ColorUNet: A convolutional classification approach to colorization

Vincent Billaut, Matthieu de Rochemonteix, Marc Thibault

CS231n Final Project, 03/12/2018

Colorization as classification

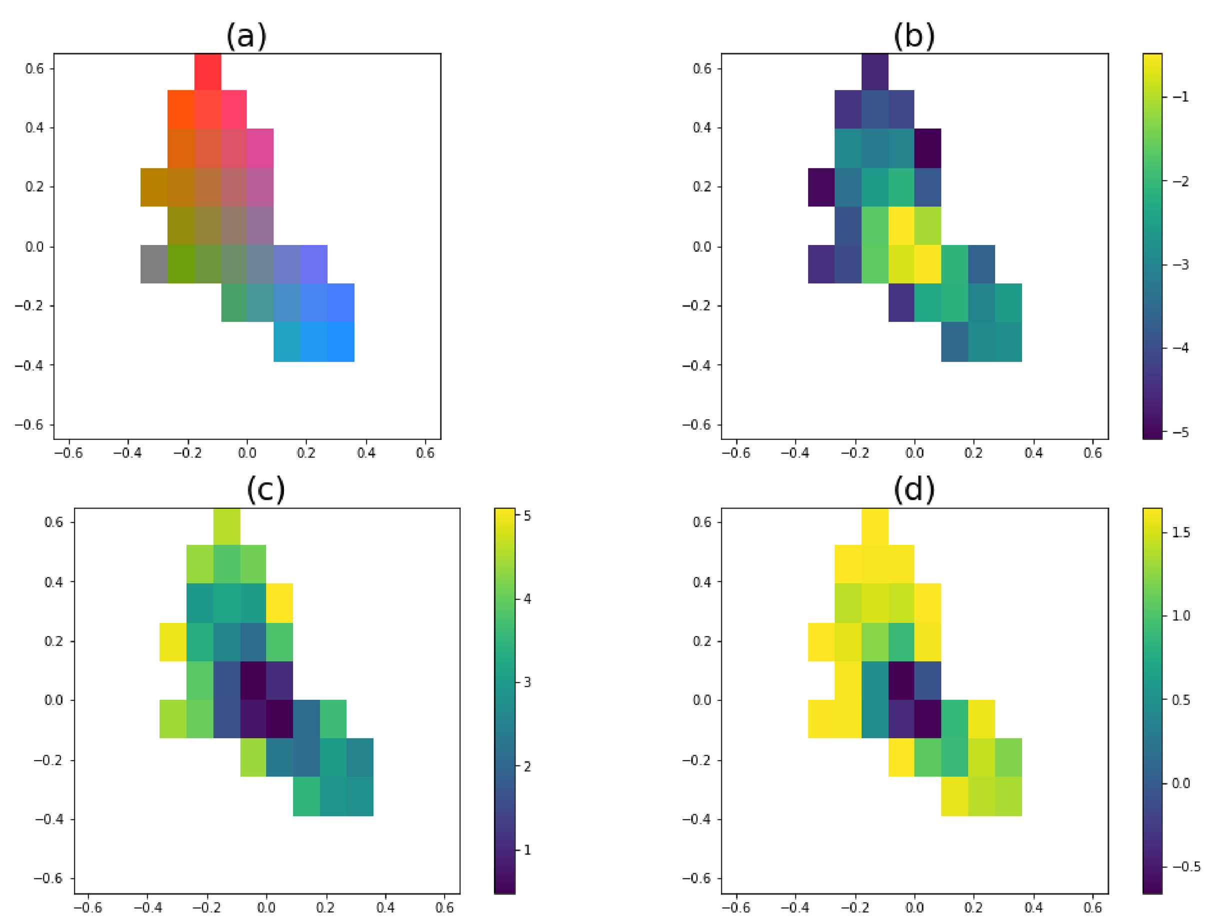

Figure 1: (a) Color map, showing the mean color (*chrominance*) of each of the selected bins (32 bins to select). (b) Frequency map (log-scale), shows the empirical frequency of the colors within each bin. (c) Inverse-frequency map. (d) Weight map (log-scale), shows the weights assigned to each bin after rebalancing.

