# Project Deliverable 3

1. Final Result

In the *Deliverable 2*, the accuracy of my model is 51.91% with a vanilla Neural Net, which is still not satisfying for a four-class classification problem (the chance of random guess is 25%). In fact, the "normal" neural net is not suitable for image classification in practice. The images we deal with in real life are very large, and therefore a fully connected network will be very expensive with millions of parameters. Indeed, it is also not necessary to connect every node in the neural net. For a natural image of a rose, the pixel of a petal is correlated to nearby pixels of this petal, rather than the pixels of a bee at the edge.

Therefore, I switched to Convolutional Neural Networks (or CNNs), which could turn the input images into a set of features that is more efficiently interpreted by the neural net.
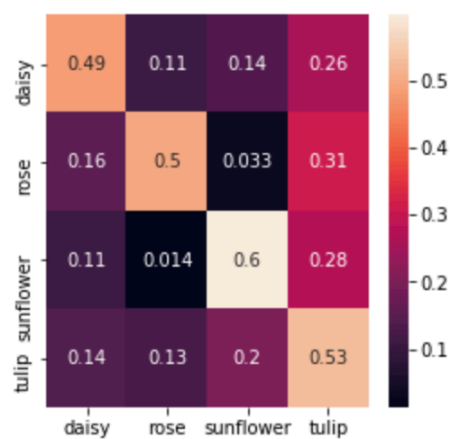
As a result, I achieved an accuracy of 77.19%, improving the performance by 48.70%.

2. Analysis of Result

2.1 Confusion matrix (Graph 1)
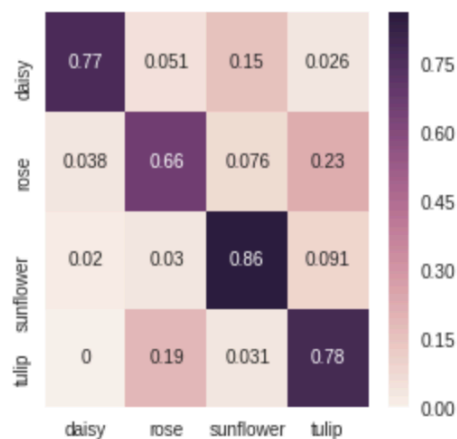
Graph 1: CNN

Graph 2: NN

In the *Deliverable 2*, I found the precision of prediction with tulips is relatively low compared to others. In particular, other flowers are misclassified as tulips with an average probability of 28.33% (Graph 2), which is abnormal. Thus, I checked the test set of tulips and found that many landscape photos of Amsterdam were included in the set. These unwanted pictures might be a consequence of collecting images from Google and require to be deleted manually.

The confusion matrix of my model shows that probability of a misclassification is low in most of time.

2.2 Precision, Recall & F1 Score

	precision	recall	f1-score	support
0.0	0.92	0.77	0.84	78
1.0	0.73	0.66	0.69	79
2.0	0.81	0.86	0.83	99
3.0	0.63	0.78	0.70	64
micro avg	0.77	0.77	0.77	320
macro avg	0.77	0.77	0.77	320
weighted avg	0.78	0.77	0.77	320

- Precision: The question this metric answer is: Of all flowers that labeled as a particular type of flowers, how many are actually this type of flowers? High precision relates to the low false positive rate. I have got 78% weighted average precision, which is pretty good.

- Recall: The question recall answers is: Of all the flowers of one particular type, how many did we label? I have got 77%, which is also acceptable.

- F1 score : F1 Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. I have got 77% for this metric.

2.3 Comparison

By adding more layers (from 2 to 4), I successfully increase my model's accuracy from around 63% to 70%. However, I came across the problem that the loss and accuracy of validation set begin to remain stable no matter how the loss and accuracy of train set change. So I added 3 dropout functions to avoid the problem of overfitting, resulting an 7% increase in accuracy.

2.4 Future Improvement

Preprocess the image by distinguishing the flower from background using OpenCV.
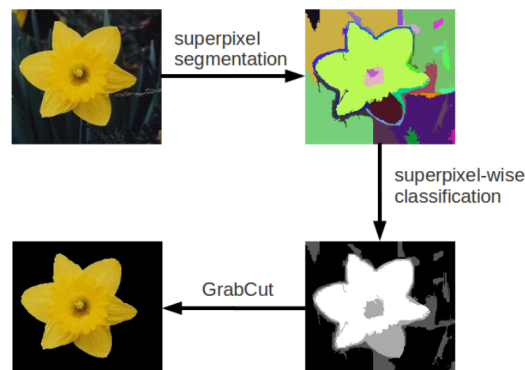
Source: Yuning CHAI, Victor LEMPITSKY, Andrew ZISSERMAN. BiCoS: A Bi-level Co-Segmentation Method for Image Classification. ICCV, 2011.

3. Application

I have already built an simple web app using Flask. Users can upload their images to the website and then get the prediction. I'm still trying to beautify the webpage using CSS.

Furthermore, I plan to combine this classification model with an NLP model. Then, users can get a short poem by AI based on the image they upload. (This is only a general idea! Something I could work on in summer!)