Extra Credit HW- Implement your on DeepDream

Valentin DE BALTHASAR DE GACHEO (M10715803)
Dept. of Computer Science and Information Engineering
National Taiwan University of Science and Technology
Taipei, Taiwan
valentindebalthasar@gmail.com

Keywords— DeepDream, Convolutional Neural Network, Deep Neural Network, Image Generation

I. Introduction

The goal of this experiments is to use three convolutional neural networks architectures to create DeepDream programs. A DeepDream program is a computer vision program created by Google to find and enhance patterns in an image, creating a dream-like hallucinogenic appearance in the input image. The results of two models are presented in this paper, the third model program is available to users who wish to make their own experiments.

II. Method

A. Material Used

1) Data

Every Convolutional Neural Network architectures are trained using the ImageNet dataset. ImageNet is a large visual database designed for use in visual object recognition software research. More than 14 million images have been hand-annotated to indicate what objects are pictured some images (more than 1 million), bounding boxes are also provided.

2) Architectures

In this experiments, three types of architectures are used. The first one is named “AlexNet”. The second and third ones are named “Inception h5” and “Inception v4”, which evolved from GoogLeNet and merged with ResNet idea, with higher accuracy.

3) Libraries

In this project, local environment is created to avoid conflicts between libraries. In addition to that, *numpy* is used to perform math calculations, *partial* submodule is used to create new versions of functions with many arguments, *pillow* is used as image library to modify images, and *tensorflow* is the machine learning library used in this project.

B. Peeking inside Neural Networks

1) Trained Neural Network used as Generator

Once trained, the convolutional neural network are run in reverse. The optimization part resembles *backpropagation*, however instead of adjusting the network weights, the weights are held fixed and the input is adjusted. Therefore, starting from an image, the network is forced to generate an image corresponding to the objects it learned. This usage resembles the activity of looking for animals or other patterns in clouds.

2) Layer enhancement

Starting from *Fig.1*, a picture of me taken last summer, layer enhancement is performed by picking one of the convolutional neural network layers to let the neural network enhance whatever it detects. Each layer of the convolutional neural network deals with features at a different level of abstraction, so the complexity of features generated depends on which layer is picked.

Fig. 1. Original picture

After many trials, the following parameters are chosen and will remain constant during the experiments:

- number of iterations: 47
- step : 4.4
- number of octaves : 4
- octave scale : 1

The image is split into a number of octaves, and the details are generated octave by octave (see Fig. 2):

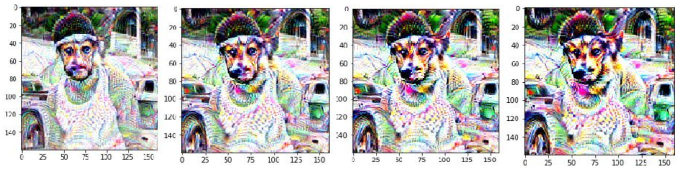


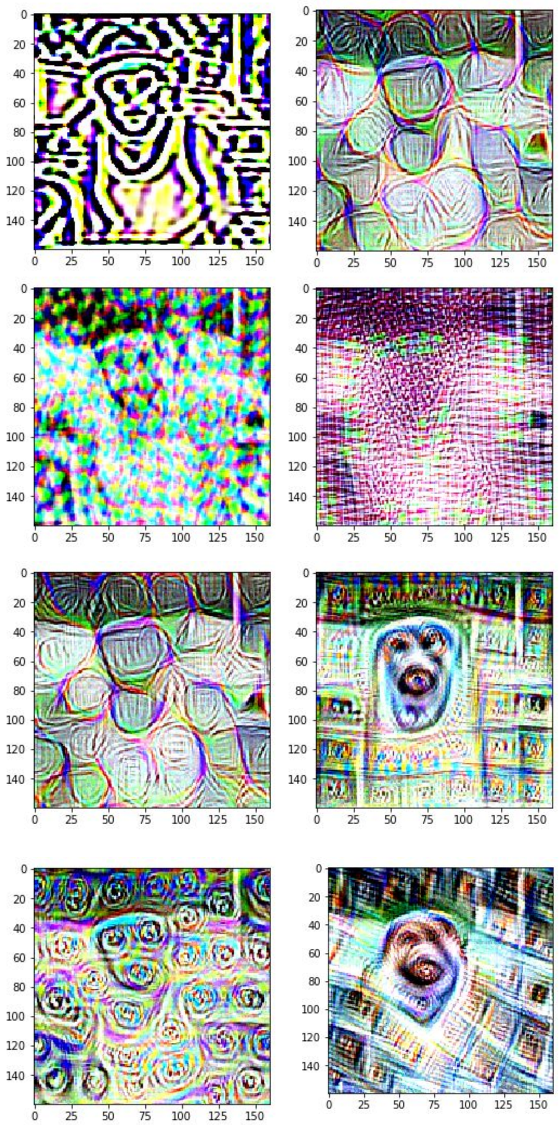
Fig. 2. Details generated octave by octave

