PY895 Project: Can AI remote view?

Zhijie Feng

*Physics Department, Boston University.**

(Dated: January 23, 2021)

This is a project report for the course PY895 machine learning for physicists, Fall semester 2020. Remote view is an proclaimed supervision-like ability of seeking information about an distant and unseen object by mental sensing. In remote viewing terminology, these objects are called target with randomly assigned code called the ID of the target, and successful remote view should result to some features of the photo. I found this procedure related to feature extraction of CNN, and thus it is an interesting exercise to experiment with the possibility of "teaching AI to remote view". In this paper, we propose a possible learning mechanism for remote view to seemingly work and test it with suitable supervised machine learning algorithm. It turns our proposal may be wrong, but the result reveal another possible mechanism.

I. HYPOTHESIS

We start by assuming some implicit cheating of the game as a possible reason for people to feel that they successfully learnt remote view after some training. If there is no information in the ID, the ID should be totally random, then it will be real superpower for someone to map from this ID to the correct image. However, it may be that the ID are not generated, but chosen to contain information about the image. From where we got our data, the ID contains 4 letter followed by 4 numbers. The total possible outcomes is $24^4 x 10^4 = 331,776,000$. The Shannon entropy is $log_2(331,776,000) = 28.3056$, which means it contains the information around 28 yes or no question. It is more than enough to provide enough description of "features" in the hidden images, that can be counted as successful remote view. For example, it is rough" or smooth, cold or hot, straight or curvy. Moreover, the advocates claims that the target images have to be complex enough, which means they usually contains a lot of features, so it may give the impression that remote viewing is getting many of features correct.

II. DATA COLLECTION

We obtain data from a website developed by a remote view advocate https://zvarik.cz/en/arv-remote-viewing-online-tool-randomizer. Since each click on a button or refresh of this website gives one labeled data- a ID with the image, we obtain 8689 data by using very basic web crawler technique from "request" package of python. The images are in jpg format and varies in size, and the ID are the file names.

* zjfeng@bu.edu

III. METHOD

To match a picture with text, we need to both extract information from the picture and the text. Therefore, two feature extraction process are required, and a mapping two result will be trained.

For simplicity, we use pre-trained network for image recognition to extract the features of image, and perform a clustering of features. The lawful way is to perform a k-means clustering on the training data set, then the test data set can be classified simply by finding the nearest neighbour of cluster. An corrupted way is to perform any clustering algorithm on all data, which mimic a person who have seen all possible pictures for remote view.

Then we treat the cluster as labels of the ID of the images, so that we can train another network to classify the "target ID" by the features. Theoretically, it should work better if there are a lot of feature classes, each filled evenly by both the images and ID. But we only choose two clusters as a prototypical model. For instance, we can correctly group the photos and corresponding ID by features, the probability of guessing the correct photo from the cluster should be doubled compared to guessing from all photos.

Finally, we can test the performance of classification, either by testing the classification accuracy on the test data set, or by an humanly way of giving machine a ID and two choices of photo, an let it determine which one is the correct one. The latter is the mimic version of a actual remote view training game called RV tournament, which can be found on Google Play/Apple store. The developer of that game claims that the average percentage of getting the correct photo is around 60% and use it as a justification that remote view is real.

A. Image feature extractor

We used pretrained networks without final layer as the feature. In this project we have chosen ResNet50 for it's performance on image clustering. The values of feature elements distributed sparsely, thus we use singular value

