Image Colorization with Deep Learning

Final Presentation

Michele Conti
Marco Aurelio Sterpa

Brief recap

- **Task**

  The model is fed black and white images. Its task is to produce plausible colorizations of the inputs.

- **How?**

  Images are produced by a convolutional neural network, whose structure is composed by encoding blocks followed by decoding ones.

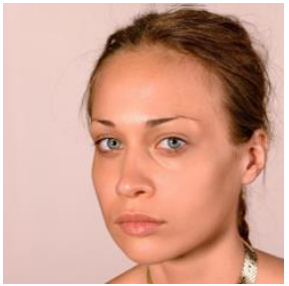
Data

CelebA (subset composed by 7200 images, cropped and aligned)

Low variance, lots of edges, highly possible to fall into edge pollution. Possible to obtain good results and draw meaningful conclusions with standard metrics.

Fruits (1300 fruit images crawled through Bing queries)

High variance between different classes, strong colors, strong edges. Possible to test VGG accuracy.



Approach

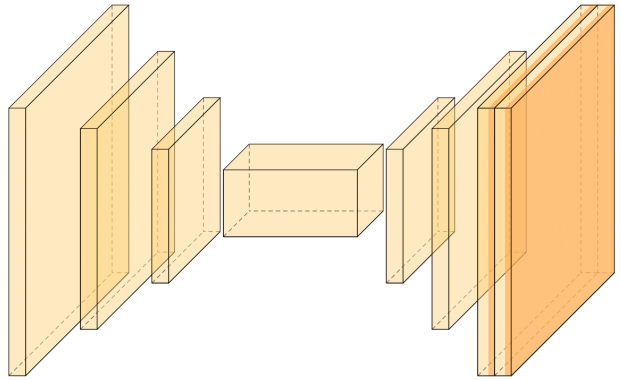
We trained our model in using two different loss functions:

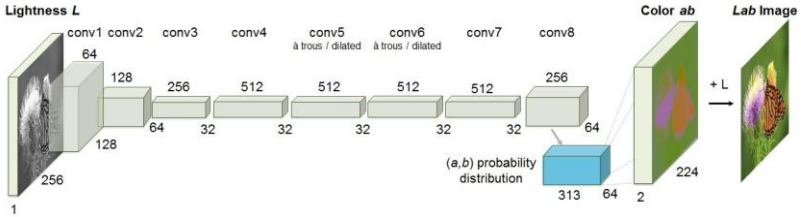
- **Regression**: Euclidean loss

$$L_2(\hat{Y}, Y) = \frac{1}{2}\sum_{h,w}\left\|Y_{h,w} - \hat{Y}_{h,w}\right\|_2^2$$

- **Classification**: Rebalanced multi cross entropy

$$L_{cl}(\hat{Z}, Z) = -\sum_{h,w} v(Z_{h,w}) \sum_{q} Z_{h,w,q} \log(\hat{Z}_{h,w,q})$$





Models

Classification:

- Multinomial classification with rebalancing (add value to **rare colours**)
- Multinomial classification without rebalancing (promotes **natural colorizations**)

Regression:

- Regression model with L2 loss (**regrets strong colorizations** to avoid high penalizations)

Multimodality

Image colorization is a **multimodal** problem: what does it mean for a coloration to be **correct**?

Ground truth

Model output

Metrics

- Standard metric

Raw accuracy: compare accuracy against ground truth images, b&w and randomly colorized images.

- Object detection

VGG accuracy: test an off-the-shelf classificator to check whether colorization helps with classification.

- **Human intervention**

Check if the output of the model feels right, natural and beautiful.

AuC (%)

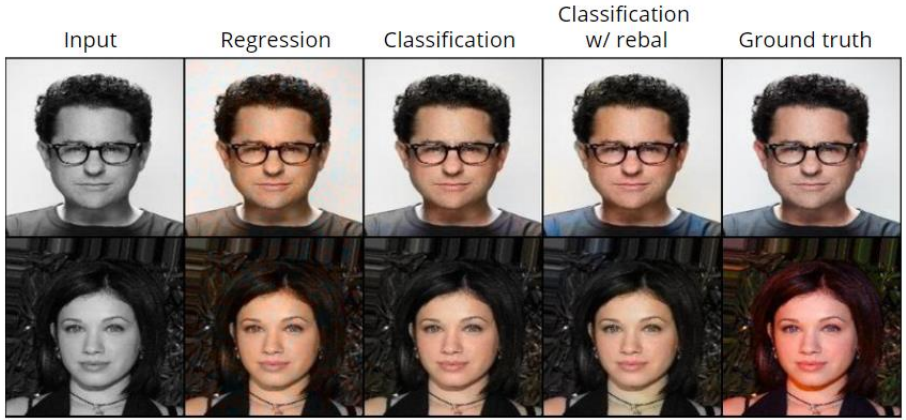
Dataset	CelebA	Fruits
Ground Truth	100	100
Gray	88.8	76.9
Random	87.4	74.1
L2	91.5	76.9
Class	92.2	**84.5**
Class + Rebalance	**92.3**	84.3

VGG Accuracy (on Fruits)

Dataset	Fruits
Ground Truth	44
Gray	33.2
Random	28.2
L2	31.7
Class	35.9
Class + Rebalance	**39.8**

Human intervention

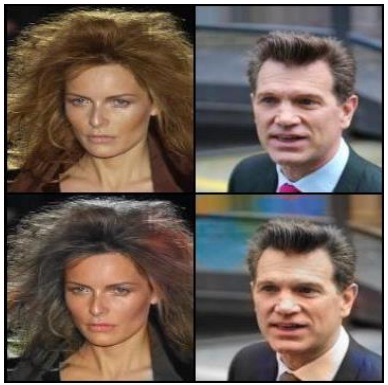
Obtaining good results in standard metrics doesn't automatically translate in good colorizations. In the end, the best option is to visualize model outputs themselves to evaluate them.



Conclusion

Best model: **classification with rebalancing**
But, it still has its limits:

- Color biases
- Color deficiency
- Inconsistent coloring
- Edge pollution
- Color bleeding



References

[1] S. Guadarrama, R. Dahl, D. Bieber, M. Norouzi, J. Shlens, and K. Murphy. Pixcolor: Pixel recursive colorization. arXiv, pages 1-17, 2017.
[2] C. Li and M. Wand. Combining markov random elds and convolutional neural networks for image synthesis, 2016.
[3] Z. Liu, P. Luo, X. Wang, and X. Tang. Deep learning face attributes in the wild. In Proceedings of International Conference on Computer Vision (ICCV), December 2015.
[4] S. Pahal and P. Sehrawat. Image colorization with deep convolutional neural networks. In G. S. Hura, A. K. Singh, and L. Siong Hoe, editors, Advances in Communication and Computational Technology, pages 45-56, Singapore, 2021. Springer Singapore.
[5] R. Zhang, P. Isola, and A. A. Efros. Colorful image colorization. Lecture Notes in Computer Science (including subseries Lecture Notes in Articial Intelligence and Lecture Notes in Bioin-formatics), 9907 LNCS:649-666, 2016.