Facial Landmarks Detection using a Cascade of Recombinator Networks

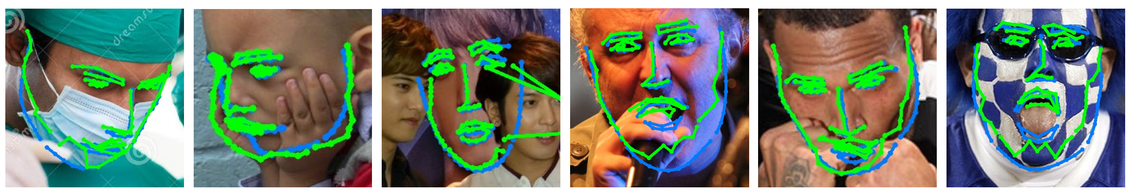
Pedro Diego López[1], Roberto Valle[1], Luis Baumela[1]
[1] Univ. Politécnica de Madrid
http://www.dia.fi.upm.es/~pcr/research.html

MADRID CIARP 2018

Problem Definition and Contribution

Facial landmarks detection is a crucial step for many face analysis problems such as verification, recognition, attributes estimation, etc.

- CNNs are very robust but they lack accuracy because they cannot enforce the landmarks to represent a valid face shape of a human face.



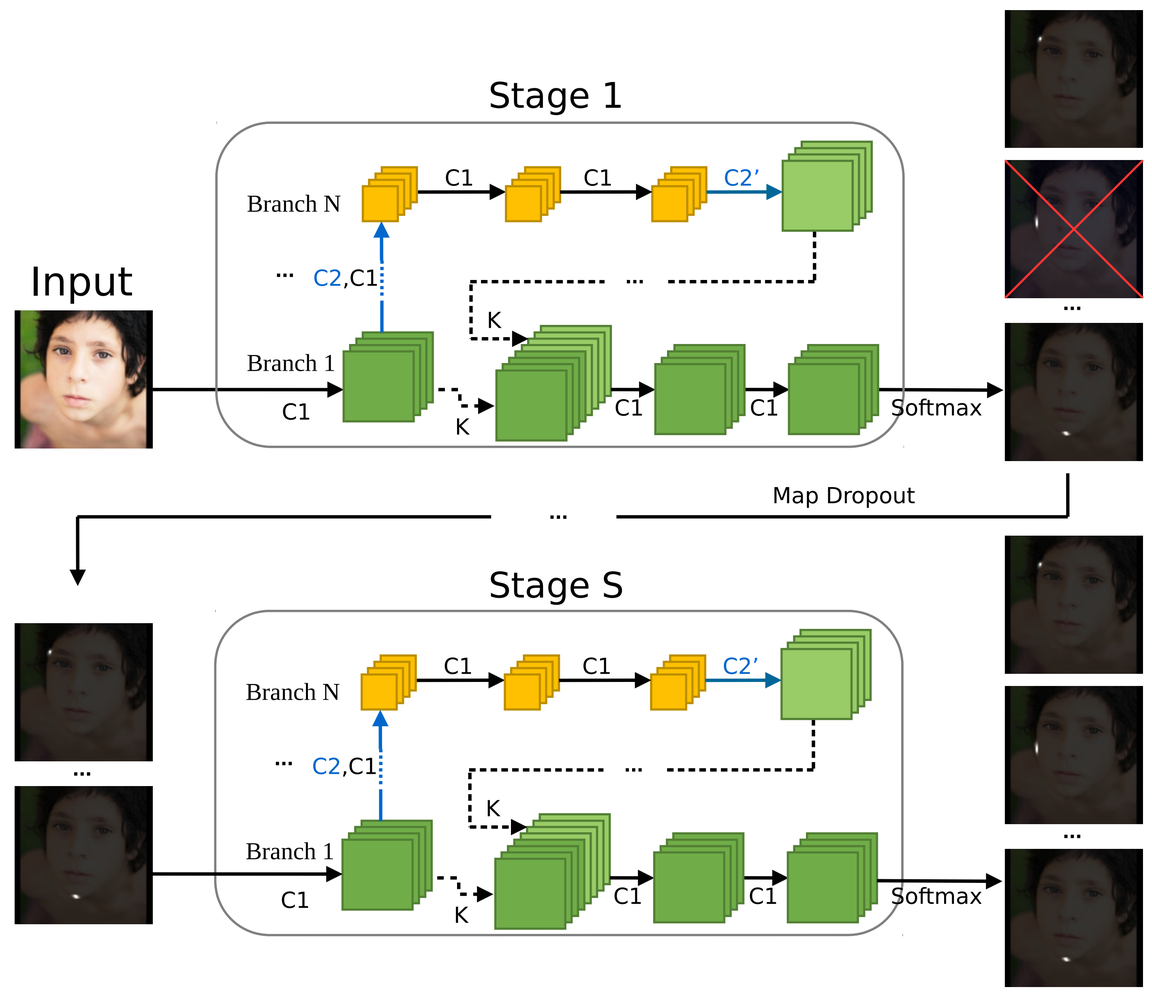
We investigate the use of a cascade of CNN regressors and a map dropout layer to learn position of landmarks using the location of its neighbors.

