Computer Vision

Team 6/6

IMAGE COLOURING

Aritra Banerjee
2018102006

Jayati Narang
2018101066

Sahil Bhatt
2018111002

Srivathsan baskaran
2018101049

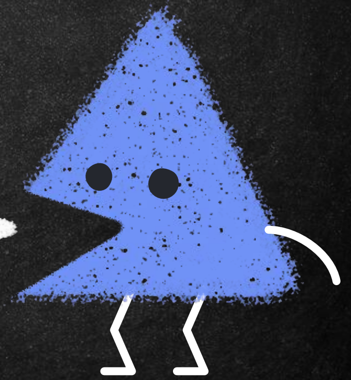
OVERVIEW

We have implemented the paper “*Let there be Color!: Joint End-to-end Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification*”.

Paper link : Let there be Color!

Our main objective was to create a model which can automatically colorize black and white images.

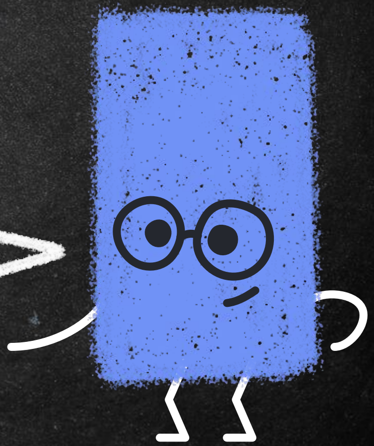
We also want our model to be able to classify images based on the scenery they depict.



RELATED WORK

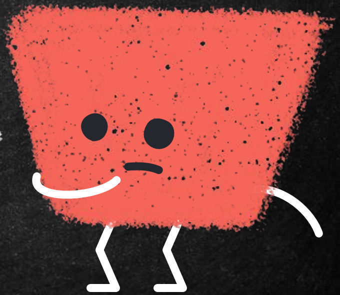
There has been a lot of work done a priori in this area. Some notable ones are:

1) Chrominance blending based on geodesic distances
2) Texture similarity based coloring *(by Luan)* which is used in the coloring of cartoons
3) Graph-cut based optimization techniques for coloring styles
4) Example based colorization robust to illumination differences *(by Liu)*. Here, user does not need to give the images, (available from web search)
5) An automatic approach *(Cheng, 2015)* in which multiple features are extracted and patches colorized with the help of neural networks. But the networks require a high-performance segmentation model to work



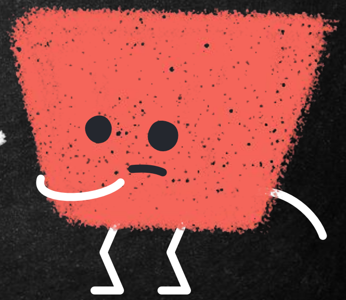
Why This Method ?

- Traditional colorization requires significant user interaction.
- This paper uses a fully automated data-driven approach for colorization of grayscale images.
- In this, we use a combination of global image priors, which are extracted from the entire image, and local image features, which are computed from small image patches, to colorize an image automatically.
- Global priors provide information at an image level such as whether or not the image was taken indoors or outdoors, whether it is day or night, etc.



Why This Method ?

- Local features represent the local texture or object at a given location.
- The approach is to use CNN. The global and local features are extracted jointly and then fuse together to perform the final colorization.
- This method requires neither pre-processing nor post-processing.
- To predict the chrominance, CIE L*a*b colorspace is used.



PROPOSED METHOD

Our model consists of four main components:

1. A Low-level features network
2. A Mid-level features network
3. A Global features network
4. A Colorization network

The components are all tightly coupled and trained in an end-to-end fashion. The output of our model is the chrominance of the image which is fused with the luminance to form the output image.

