Applying Machine Learning Technique to Bird Species Classification and Beyond

* Introduction
  + This project aims to design and construct automated bird monitorization station that help estimate the population of bird species without human intervention. Through the research, I spent first half of research time familiarizing myself with the project by experiencing the “BirdID” tool developed by previous student and increased its accuracy to 94%, and the second half of research time exploring a more popular approach, Convoluted Neural Network(CNN) in image recognition field
  + In order to achieve the goal, the task can be listed in following steps
    - Acquiring data
      * We need fixed cameras at interested zone towards feeders, where images can be captured automatically when camera sensors sensed the arriving of birds
      * Then raw picture data are collected and classified manually based on bird species in order to prepare for the training stage.
    - data preprocessing
      * This step aims to transform the image data into desired form
      * Some technique used in this step includes image segmentation, data augmentation and green background subtraction
    - Feature extraction
      * This step starts from an initial set of pictures and aims to identify find the most interesting clusters of points (ideally, the edge of the bird)
    - Fitting the extracted feature into Supervised learning model
    - Post processing and accuracy analysis
* Step One: Acquiring data and data preprocessing
  + As our first challenge, we were presented with a set of images of different sizes, hues, lighting conditions, we want to normalized those feature input in order to eliminate those noise in training stage.
  + convert the color image into grayscale
  + Standardize the images into fixed scale such as 128x128 pixels
  + Main Component Analysis
  + Data Augmentation
  + Green Background Subtraction
* Step Two: Feature extraction
  + Mainly rely on matlab computer vision system toolbox
  + It contains several feature detection methods including SURF, MSRF, HOGFeatures
  + By trying one or combining several feature detection methods together, I end up rely on SURF and MSER
* Step Three: Training the bird recognition model
  + BirdID - SVM
    - Support vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. In its simplest term, an SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. Examples of type A then are being mapped to one side of the gap and examples of type B are mapped to the other side of the gap
    - Since we need to separate not just two types of birds but 10 types, we need to use multi-class SVM, which we build binary classifiers which distinguish between every pair of classes.
  + BirdID\_lasagne - CNN
    - Convoluted Neural Network takes a different approach. Inspired by neuroscience, this approach transform each image into input layers to obtain a better representation of the original image.
    - Convolutional Layer
      * The Convolutional layer is the core building block of a CNN. The layer's parameters consist of a set of learnable filters, which have a small receptive field, but extend through the depth of the input volume. During the forward pass, each filter is convolved across the width and height of the input volumn, computing the dot product between the entries of the filter and the input and producing a 2-d activation map of that filter.
    - Pooling Layer
      * In the simplest term, CNN’s pooling layer is a form of non-linear down-sampling. Pooling partitions the input image into a set of non-overlapping rectangles and for each subregion, outputs the maximum. The intuition is that once a feature has been found, its exact location isn't as important as its rough location relative to other features.
    - ReLU(Rectified Linear Units) layer
      * It is a layer that applies activation function f(x) = max(0, x). It is believed that this equation would help increase the nonlinear properties of the decision function and of the overall network without affecting the respective fields of the convolution layers.
    - Fully Connected Layers
      * After several convolution, max pooling and ReLU layers, this layer aims to connect all activations in the previous layer into one matrix in order to prepare for the loss function measurement
    - Loss Layer
      * This layer specifies how the network training penalizes the deviation between the predicted and true label through a loss function.
    - Some of the layer pattern can be following
      * INPUT -> FC
      * INPUT -> CONV -> RELU -> FC
      * INPUT -> [CONV -> RELU -> POOL]\*2 -> FC -> RELU -> FC Here there is a single CONV layer between every POOL layer
      * INPUT -> [CONV -> RELU -> CONV -> RELU -> POOL]\*3 -> [FC -> RELU]\*2 -> FC
  + Decision Tree
* Step Four: Post processing and accuracy analysis
  + For both SVM and CNN, we use same set of image data with different layer patterns/ different patterns to test out accuracy and look for the best set of accuracy
* Research Extended
  + From Image Classification to Walmart Trip Type Classification
  + After successfully tackled problems of image classification, I shifted my interests into another type of classification problem mainly dealing with data rather than pictures
  + Kaggle’s Walmart Trip Type Classification Problem
  + Kaggle is a platform for predictive modeling and analytics competitions on which companies and researchers post their data and data miners from all over the world compete to produce the best models. Hosted by Walmart, this Trip Type Classification Problem aims to improve customers’ shopping experiences by segmenting their store visits into different trip types. Basically, Walmart provides 65 millions lines of data about existing customers purchasing history and trip type and want the researcher to train models based on those data in order to predict future customers trip type based on their behaviours.
* The steps
  + Acquiring data
  + Data munging
  + Fitting models
  + Model evaluation

