Project Plan

1. **Team Members:**
   Chen Shen, Yaochen Liu

2. **Project Title:**
   Object Detection Based on FCN

3. **Project Abstract:**

As is known to all, object detection and segmentation is one of the most important part in Computer Vision. Meaningfully, to obtain the best performance, profoundly involving deep Learning especially convolutional neuron network becomes a must.

In this project, a method of detecting and labeling object from a given photograph such as means of transportation, constructions and animals will be designed and improved based on the Fully Convolutional Network (FCN). Since developed, FCN has been intensively utilized owing to its advantages in acceptability of images with any size and efficiency compared with conventional CNN.

In order to train the convolutional network, numerous datasets ranging from vehicles and bicycles in means of transportation to elephants and cats in animals will be collected and created from Flickr website as well as numerous famous datasets.

In this project, a model based on fully connected layers will be constructed to learn representations and make decisions based on local input dataset. In addition, an Keras library in-built model named VGG-16 will be implemented in the project to improve the performance of the model designed.

After the construction of prototype model, several photographs taken by members of the project will serve as input data to evaluate the accuracy and performance of the model. An output photograph which recognizes, and marks objects shown in the input photograph with given labels will be automatically generated. More improvements and optimizations of the model are expected in the future.

4. **Milestones:**
   4.1 Part of the datasets used will be generated by modifying the *demo08_2_flickr_images* provided by Prof. Yao Wang. In order to obtain a better result, a well-designed dataset may be used as well, for example, PASCAL VOC or MS COCO.
   4.2 All of the images will be normalized and regularized so that the network could work properly.
   4.3 More than 20 Photographs used for training will be taken as the input data for test.
   4.4 Prototype model based on FCN and VGG-16 will be constructed by training the datasets generated.
   4.5 Photographs will be tested, and output data will be collected and analyzed to further improve the model.

4.6 Optimizations such as autoencoder and development from static photographs to video clips are expected to be conducted.

5. References:

[1] J. Long, E. Shelhamer and T. Darrell. *Fully Convolutional Networks for Semantic Segmentation.* Retrieved April 18, 2018, from https://people.eecs.berkeley.edu/~jonlong/long_shelhamer_fcn.pdf

[2] E. Shelhamer, J. Long and T. Darrell. *A Fuller Understanding of Fully Convolutional Networks.* Retrieved April 18, 2018, from http://www.micc.unifi.it/bagdanov/pdfs/FCN-presentation.pdf

[3] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). *Imagenet: A large-scale hierarchical image database.* In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 248-255). IEEE.

[4] Krizhevsky, Alex, I. Sutskever, and G. E. Hinton. *"Imagenet classification with deep convolutional neural networks."* Advances in neural information processing systems. 2012.

[5] K. Simonyan and A. Zisserman. *VERY DEEP CONVOLUTIONAL NETWORKSFOR LARGE-SCALE IMAGE RECOGNITION.* In arXiv technical report, 2014, from https://arxiv.org/pdf/1409.1556.pdf

[6] S. Ioffe, C. Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.* Retrieved April 18, 2018, from https://arxiv.org/pdf/1502.03167v3.pdf