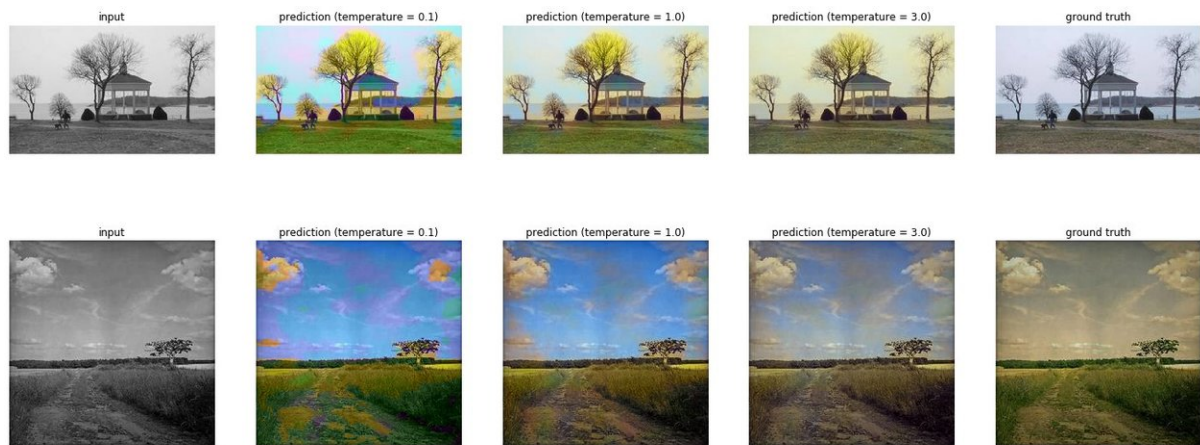# ColorUNet: A new deep CNN classification approach to colorization

November 16 2018, by Ingrid Fadelli

Sample predictions of ColorUNet on the validation set, for bland input images. ColorUNet's output images are more colorful than the ground truth (original) images. The bottom example is an old photograph with worn out tones. Credit: Billaut, De Rochemonteix nd Thibault.
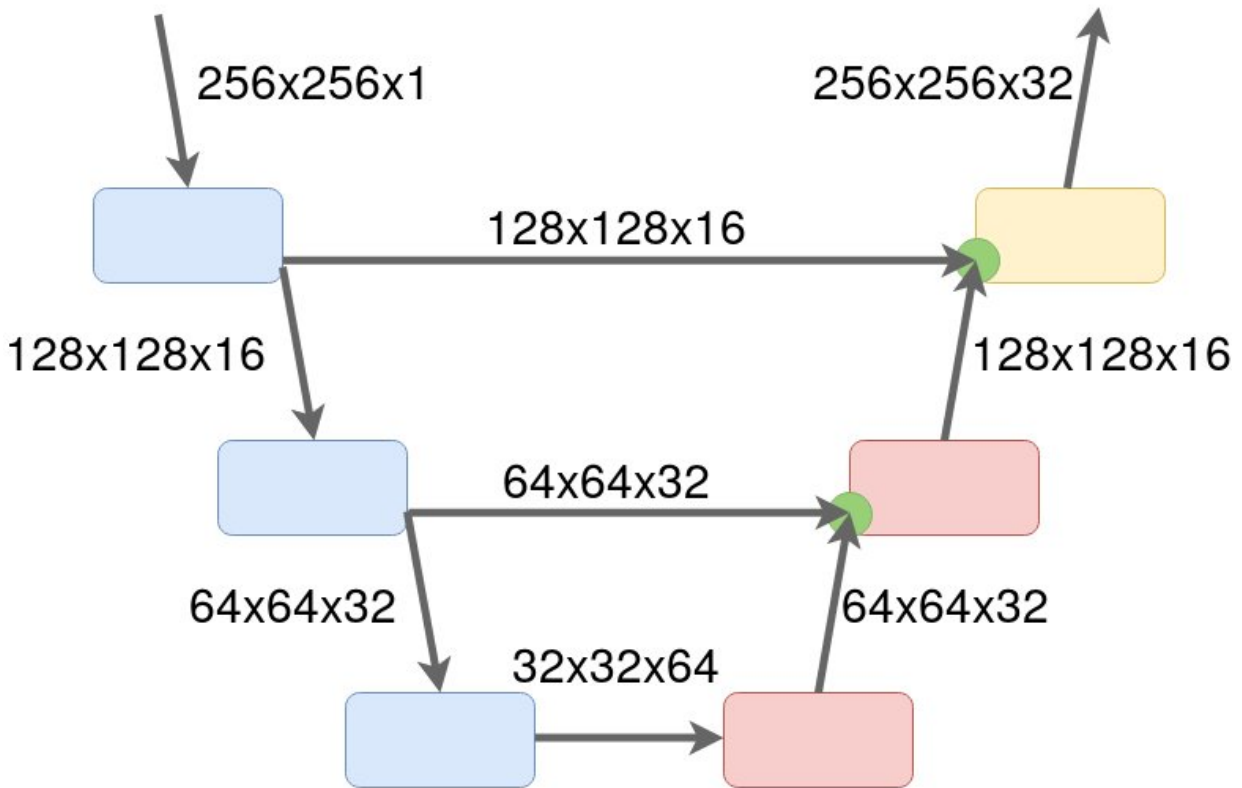
A team of researchers at Stanford University has recently developed a CNN classification method to colorize grayscale images. The tool they devised, called ColorUNet, draws inspiration from U-Net, a fully convolutional network for image segmentation.