We approach the colorization problem as a **classification problem** [2]. The main purpose is to boost the possibility of a rare color being predicted, and therefore we use the following tricks:

- Use *rebalancing* to give larger weights to rare colors in our loss function.
- Use an *annealed-mean* to output predictions y from the probability distribution $\mathbf{Z}$ over our n bins to the full original color space. To achieve this we use a temperature parameter $T > 0$ in the following softmax-like formula for one pixel

$$y = f_T(\mathbf{z}) = \frac{\exp(\log(\mathbf{z})/T)}{\sum_i \exp(\log(\mathbf{z}_i)/T)}$$

  where $\mathbf{z}$ is the n-dimensional probability vector of a given pixel over the n bins, and the sum in the denominator is over all the bins.

Dataset

To train our model, we used subsets of the SUN and ImageNet datasets. We selected 8 categories from ImageNet and 14 categories from SUN, that correspond mainly to nature landscapes. Our final training set is composed of 13,469 images, downscaled to a 256×256 resolution.
We used data augmentation to increase the robustness of training, namely :

- Flipping the image along the horizontal axis;
- Adding noise to the image (with different intensities);
- Selecting random crops of the image.

This procedure increases the size of the dataset sevenfold

ColorUNet outputs: influence of temperature

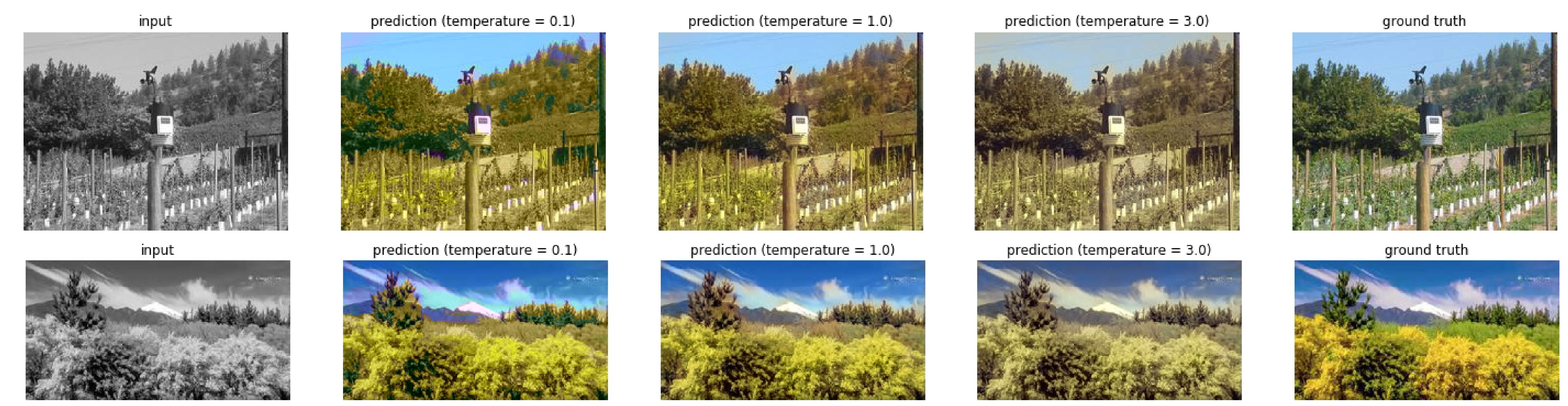

Figure 2: Examples of images colorized by ColorUNet. Temperature enables a trade-off between vivid colors and conservative colorizing.

