

# REMember: A Deep Dream Batch Program for Multi-Frame Formats\*

Madison Blake<sup>1</sup>, Josh Hickman<sup>2</sup>, Axel Hranov<sup>3</sup>, and Matthew MacNeil<sup>4</sup>

**Abstract**—This is an electronic document serving as a proposal for our CS 445 group’s Final Project. This proposal will provide a brief motivation of the project, detailed discussion of the data that will be obtained or used in the project, along with a time-line of milestones and expected outcomes.

## I. INTRODUCTION

DeepDream is a computer vision program created by Google. It finds and enhances patterns in images via algorithmic pareidolia using a convolutional neural network. This program popularized the term (deep) “dreaming” to refer to the generation of images that produce desired activation in a trained deep network. Now this term is used to refer to a collection of related approaches.

## II. PROCEDURE FOR PAPER SUBMISSION

### A. Maintaining the Integrity of the Specifications

It must be one and a half to two pages in IEEE format. This proposal is to provide a brief motivation of the project, detailed discussion of the data that will be obtained or used in the project, along with a time-line of milestones, and expected outcome.

### B. Submission and Presentation

Final project proposals are due October 4th. Presentation dates are not outlined on our class website, but we are eager to present our concept to our peers at the earliest convenience so that we may improve the scope and outcome of our project.

## III. MOTIVATION

Deep Dream has been around for several years, spawning several websites that allow users to upload their own images to have it be processed by a neural network. These websites include various parameters to tailor the image to the user’s desires. The images generated themselves are motivating as they correlate to the original image in a way that is extrapolated and has a life of its own.

Some developers have gone as far as using Deep Dream to process individual frames in a video, producing an animated version of the aforementioned deep dream products. From this arose the motivation to create this project. While there exist on-line tools to process individual images, there is not one to easily batch together frames from videos, gifs, and webms.

Our goal is to create a script using a higher level programming language that will process a given video (in a variety of formats) by grabbing each individual frame, and running it through the neural network, applying the parameters specified by the user.

## IV. DATA

### A. Collection

Collection of data on this project is mostly limited to the image data that the neural network matches the given input on. The network operates on a trained set of data that is used as the basis for the image alterations. Fed an image, the network searches for parts that match the images that it was trained on, merging the sections that are within a specified tolerance. There are quite a few recurring patterns that the most basic settings will implant into an image. These include dogs, birds, spires, eyes, and vehicles. The network itself is not limited to this single selection, and other trainings can be used to produce various effects.

A possible feature that may be integrated into future version is the ability to accept URL’s, allowing a user to input, for example, a Youtube or Vimeo link for which to process. This can have legal ramifications, and will be require further research before it can be implemented.

### B. Uses

The use of the collected data will revolve around replacing minute sections of the original image with the matched images from the neural network. On a pixel-by-pixel scale, these alterations will be hardly noticeable (due to the network finding the closest matches); however, when viewed as a whole, there will be a visible alteration. This new image can then be pumped back into the algorithm for another iteration (either with the same settings or different parameters). Doing this over and over results in a more clearly altered image with patterns that more closely match the trained images. Once the script finishes, the final product can either replace the original source file, or be saved as a separate file with the same file extension. A stretch goal for our product would be to have the ability to convert file formats after processing.

\*This work was not supported by any organization

<sup>1</sup>Madison Blake is a senior in the University of Tennessee Knoxville’s CS department, currently taking CS 445 (Fundamentals of Digital Archeology) with Audris Mockus [web.eecs.utk.edu/~miblake6](http://web.eecs.utk.edu/~miblake6)

<sup>2</sup>Josh Hickman is a senior in the University of Tennessee Knoxville’s CS department, currently taking CS 445 (Fundamentals of Digital Archeology) with Audris Mockus [web.eecs.utk.edu/~jhickman](http://web.eecs.utk.edu/~jhickman)

<sup>3</sup>Axel Hranov is a senior in the University of Tennessee Knoxville’s CS department, currently taking CS 445 (Fundamentals of Digital Archeology) with Audris Mockus [web.eecs.utk.edu/~ahranov](http://web.eecs.utk.edu/~ahranov)

<sup>4</sup>Matthew MacNeil is a senior in the University of Tennessee Knoxville’s CS department, currently taking CS 445 (Fundamentals of Digital Archeology) with Audris Mockus [web.eecs.utk.edu/~mmacneil](http://web.eecs.utk.edu/~mmacneil)

## V. MILESTONES

This program will go through several iterations that will each add functionality:

### A. *Alpha v1*

The first iteration will act as a proof of concept and will be a simple command line driven program that can take a single frame or image and process it using hard coded settings.

### B. *Alpha v2*

The second iteration will add the ability to specify various settings to be passed onto the neural network, which can be supplied as arguments to the program.

### C. *Alpha v3*

This version will allow for additional file formats to be given, while still extracting a single frame to ensure that the program can handle the various formats.

### D. *Alpha v4*

Once the previous version is functioning as needed, the next step will be to allow the program to crawl through the given file and process multiple frames. These individual frames will each be dumped to its own file for testing purposes. The next iteration after this one will push the program into the Beta stage.

### E. *Beta v1*

The first beta iteration will allow for the processed frames to be combined back into the same file format that was pumped in. This version will still be command line based. It is possible that we will only handle one file format type at this stage.

### F. *Beta v2*

The second beta will accept additional file types and ensure that they are processed and reconstructed properly

### G. *Release v1.0*

The initial release version will be the culmination of previous versions, delivering a product that can handle multiple file formats, and process the entire series of frames. This version will still be command-line based.

### H. *Release v2.0*

Ideally, the second release version will be machine-independent, and support use of multiple OS's (namely Windows and Linux).

### I. *Release v3.0*

Future version goals will include creating a GUI to more easily allow the user to select files and set parameters rather than having to deal with command-line arguments

## VI. OUTCOMES

By the conclusion of this project, we would have produced a fully developed batch program to input videos, webms, GIFs and similar files and modify them via the Deep Dream Neural network according to the specifications of the user via a general user interface. This can expand the scope of the Deep Dream neural network in terms of image processing, classification, and identification by applying in the context of multiple frames to develop a more articulate context as to the nature of the video and its particular visual elements. The videos used in the program could also provide a new training set for the furthering development of the Deep Dream neural network and computer vision.

## ACKNOWLEDGMENT

We wish to acknowledge the distinguished University of Tennessee in the glorious city of Knoxville. Our success is due to the honorable Doctor Audris Mockus's guidance and influence who gives inspiration to all in the field of Digital Archeology among the other disciplines of Computer Science.