



Python Libraries for Image Recognition

Sklearn & TensorFlow are used for image recognition of different complexity
Team 7

Our Motivation

- As future data analysts, we want to be able to master mainstream data analytics tools especially machine learning libraries.
- Among all options out there in the market, scikit-learn is the most easy-to-use and general-purpose machine learning platform in Python.
- We will take this opportunity to start out our ML journey by implementing a primary image clustering task using sklearn's packages. As a supplement, TensorFlow's advanced image recognition package will also be introduced.

Our Hypothesis

- TensorFlow can outbeat sklearn in Minnesota scenery image recognition since TensorFlow supports more complex computation.

Our Dataset

- 2481 pictures of 'Minnesota natural scenery' parsed from Google Image
- The raw images are analyzed through two approaches: RGB channels under Sklearn and convolutional neural network under TensorFlow. The latter supports more specific classifications..

Our Testing

Sklearn's image clustering using RGB features

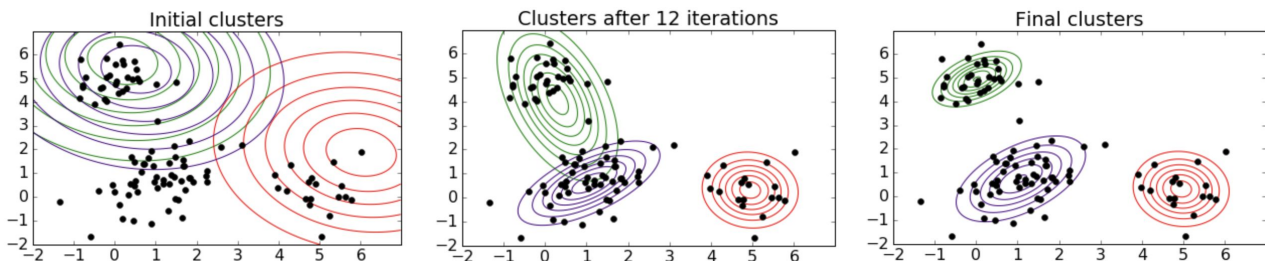
K-Means

- Package
 - `from sklearn.cluster import KMeans`
 - `KMeans(n_clusters = n)`
- In a 3D space with R, G, B axis, initialize K centroids and cluster all data points around the nearest centroids based on Euclidean distance.

- K-Means lacks accuracy when a certain data point is plot between cluster boundaries. It can only make hard assignment on each data point, while the data point may contain information which belong to multiple clusters.

Gaussian Mixture Model

- Package
 - from sklearn import mixture
 - mixture.GMM(n_components=4, covariance_type='full')
 - multivariate_normal.pdf(data, mean, covariance) * weight
- A 3D space with R, G, B axis is not enough anymore for GMM. A 4th dimension denoting probability is needed. Initialize N multivariate Gaussian distributions and adjust parameters of μ and Σ using EM algorithms, until each distribution captures the maximum likelihood of data points.
- GMM provides soft assignment for each picture. However, as number of features increase, data points may mix together thus making multivariate distribution of any kind fails to capture clusters.



TensorFlow's CNN image classification using Inception v3 model

- TensorFlow is an open source software library for machine learning written in Python and C++. Inception is a pre-trained 22 layers deep neural network built on TensorFlow used for recognize day-to-day object. Inception-v3 is trained for the ImageNet Large Visual Recognition Challenge using the data from 2012. By training on a 1000 categories labeled dataset ImageNet, Inception has high accuracy in image recognition.
- We use the pre-trained Inception v3 model to recognize object. The model gives the output of the probability of an object is in the image by using softmax regression. CNN can generate different filters (also known as neuron or kernel) through the training process and use them to create activation map or so-called feature map. The features that a CNN can recognise is determined by the numbers of layers and the number of filters in each layer.
- CNN can also perform non-linear algorithm by using ReLU function in the convolution layer, which couldn't be performed by GMM model and K-means model.

References

- [1]. 2.1. *Gaussian mixture models*. Retrieved August 13, 2017, from http://scikit-learn.org/stable/modules/generated/sklearn.cluster.k_means.html#sklearn.cluster.k_means
- [2]. *A Beginner's Guide To Understanding Convolutional Neural Networks*. Retrieved August 13, 2017, from <https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/>
- [3]. Ajay, (2017). EMOTION DETECTION USING MACHINE LEARNING. *International Journal of recent trends in engineering and research*, 2455-2457, 28-32
- [4]. *Image Recognition*. Retrieved August 13, 2017, from https://www.tensorflow.org/tutorials/image_recognition
- [5]. *Python-OpenCV 处理图像（五）：图像中边界和轮廓检测*. Retrieved August 13, 2017, from http://blog.csdn.net/qq_26898461/article/details/50454547
- [6]. *Rectifier (neural networks)*. Retrieved August 13, 2017, from [https://en.wikipedia.org/wiki/Rectifier_\(neural_networks\)](https://en.wikipedia.org/wiki/Rectifier_(neural_networks))
- [7]. *sklearn.cluster.k_means.()*. Retrieved August 13, 2017, from http://scikit-learn.org/stable/modules/generated/sklearn.cluster.k_means.html#sklearn.cluster.k_means
- [8]. *Softmax function*. Retrieved August 13, 2017, from <https://zh.wikipedia.org/wiki/Softmax%E5%87%BD%E6%95%B0>
- [9]. *Understanding Convolutional Neural Networks for NLP*. Retrieved August 13, 2017, from <http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/>
- [10]. *Visualizing what ConvNets learn*. Retrieved August 13, 2017, from <http://cs231n.github.io/understanding-cnn/>
- [11]. *卷积网络*. Retrieved August 13, 2017, from https://exacity.github.io/deeplearningbook-chinese/Chapter9_convolutional_networks/
- [12]. *浅谈深度学习中的激活函数 - The Activation Function in Deep Learning*. Retrieved August 13, 2017, from <http://www.cnblogs.com/rgvb178/p/6055213.html>