"As part of Stanford's Computer Vision class, we worked on this project for several months," Vincent Billaut, one of the researchers who carried
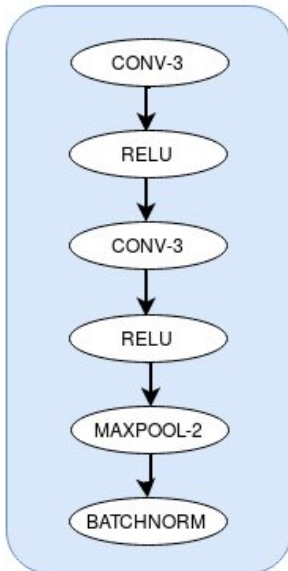
out the study, told TechXplore. "Our objective was to reproduce state-of-the art results using a lightweight model, rather than enhancing existing models by increasing the size of the training set or their computational complexity, a very common approach in CV problems. We wanted our results to be easy to evaluate and visually appealing, because besides useful and impactful applications, CV is also about cool stuff."

Billaut and his colleagues decided to approach the task of automatically colorizing grayscale images from the angle of classification, working with a finite set of color possibilities. Their model followed a loss and prediction function, favoring colorful images over realistic ones.
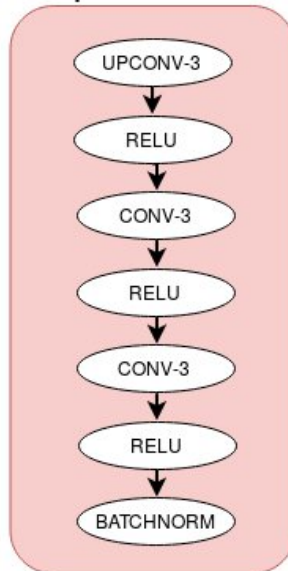
"Instead of trying to predict the colors directly via a regression task, we split all the colors into bins, with a classification task," Marc Thibault, another researcher involved in the study, told TechXplore. "Formulating the problem as a classification task allows us to have better control over how colorful we want our output to look, by fine-tuning how we predict a color from the output of the network."
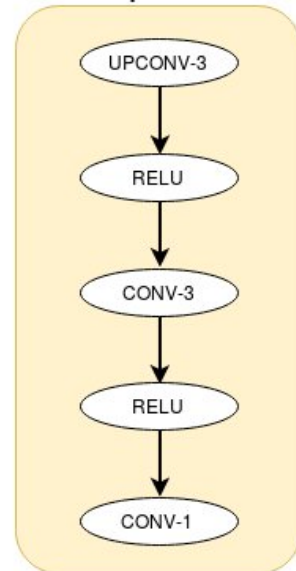
256x256x1

256x256x32

128x128x16

128x128x16

128x128x16

64x64x32

64x64x32

64x64x32

32x32x64

DownConv cell

CONV-3

RELU

CONV-3

RELU

MAXPOOL-2

BATCHNORM

UpConv cell

UPCONV-3

RELU

CONV-3

RELU

CONV-3

RELU

BATCHNORM

Output cell

UPCONV-3

RELU

CONV-3

RELU

CONV-1

CONCAT

The architecture of ColorUNet. Structure of the ColorUNet. The researchers use 3 types of cells: DownConv Cells that use 2 stacked convolutional layers to have a large perceptive field and a maxpooling to downsample the image, UpConv cells that use 1 ConvTranspose Layer to upsample the image and then 2 convolutional layers, and an Output cell that is a simplified version of the UpConv cell. Credit: Billaut, De Rochemonteix and Thibault.