- We present a loss function that is able to handle missing landmarks. It allows an aggressive data augmentation that provides state-of-the-art results in the 300W, COFW and AFLW data sets.

$$\mathcal{L} = \sum_{i=1}^{N} \left(-\frac{1}{||\mathbf{w}_i^g||_1} \sum_{l=1}^{L} (\mathbf{w}_i^g(l) \cdot \mathbf{m}_i^g(l) \cdot \log(\mathbf{m}_i(l))) \right), \qquad (1)$$

Architecture diagram

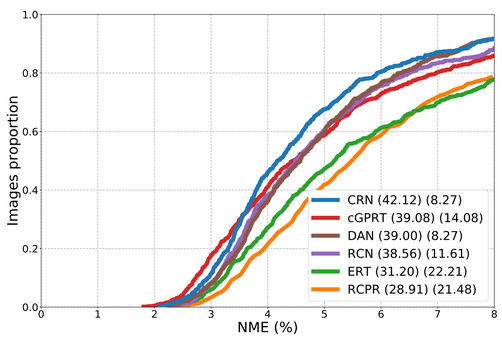
- It is composed of S stages where each stage represents a network that combines features across multiple branches B.
- The output of each stage is a probability map per each landmark providing information about the position of the L landmarks in the input image.
- The maximum of each probability map determines the landmarks positions.

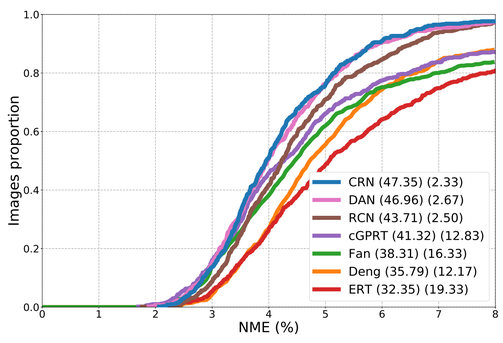


Experiments on 300W

300W public	Common pupils NME	Common corners NME	Challenging pupils NME	Challenging corners NME	Full pupils NME	Full corners NME	Full corners AUC_8	Full corners FR_8
RCN [2]	4.70	-	9.00	-	5.54	-	-	-
RCN+DKM [2]	4.67	-	8.44	-	5.41	-	-	-
DAN [3]	4.42	3.19	7.57	5.24	5.03	3.59	55.33	1.16
TSR [4]	4.36	-	7.56	-	4.99	-	-	-
RAR [6]	4.12	-	8.35	-	4.94	-	-	-
SHN [7]	4.12	-	7.00	4.90	-	-	-	-
CRN (S=**1**)	4.26	3.07	8.69	6.01	5.09	3.62	55.62	2.75
CRN (S=**2**)	4.12	2.97	7.90	5.47	4.83	3.44	57.44	1.88