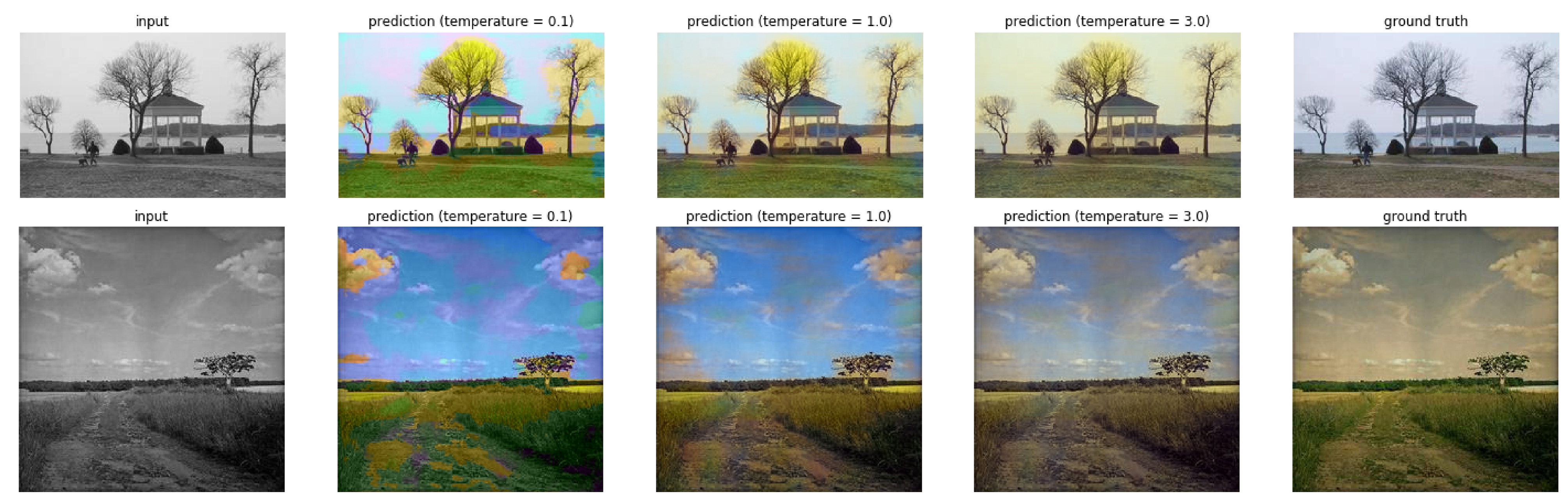

Figure 3: Sample predictions of the ColorUNet on the validation set, for bland input images. The ColorUNetâĂŹs output is more colorful than the ground truth. The bottom example is an old photograph with weared tones.

Understand ColorUNet: Confidence maps

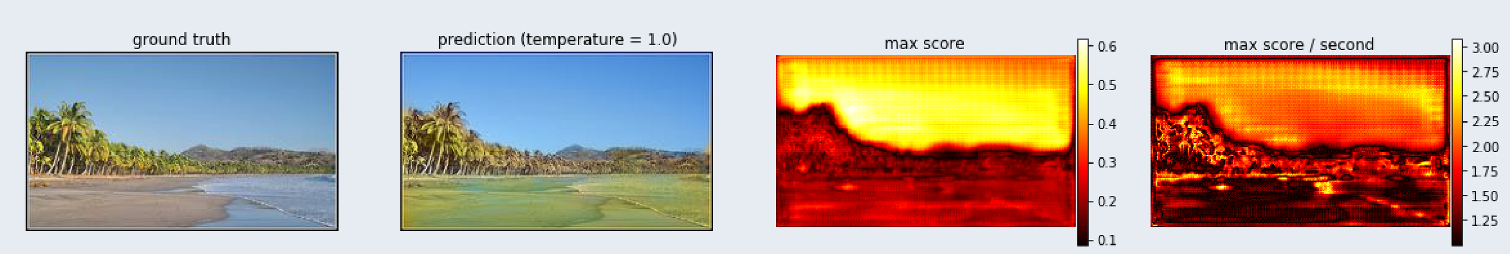

Figure 4: Confidence maps for the colorization of an image.

The model outputs a probability distribution for each pixel, $(p_i)_{i=1...32}$
Confidence maps are defined, for each pixel, as $C_1 = p_{(1)}$ $\quad C_2 = \frac{p_{(1)}}{p_{(2)}}$

These results have been obtained using NVIDIA K80 GPUs. Total training time was roughly 10 hours. Batch size used is 64.

[1] O. Ronneberger, P. Fischer, and T. Brox.
U-net: Convolutional networks for biomedical image segmentation.
CoRR, abs/1505.04597, 2015.

[2] R. Zhang, P. Isola, and A. A. Efros.
Colorful image colorization.
In *European Conference on Computer Vision*, pages 649–666. Springer, 2016.