The layers' enhancement is performed from low-level layers to middle and higher-level layer.

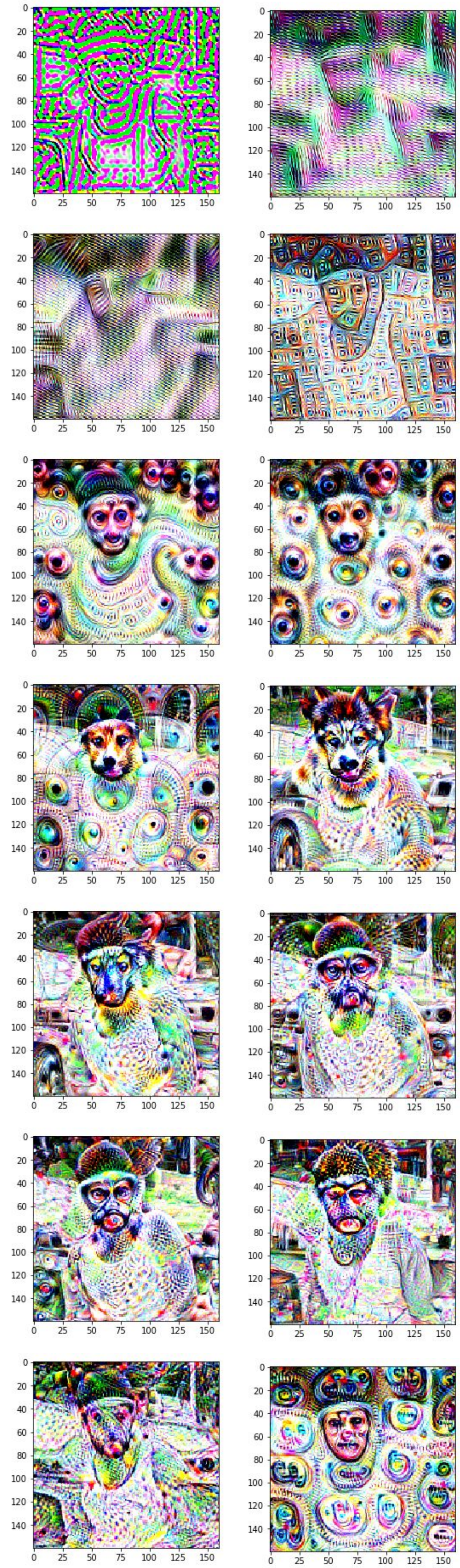
III. Result

A. Layer enhancement: Towards the comprehension of Neural Network features extraction.

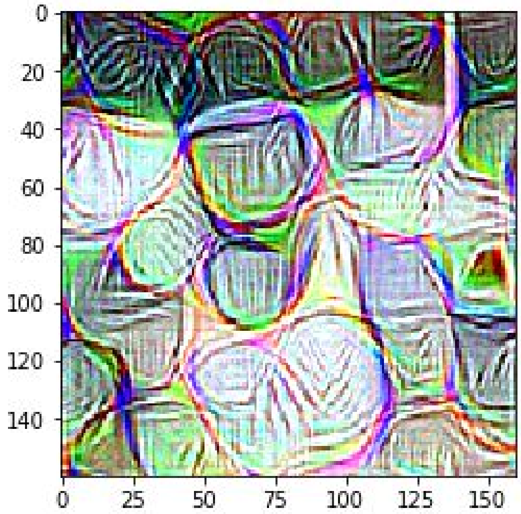
1) AlexNet

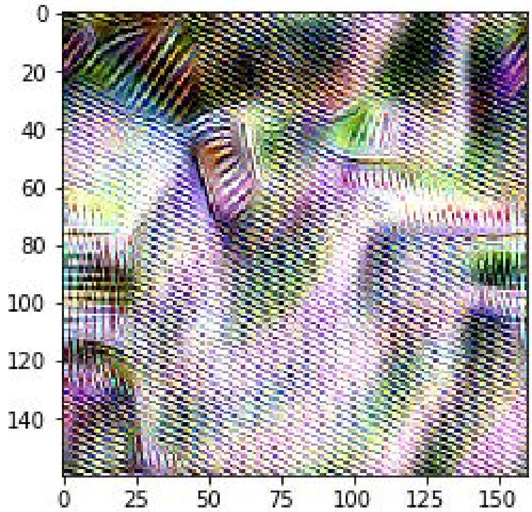