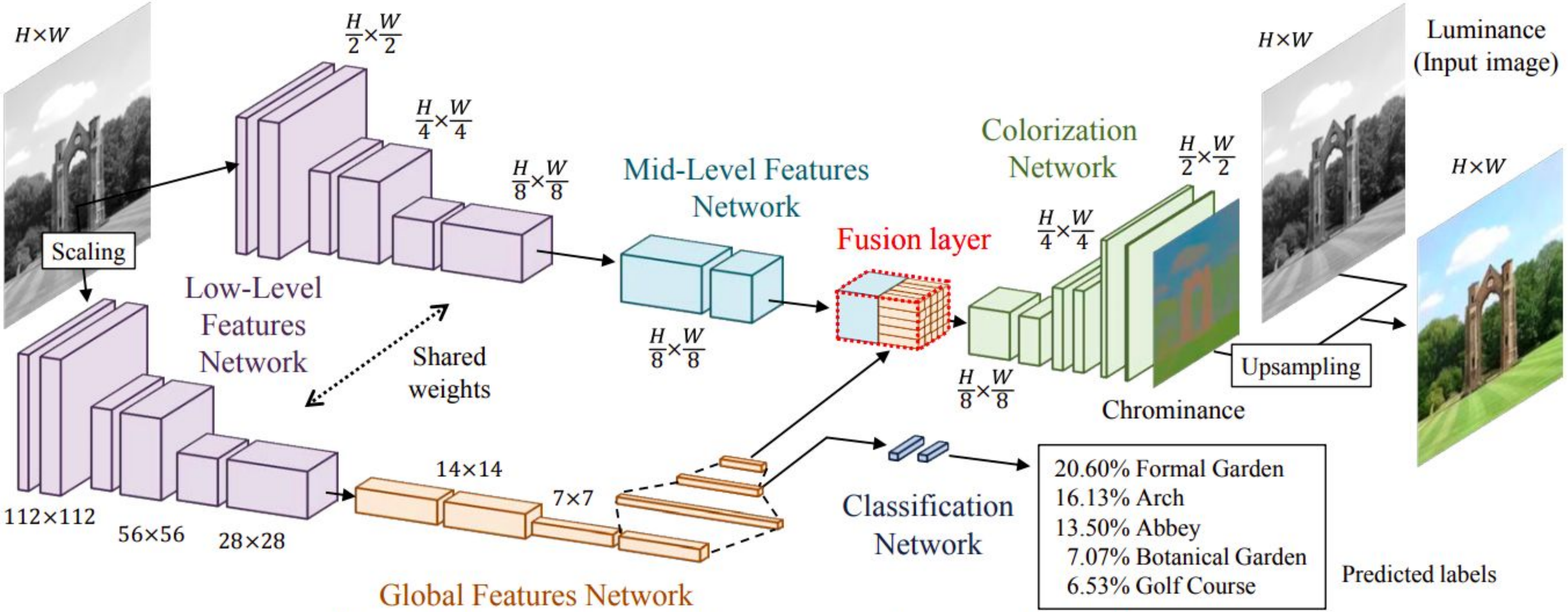


Figure 2: *Overview of our model for automatic colorization of grayscale images.*

Table 1: *Architectures of the different networks used in our model. Fully-Connected (FC) layers refer to the standard neural network layers from Eq.* (1), *convolution layers use Eq.* (2). *The output layer consists of a convolutional layer with a Sigmoid transfer layer instead of a ReLU transfer layer. Outputs refers to the number of output channels for the output of the layer. Upsample layers consist of using the nearest neighbour approach to increase the resolution of the output by a factor of* 2.

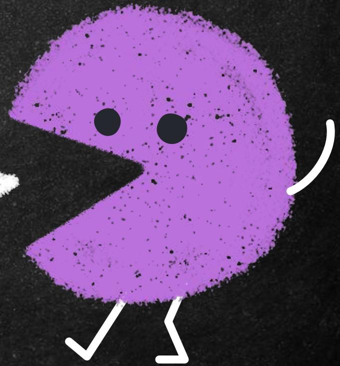
Image taken from the reference paper

Low-Level Features:

We use a 6-layer Convolutional Neural Network to obtain low-level features directly from the input image. The convolution filter is shared between the global features network and the mid-level features network.We use 3x3 convolution kernels exclusively and a padding of 1 x 1

Global Image Features:

The global image features are obtained by further processing the low-level features with four convolutional layers followed by three fully-connected layers resulting in a 256-dimensional vector



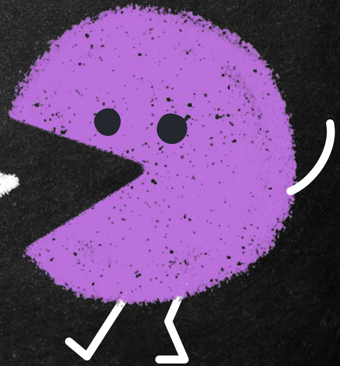
Mid-Level Features:

The mid-level features are obtained by processing the low-level features further with two convolutional layers.The output is 256-channel mid-level features.

The mid-level features networks are fully convolutional networks, such that the output is a scaled version of the input.

Fusion Layer:

This layer combines the global image features, with the (mid-level) local image features,to produce a H/8 x W/8 x 256-dimensional volume, we introduce a fusion layer. The resulting features are independent of any resolution constraints that the global image features might have.

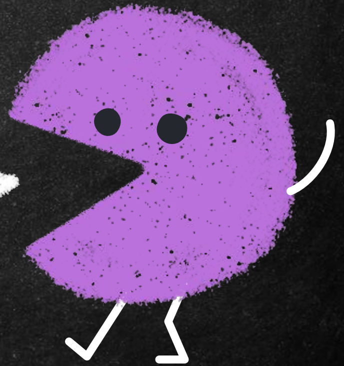


Colorization Network:

We process the fused features using a convolution layer and an upsampling layer by alternating them.The output layer consists of a convolutional layer with a Sigmoid transfer function that outputs the chrominance of the input grayscale image. We compute the MSE between the output and target output which is then back-propagated through all the networks to update all the parameters of the model.

Colorization with Classification:

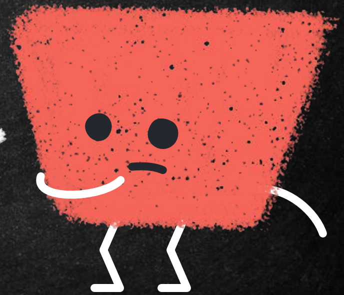
We optimize the training for colorization by also doing classification with labels. These labels correspond to a global image tag and can be used to guide the training of the global image features. The color loss affects the entire network, while the classification loss only affects the classification, global, and the shared low-level features network



Results

Features

- Our model performs best for image size of 224 X 224 .
- Since training occurs for a large dataset, we make the process efficient by keeping a uniform size for all images (224 X 224).
- We randomly flip the images horizontally with 50% probability to ensure robustness of the model.



Thank you