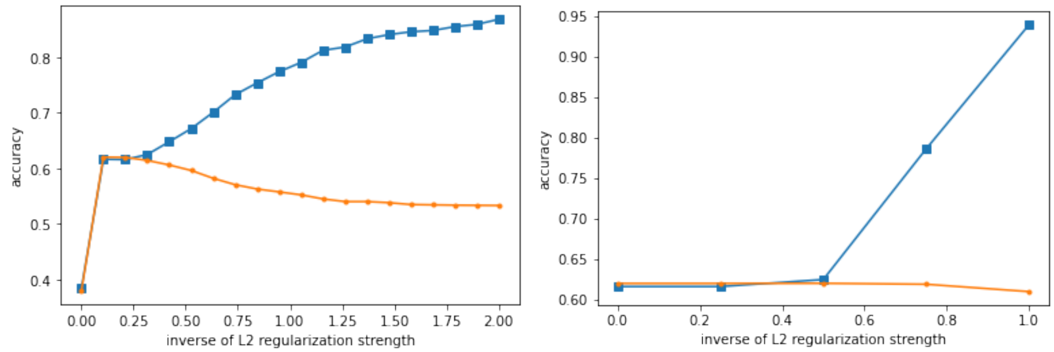

FIG. 1. The performance of logistic regression (left) and support vector machine (right) as text classifier with different regularization.

decomposition (SVG) to reduce the dimension of feature vectors. Then, we perform a simple Kmeans on the dimension reduced feature vectors to group the photos into two clusters. Kmeans is chosen to avoid manipulation of test data. As we can measure test data's distance from Kmeans directly, but not with other algorithm.

B. Text feature extractor and Classifiers

We then formulate the text-image matching problem as a text classification task, identifying the correct cluster that the image linked to the ID belongs. To feed the ID into a classifier with more information preserved, we first separate each ID to their "character N-grams". We only used up n=2 here: "unigram" and "bigram". For instance string "ABC" gives tokens "A, B, C, AB, BC".

We used three different common classifiers of text to compare their performance, including logistic regression, support vector machine and recurrent neural network of output dimension 64. The label of each ID are the cluster that the corresponding image belongs to.

IV. RESULT

The three classifiers, with optimized regularization, all gives around 62% as the accuracy on test data. The performance of logistic regression and support vector machine for different strength of L2 regularization is shown in Figure 1. For recurrent neural network, it arrive around 60% test accuracy in the first two epoch, we did not perform hyperparameter tuning due to computational power restriction. It is better than random guessing that leads to 50% accuracy. However, upon closer inspection, we notice that the ratio of the clusters' size is also around 62%, and the decisions that the classifiers made are simply guessing that all test data belonging to the larger cluster. This shows that the hypothesis we made may be wrong, as it seems that the text cannot provide information about the picture, or at least our feature extractors are not capable to detect it. And the classification could not give better than guessing result for a game of choosing correct from several choices since that total probability of guessing correct is still $0.4\times0.4\times0.5+0.6\times0.6\times0.5+0.4\times0.6\times0.4+0.6\times0.4\times0.6=0.5$.

Nevertheless, our neural networks are indeed trained to reveal a somehow astrology-like strategy of giving some features, in both the machine learning and remote viewing sense, that are most likely to be correct for any photos. Translated to human strategy, it is simply browsing over some of the pictures of remote view tutorials, and making use of one's experience in the visual world, describe some common features shared by the most photos, and always choose the picture that seems more "usual".

V. CONCLUSION

Some people believe they can successfully learn remote viewing, a proclaimed super-power of many, who believe they can describe features of unseen images that are labeled by random text. We made an hypothesis that they obtained hidden information from text that they claimed to be random and uncorrelated to the image. In this project, we build neutral network as text classifier, and made use of transfer learning from ResNet50 as image feature extractor. The best result that the classifier can to is to classify pictures with more common features and less common features. Although we could not find evidence that our hypothesis is correct, we find out that the machine learning results is somewhat like an astrology-like mental trick of giving general descriptions that are most likely to be true for any images.

VI. FURTHER INVESTIGATION

As a check for our model, it would useful to tailer made some dataset to see if the model can give better result that make use of the information of the text. For instance, we could name the file of pictures of cats and dogs with some related keywords and let the classifier decide from the file name if the picture is cat or dog. Several improvement could be made including performing hyperparameter tuning, and increase the number of clusters. A alternative way to test if there is correlation between the file name and image in our data is to define a mutual information measure between text and images check its value.