2) Inception5h

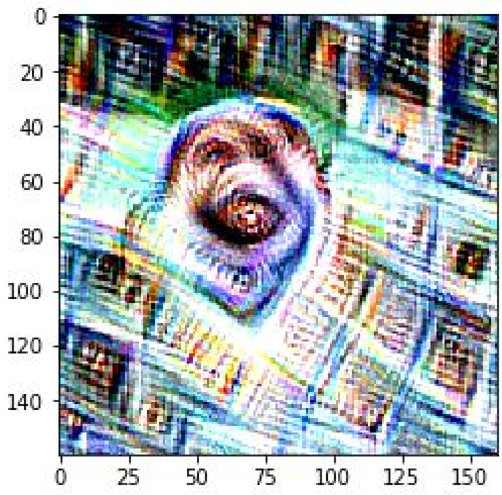


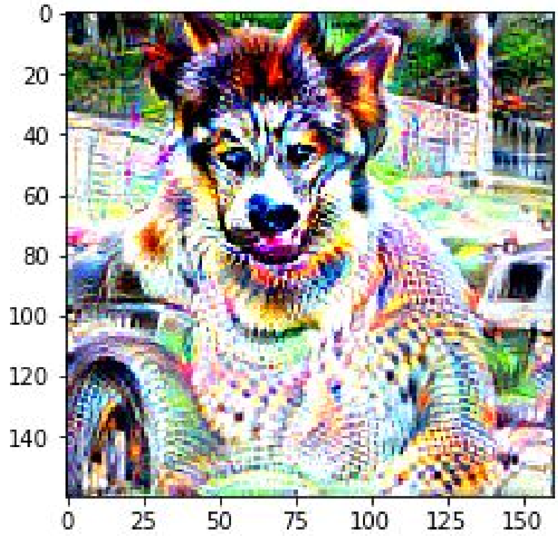
The results, for both the architectures used above, show that by picking lower-level layers, strokes or simple ornament-like patterns are produced, as those layers are sensitive to basic feature such as orientations of edges:

