RC4 Internship Program Task

Week 1

Jayeon Kang

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Task 1: Introduction to ML**

**a. What key differences can you identify between linear regression and artificial neural network ML techniques?**

Key difference between linear regression and ANN ML is Universal approximation theorem. In other word, linear regression is only used when given variables have linear relationship. However, ANN can be applicated for approximation of any continuous function.

**b. What are the similarities and differences between logistic regression and artificial neural network (ANN) techniques?**

Logistic regression and ANN with sigmoidal activation function use P(y|x) = f(x,a) to predict output. That is, two things are derived from statistical learning theory. Additionally, interaction term in logistic regression and hidden neurons in ANN makes prediction function (output) non-linear. However, logistic regression return prediction as 11 + e-(a・x) but ANN return prediction as 11 + e-(a・β + β0). ANN’s hidden layers makes difference between both prediction function form and more flexible than logistic regression.

**c. How does ANN differ from convolutional neural network technique (CNN)?**

Big difference between ANN and CNN is existence of convolutional layer. Convolutional layer is a kind of filter layer and can be shared for given spatial location. Existence of convolutional layer makes different feature learning way between ANN and CNN. ANN without convolutional layer must compute all of values in images but CNN with convolutional layer just compute only the convolutional layer region. In addition, characteristic of sharing the layer gives advantage of finding same patterns in different location CNN. However, ANN’s filter does not because the layer is fixed.

**Task 3: Introduction to OpenCV and Computer Vision**

Our images with black and white color represent values range from 0 to 255 in matrix. Thus, as adjust each element in matrix, images can be manipulated. In the case of color image, there are 3 channels: Red, Green, and Blue. Combination of these channels, image can have color. If more analyzing the color of image through OpenCV, by using . split(), multi-channels are spited into single channel. If visualizing each image, they are represented as black and white color. As interpret those color, the more black color in given region the less color in same given region of original image. That is, these colors are interpreted as how much correspond color is in the image. Color can be descripted many ways. The representative ways are Hue Saturation Value (HSV) and Red Green Blue (RGB). RGB make color image through the combination of each color value in the image. HSV also express color through 3 values but has difference perspective from RGB as representation of actual colors (Hue), intensity of color (saturation), and how the relative lightness or darkness of a color (Value) are sued. So, if convert image descripted value (color space) in OpenCV, through applying .cvtColor(), color space can be changed.

Image manipulation can be specified into 3 things: Cropping, Resizing, and Flipping. First, to manipulate image, accessing and treating pixel is essential. Images is represented from matrix filled with numbers (pixel). Row and column in matrix give information about location of element, which allows to access pixel. Thus, after appointed pixel from the information, the pixel can be adjusted by some calculations. As a result, manipulation of image can be done. Based on this approach, 3 things of image manipulation are proceeded. As extracting range of row and column to get desired region, image can be cropped. Control of scale of x and y determine resize of image. Without the location information, just only using code (-1, 0, 1) make flipping image. Not only the manipulation but annotation can be provided from OpenCV, also based on the matrix location information. If information of start point and end point is added into argument, line can be drawn. Circle, rectangle, and added text also does in same way.

To treat matrix element in image, brightness and contrast are also adjusted. As subtracting and adding value from all values in image, brightness of image can be controlled. Meanwhile, multiplying value with all values in image makes different contrast. To enhance image, masking and threshold are used. Masking is editing and processing images following condition like or, and, and xor to protect specific area of an image. On the other hand, threshold segmentate image to change the pixels of an image to make the image easier to analyze. In real life, threshold is used to approach binary mapping for image.

Furthermore, OpenCV can align distorted image, which is called image alignment. Image alignment is the technique of warping one image to line up completely with another image. The warping is called as homography. More detail, OpenCV get keypoint and descriptor as feature extraction from converted imaged into grey scale. Descriptor find best match between them. The matches are ordered, and not good matches are removed. Then, top matches are selected and used to find homogrpay. Using homogrpy function in OpenCV, distorted image is warped. Some images have that the brightest part is not seen or over-captures the darkest part. To solve this problem, High Dynamic Range (HDR) imaging is appeared. HDR imaging uses some number of aligned images with difference exposure and features. Estimating average of intensity of each image about each pixel, image is merged into HDR linear image. As a result, proper image can be returned.

OpenCV can be applied into not only image but video tracking. Tracking is the estimation of location of object and prediction of its location and feature point over time. To do this, there are two models used. First model is motion model to predict location and motion over timestep of video frame. Second model is appearance model to predict objects’ appearance. In the case of OpenCV, tracker class is used like MEDIANFLOW, GOTURN, and so on. GOTURN is based on neural network. This tracker train current and previous frame and return output about tracking current frame. This helps the model that can have higher level of description of object.