**README**

**P4P #110 – A Convolutional neural network for detecting visual texture**

**Callan Loomes, Conrad Scherb**

**Prelude:** This project is based off using convolutional networks to model the primary visual cortex in the brain; hence this compendium consists largely of a base of code to generate the datasets, and subsequently train and create models. MATLAB is used to generate the Gabor patch datasets and for user labelling. Tensorflow and Keras are used to create and train the CNN models; this is best installed and run through an Anaconda environment. Instructions for this are located in the relevant directory. Furthermore

This compendium contains some datasets of user labelled images, and some models that have already been saved and are available for use. The tools to create these however are easily accessible; care should be taken with regards to the current filenames used in the code, as there was some reorganization of the GitHub that may mean some file paths need updating.

**File Structure:**

An overview of the file structure is included below; each sub file also contains its own readmes. All code is commented sufficiently, and should be easy to parse:

* **Reading in existing data:** Contains MATLAB scripts for extracting existing data provided by Luke Hallum into images and an organized file structure that can be interpreted by tensorflow. A link to this data can be found on the following google drive for reference or download (https://drive.google.com/drive/folders/1szKk69YlkZZxvbV33aNGMd7VrNlqDKHh?usp=sharing). This data was given as an initial ground for testing, however, because there was not enough data in this subset, our own data was generated. This data was not used much for training models.
* **Generating Datasets:** Contains MATLAB code for generating datasets, including vertical, diagonal and horizontal boundaries. Also contains script for user labelling of datasets. Also contained in this file are a large collection of datasets, already ready to use for training. These include over 7500 user labelled images ~ recommend you use these for training and understanding process before labelling user data on your own.
* **CNN design and Creation:** Contains all python scrips required for altering CNN models and performing 10-fold cross validation. Also contains ability to save models and graph individual training sessions. Also within this file are a collection of already trained and saved models, which can be used for visualization of kernel output.
* **Results of testing:** Contains a spreadsheet of data outputs for the majority of models tested, excluding some initial exploratory content. Contains graphing spreadsheets, and also graphs showing training over each individual fold.
* **Visualization of Models:** Contains python scripts for visualization of Kernel output and first convolutional layers. Use the same Anaconda environment for running scripts in this folder, models from CNN design and creation can be used as examples.
* **Notes on CNN Biology:** Contains some initial notes relating what we have found into models and how it relates to the current biological literature. More concise and refined notes can be found in the student reports.
* **List of Important Papers:** Contains a variety of useful papers that can be used as a base for understanding the current literature surrounding the topic. These are the most important papers highlighted from the reports.
* **Student Reports and Submissions:** Contains all deliverables, including student reports, poster, midyear reports lit review, presentation etc.
* **Figures and Plots:** Contains a variety of figures and plots used in the reports and demonstrations.
* **Logistics and meeting minutes:** Contains some information regarding the logistics of the project, including an overall timelines.