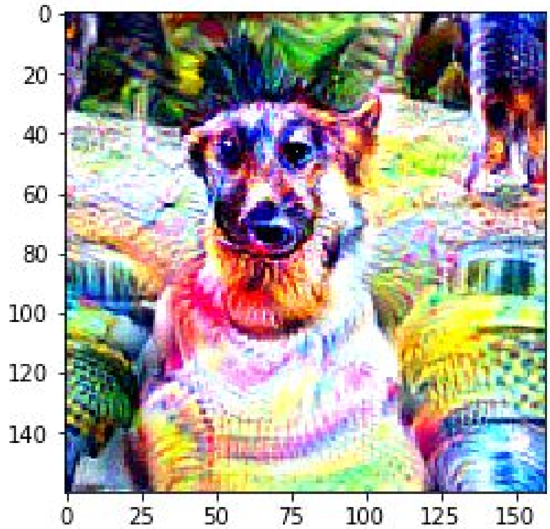


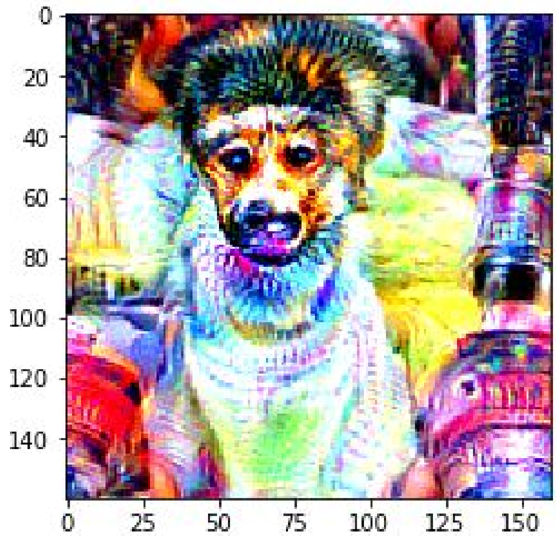


Conversely, more sophisticated and complex features or even whole objects tend to emerge if higher-level layers are picked. In this case, dog-like features emerge, which were previously learned by the neural network during training on the ImageNet dataset.









IV. Discussion

A. Generator: features amplification: a correction visualization tool for training mishaps

It is also possible to use DeepDream algorithm as a visualization tool for training mishaps. For example, if a Convolutional Neural Network is trained to recognize and classify images of smartphones, but most of them are pictures of smartphones in someone's hand, it is likely that the model's prediction on unseen smartphone pictures will not be good for a smartphone which is not located in someone's hand. Therefore, by generating images from the features learned by a Convolutional Neural Network, DeepDream can help visualizing training mishaps, and therefore further improve future predictions, after the necessary corrections are performed, as the programmer can check that the network has correctly learned the right features.

B. DeepDream algorithm, an hallucinogenic drug for artificial neural networks

The often-mentioned resemblance of DeepDream images to hallucinations induced by LSD and psilocybin suggests a functional similarity between artificial neural networks and certain layers of the visual cortex, an issue that will require further study.

Deep learning methods have transformed the world of AI. The abilities that human beings once thought were uniquely our own have begun to be shared with machines. There is no mathematical reason why networks arranged in layers should perform well at these tasks, when compared to other models. Despite the huge success of DNN, nobody is sure how they achieve their success. Henry Lin at Harvard University and Max Tegmark at MIT revealed that the reason why mathematicians have been so embarrassed is that the answer depends on the regime of physics rather than mathematics.

An image of dog or cat consists of a million pixels that can each take one of 256 grayscale values. Therefore, there can be 256^1000000 possibilities. It is necessary to compute whether it is a cat or a dog, and neural networks, with only thousands of parameters manage to perform this task with ease. There are in orders of magnitude more mathematical functions than possible networks to approximate them, and DNNs get the right answer. WHY?

The universe is governed by a small subset of all possible functions. Mathematically, laws of physics can be described by functions that have a remarkable set of simple properties. The number of orders for polynomial functions is infinite, but only a small subset of polynomials appear in our laws of physics, ranging from 2 to 4. Artificial neural networks are based on biological ones. Therefore, Henry Lin and Max Tegmark not only showed the reason why Neural Networks are so good at performing those tasks, but also why our brains can make sense of the universe.

V. Conclusion

The DeepDream algorithm presented here using two different models help us understand and visualize how CNN are able to perform difficult classification tasks, how it is possible to improve network architecture thanks to this program, finally check what the network has learned during training.

It also makes us wonder about the similarity between artificial neural networks and certain layers of the visual cortex.

[1]