The researchers trained their model on subsets of the SUN and ImageNet datasets, which contain images of landscapes. The neural network architecture they developed allowed their deep learning algorithm to extract both local and global information from each grayscale image.

"The algorithm can then decide on a region's color based on its own aspect, as well as on the context around it," Thibault said. "In general, it is crucial that AI techniques for real-life decision-making leverage both locally precise subject identification and an understanding of the broader context."

One of the key goals of the study was to develop a lightweight architecture that was scalable, but also performed as well as state-of-the-art models in colorization tasks. To achieve this, the researchers limited the task to images of natural landscapes.
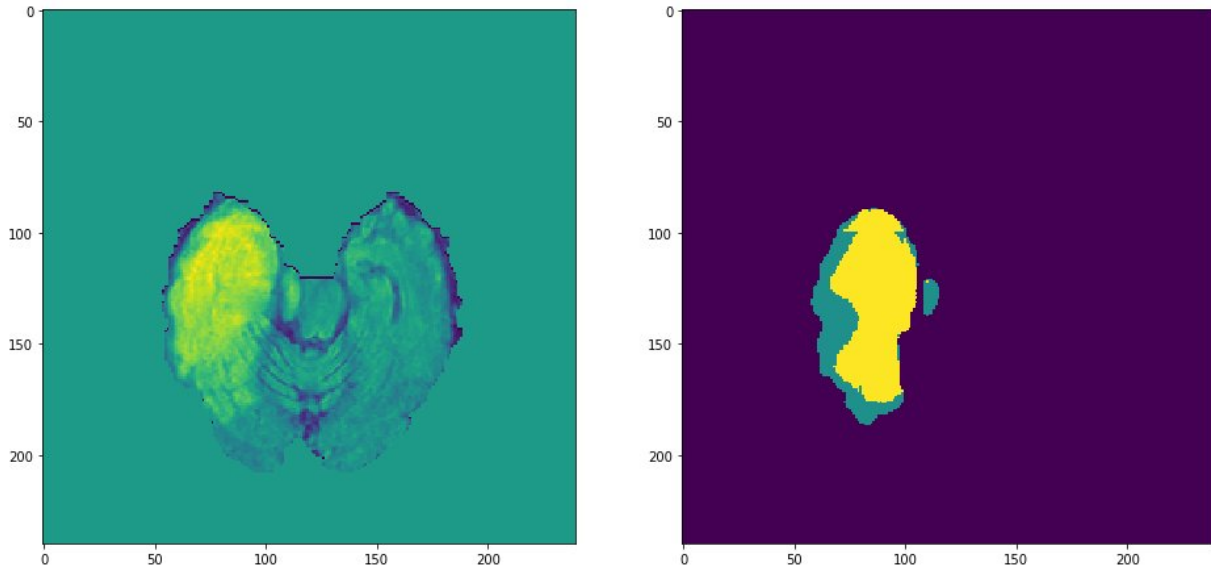
Open-source MRI picture that could be processed by ColorUNet in the future. Credit: Billaut, De Rochemonteix and Thibault.

"Most importantly, we used a U-Net architecture to enhance the performance and reduce the complexity of the model," Matthieu de Rochemonteix, one of the researchers who carried out the study, told TechXplore. "ColorUnet approaches state-of the art performance on the selected subtask. Its architecture allows for faster and more stable training, without trading off the depth and representative power of the model."

When evaluated on pictures of landscapes, ColorUNet achieved very promising results, with data augmentation significantly improving the performance and robustness of the model. The researchers also applied to model to video colorization, proposing a way to smoothen color predictions across frames without having to train a recurrent network for sequential inputs.

"The main contribution of this technique is the ability for an algorithm to understand what is going on in an image on a local scale, by feeding it the whole image's context," Thibault said. "While we showed its efficiency in image coloring, we are also working on other applications, especially in the medical domain. Within the Gevaert Lab at Stanford, we have applied this method to tumor detection for glioma (brain cancer) patients based on MRI scans. Research is flourishing in this field, with more and more CV techniques being applied to medical imaging."

  More information: ColorUNet: A convolutional classification approach to colorization. arXiv:1811.03120 [cs.CV]. arxiv.org/abs/1811.03120