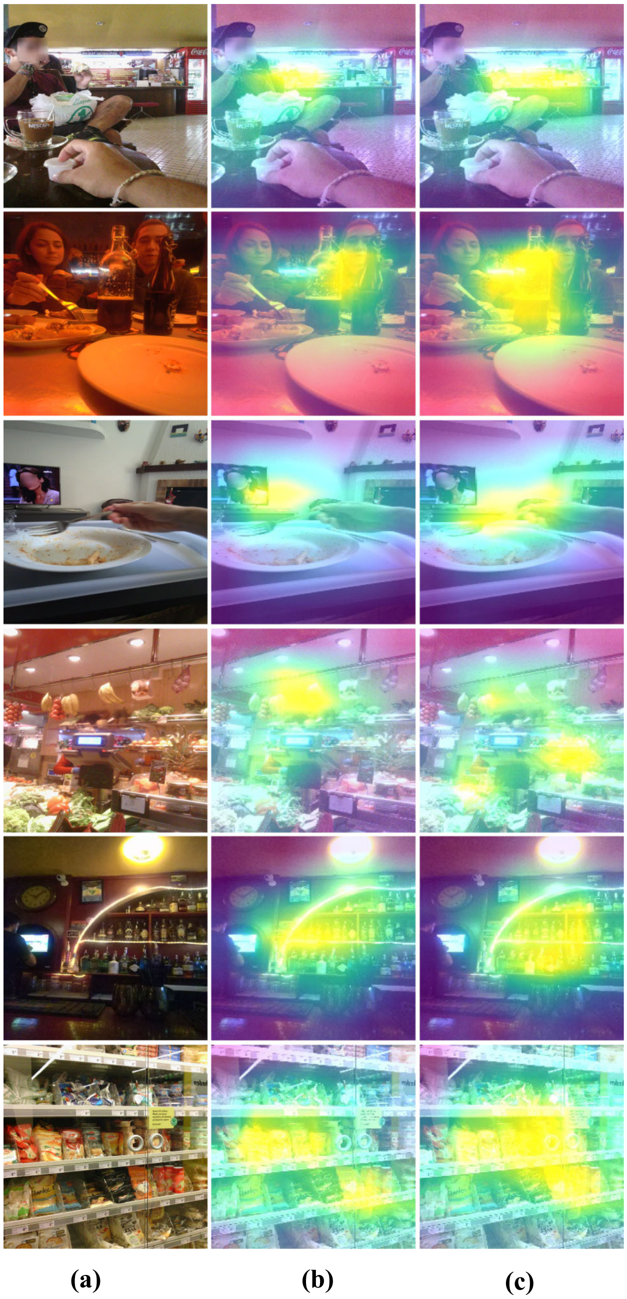

Fig. 3 **a** Input images, **b** heat maps generated by Grad-CAM for the baseline FT-ResNet50, and **c** heat maps generated by Grad-CAM for the model trained with the proposed training scheme (0-occlusion)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.