Model architecture

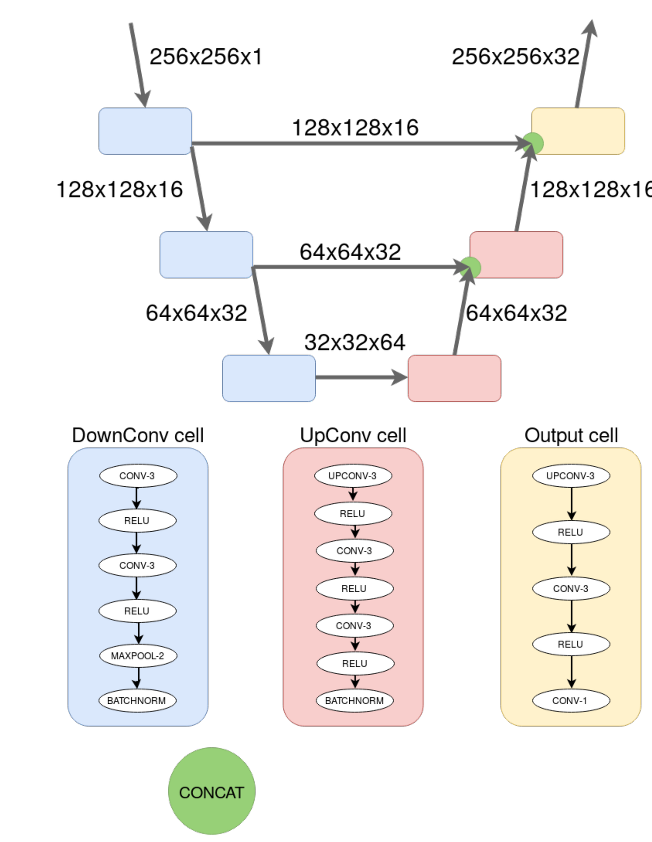

Figure 5: Structure of the ColorUNet.

After trying a simple convolutional architecture and a deep bottleneck architecture, we implemented an architecture based on the approach of [1], to improve both the gradient flow and the upscaling process.

Results analysis

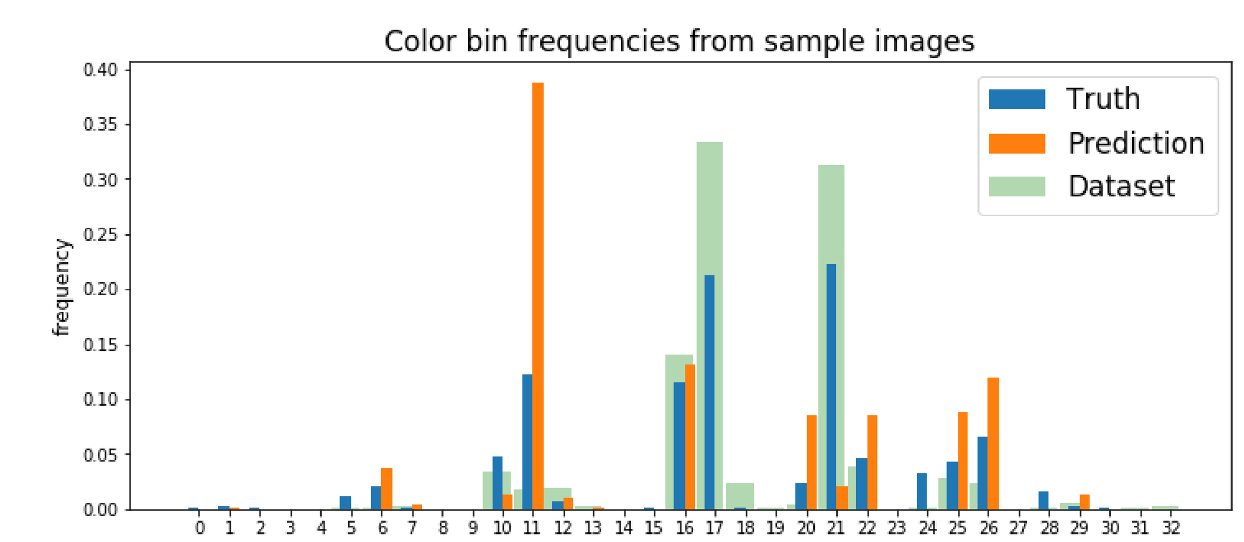

Figure 6: Histograms of the frequencies of the 32 selected color bins.

Comparison to baseline

Figure 7: Left: our network. Right: baseline