300W private	Indoor corners NME	Indoor corners AUC_8	Indoor corners FR_8	Outdoor corners NME	Outdoor corners AUC_8	Outdoor corners FR_8	Full corners NME	Full corners AUC_8	Full corners FR_8
DAN [3]	-	-	-	-	-	-	4.30	47.00	2.67
SHN [7]	4.10	-	-	4.00	-	-	4.05	-	-
CRN (S=**1**)	4.42	45.91	1.66	4.45	45.25	2.66	4.43	45.59	2.16
CRN (S=**2**)	4.28	47.36	2.66	4.25	47.32	2.00	4.26	47.35	2.33

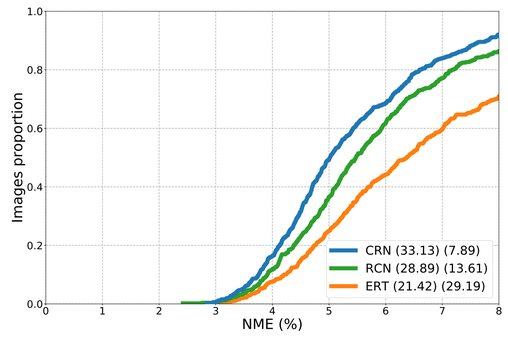


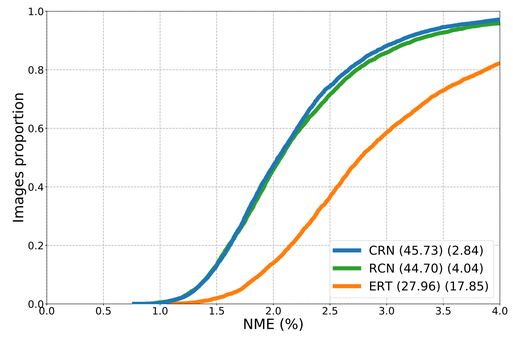


Experiments on COFW and AFLW

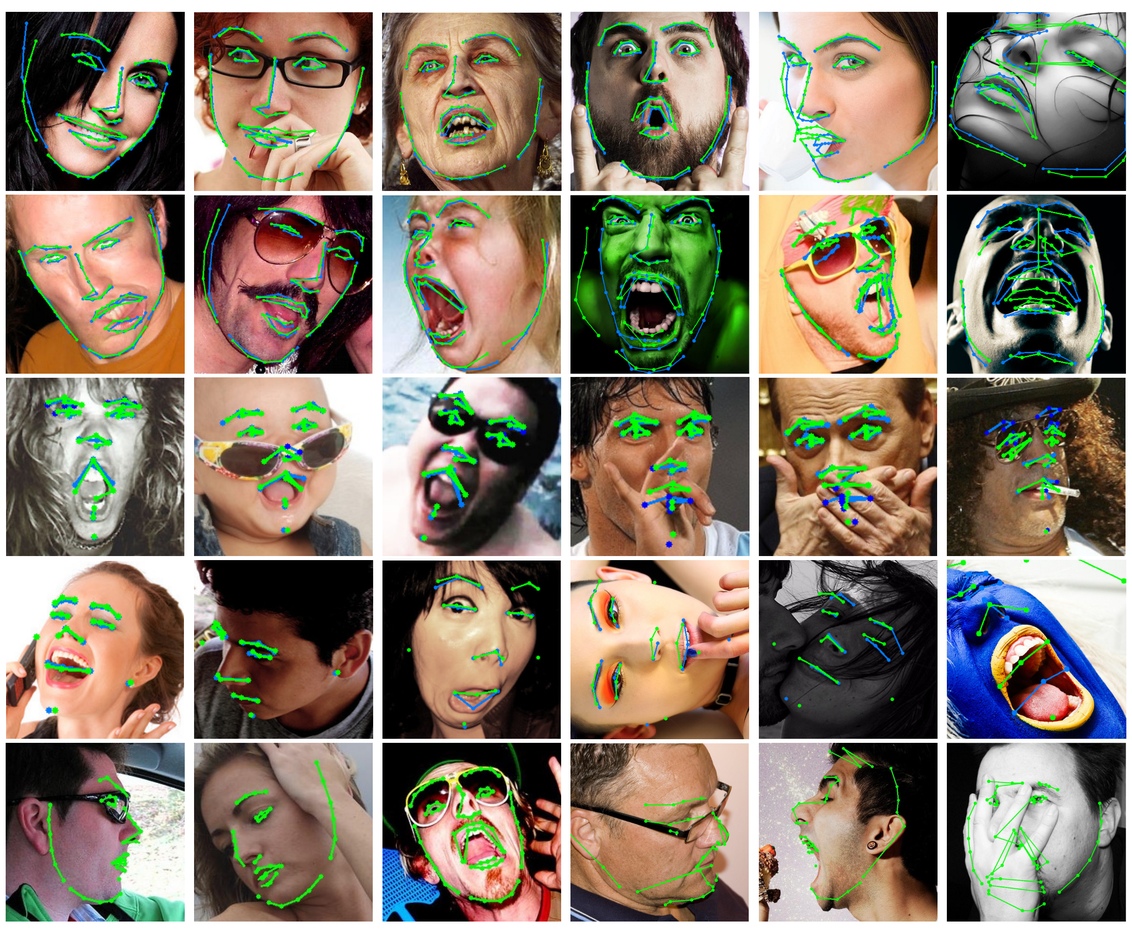
COFW	pupils NME	pupils AUC_8	pupils FR_8
RAR [6]	6.03	-	-
Wu *et al.* [5]	5.93	-	-
SHN [7]	5.6	-	-
CRN (S=**1**)	5.75	30.91	11.04
CRN (S=**2**)	5.49	33.13	7.88

AFLW	height NME
Bulat *et al.* [1]	2.85
CCL [8]	2.72
TSR [4]	2.17
CRN (S=**1**)	2.29
CRN (S=**2**)	2.21





Results



Acknowledgements: The authors gratefully acknowledge funding from the Spanish Ministry of Economy and Competitiveness under project TIN2016-75982-C2-2-R.

References

[1] Bulat, A., Tzimiropoulos, G.: Binarized convolutional landmark localizers for human pose estimation and face alignment with limited resources. In: Proc. International Conference on Computer Vision (ICCV) (2017)

[2] Honari, S., Yosinski, J., Vincent, P., Pal, C.J.: Recombinator networks: Learning coarse-to-fine feature aggregation. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)

[3] Kowalski, M., Naruniec, J., Trzcinski, T.: Deep alignment network: A convolutional neural network for robust face alignment. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017)

[4] Lv, J., Shao, X., Xing, J., Cheng, C., Zhou, X.: A deep regression architecture with two-stage re-initialization for high performance facial landmark detection. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017)

[5] Wu, Y., Ji, Q.: Robust facial landmark detection under significant head poses and occlusion. In: Proc. International Conference on Computer Vision (ICCV) (2015)

[6] Xiao, S., Feng, J., Xing, J., Lai, H., Yan, S., Kassim, A.A.: Robust facial landmark detection via recurrent attentive-refinement networks. In: Proc. European Conference on Computer Vision (ECCV) (2016)

[7] Yang, J., Liu, Q., Zhang, K.: Stacked hourglass network for robust facial landmark localisation. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017)

[8] Zhu, S., Li, C., Change, C., Tang, X.: Unconstrained face alignment via cascaded compositional learning. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)