**SRM: A Style-based Recalibration Module for Convolutional Neural Networks**Adi Horowitz, Gilad Deutch

1. **Abstract**

*In our project paper, we aim to improve the representational ability of a CNN, by considering the style of the image, as defined below. SRM stands for a Style-based Recalibration Module. This module is directly fed into existing CNN architectures. By incorporating the relative importance of individual styles into feature maps, SRM learns to suppress the contribution of nuisance styles, and by doing that, the network can concentrate more on meaningful features, which boost its representational power. Finally, we added some improvements of our own. We examined the option of adding new features to expand the style features vector. Later on, we tested the importance of the SRM layer location inside the network. We performed experiments on a few general vision tasks and show the improvement of results when using SRM over other approaches such as Squeeze-and-Excitation (SE). We also show the results of SRM with our own modifications and additions. Throughout the experiments, SRM outperforms recent approaches though it requires orders of magnitude less additional parameters.*

1. **Intro**

It has been known that CNNs are capable of handling not only the content (i.e. shape) but also the style (i.e. texture) of an image. We aim to enhance the utilization of styles in a CNN to boost its representational power, in order to leverage the capabilities of CNN network with a wide range of vision classification problems.

SRM ﬁrst extracts the style information from each channel of the feature maps by style pooling, as defined below. Then, the module estimates per-channel recalibration weight via channel-independent

style integration. As part from our work, we tried to include more data in the style vector, like the correlation between different channels, the median, etc.

We will start by mentioning some relative work. First of all, Gatys et al discovered that the feature statistics of a CNN effectively encode the style information of an image [[5](#_[4]_Image_Style)], which laid the foundation of neural style transfer. Then,

1. **Methods**

Other types of style features such as the correlations between different channels can be also included in the style vector, discovered by Gatys et al. [[4](#_[4]_Image_Style)]. In the paper, the writers focused on the channel-wise statistics for efﬁciency and conceptual clarity. We try to add the information about different channels statistics and see if we can benefit from it.

By including the feature correlations of multiple layers, we obtain a stationary, multi-scale representation of the input image, which captures its texture information but not the global arrangement.

We used convolution layer with kernel size equals the number of style features. In this case the CNN essentially becomes fully connected channel-independent layer.

Moreover, we try to apply the recalibration module in different stages of the CNN representation.

Our assumption, based on Figure 1, is that this kind of recalibration module would be more efficient when applying it into a low-level features representation.

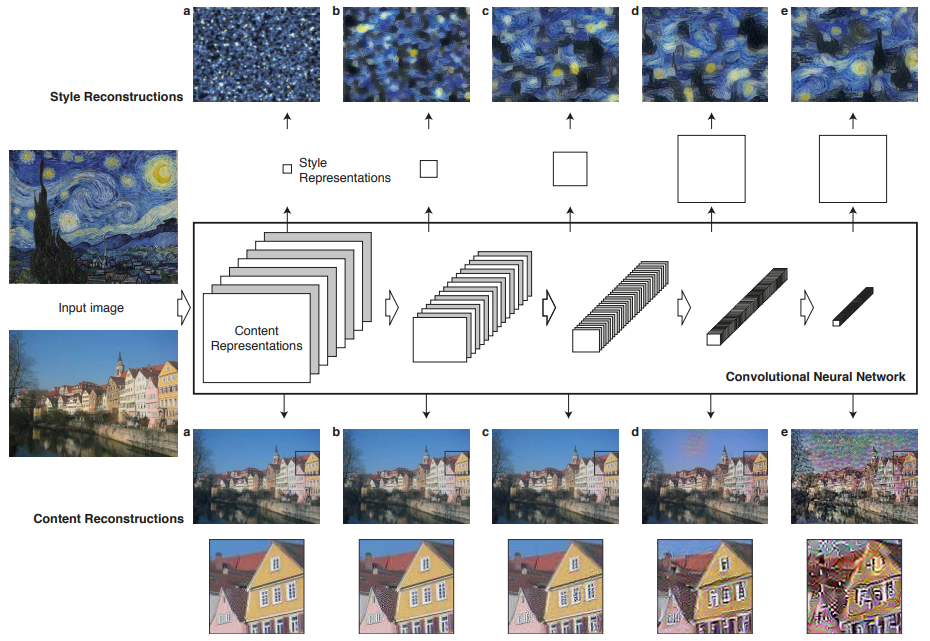


Figure 1. taken from [4]. A given input image is represented as a set of filtered images at each processing stage in the CNN. While the number of different filters increases along the processing hierarchy, the size of the filtered images is reduced by some downsampling mechanism (e.g. max-pooling) leading to a decrease in the total number of units per layer of the network. Content Reconstructions. We can visualise the information at different processing stages in the CNN by reconstructing the input image from only knowing the network’s responses in a particular layer. **The style is most pronounced at low-level features representation.**

Nevertheless, we experimented with the location of the SRM module in the neural net in order to approve this hypothesis.

Another modification to the original SRM model that we tried to experiment with is using the median. As mentioned, in the original paper an image style was defined as the average and std of each of its channels. Intuitively, we can see that this definition is a rough approximation of the image distribution. In order to better represent the image's distribution, we also used the median in addition to the mean and std.

Furthermore, we combined the approach using the feature correlations, and the approach using the median to create a SRM model that uses both the median and the feature correlations (in addition to the mean and std, of course).

We present a style pooling approach which is superior to the standard global average or max pooling in our setting, as well as a channel-independent style integration method which is substantially more lightweight than fully connected counterparts yet more effective in various scenarios.

1. **Implementation and experiments**

In order to verify the effectiveness of SRM, we conduct several experiments using different general object classiﬁcation problems and datasets.

First, we evaluate the performance of SRM on the CIFAR-10 dataset, in comparison with the state-of-the-art method Squeeze-and Excitation (SE).

Later on, we also evaluate the performance of SRM on the ImageNet-1K.

We implemented all competitors to compare under consistent settings for fair comparison.

1. **Results**

Throughout the experiment, SRM outperforms recent approaches though it requires orders of magnitude less additional parameters.

boosting meaningful features while suppressing weak ones,

**References –**

[1] [Recalibrating Fully Convolutional Networks with Spatial and Channel ‘Squeeze & Excitation’ Blocks](https://arxiv.org/pdf/1808.08127.pdf)

**[2]**  [P-CNN: Part-Based Convolutional Neural Networks for Fine-Grained Visual Categorization](https://ieeexplore.ieee.org/document/8789527)

# [3] [Visualising Filters and Feature Maps for Deep Learning](https://towardsdatascience.com/visualising-filters-and-feature-maps-for-deep-learning-d814e13bd671)

# [4] [Image Style Transfer Using Convolutional Neural Networks](https://zpascal.net/cvpr2016/Gatys_Image_Style_Transfer_CVPR_2016_paper.pdf)

[5] L. A. Gatys, A. S. Ecker, and M. Bethge. Texture synthesis using convolutional neural networks. In NIPS, 2015. 1