**The list and the definitions of the base line algorithms to compare your results**

* StixelNet

(Levi, Garnett, Fetaya, & Herzlyia, 2015) proposed a convolutional neural network named StixelNet, which is a multi-layer CNN, the architecture is shown in Figure 1. It is a 5 layer network in which the first two layers are convolutional and the last three are fully connected. The ReLU non-linearity is applied to the output of all layers except the last one, and for the convolutional layers a Max-pooling layer follows the ReLU operation as well. Given a single RGB image vertical stripe Is of dimensions (w; h; 3), this network can find the pixel location y of the bottom point of the closest obstacle in the center column of Is.



Figure 1 StixelNet Architecture

The segmentation task is done in three step, which is shown in Figure 2. The first two, StixelNet (trained on the road segmentation task) followed by a Conditional Random Field (CRF). The final stage performs a graph-cut segmentation on the image to achieve higher accuracy by enforcing road boundaries to coincide with image contours.



Figure 2 Road Segmentation algorithm flowchart

* Convolutional Neural Network with Context Fusion

(Brust, Sickert, Simon, Rodner, & Denzler, 2015) proposed a Neural Network that combines a 3-layer CNN with context information in the last two layers, the architecture is shown in Figure 3. In total, the network is constrcuted with 2 convolutional layer, 1 max-polling layer and 3 fully connected layers. An important architectural choice is the incorporation of the absolute position of the patch as an input in one fully connected layer, which allows for learning a spatial prior of the road category.



Figure 3 Network Architecture

* Fully Convolutional Neural Network with Large Context

(Mendes, Frémont, & Wolf, 2016) applied a fully convolutional network for road detection problem, the network archtercture is shown in Figure 4. It consists of extracting patches around a pixel or region of the image and classifying those patches using a trained CNN. The output class it attributed to the pixel or region from which the patch was centered. The main advantage of this method is it classify the pixel or region using not only its information but also information about its surrounding, that is, contextual information.

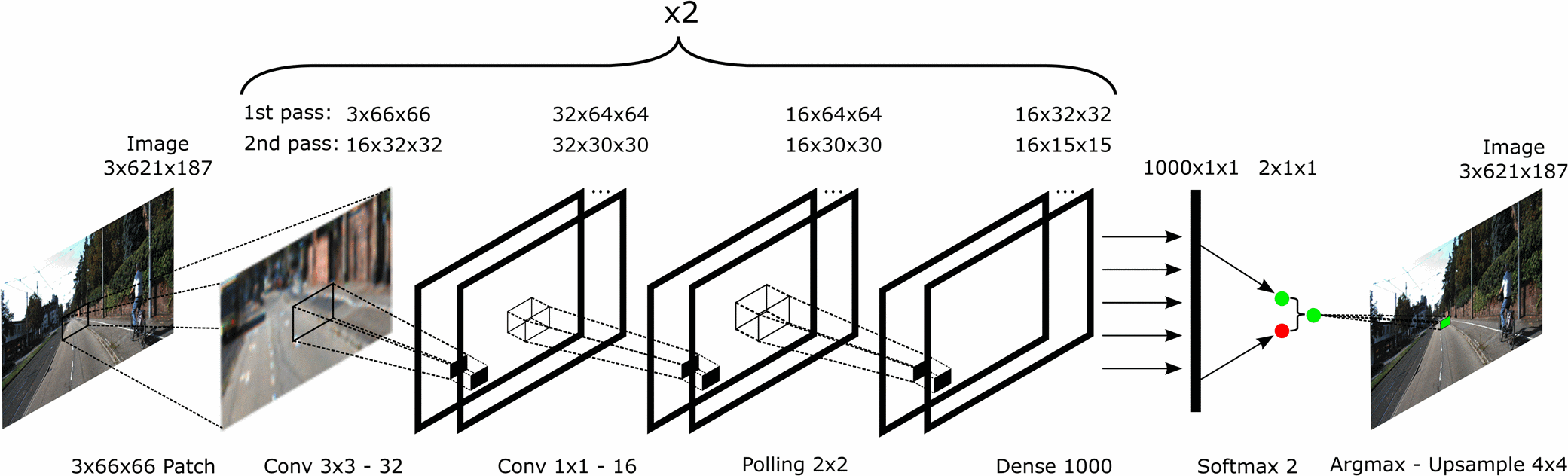


Figure 4 Network Architecture

* Up-Convolutional network

(Oliveira, Burgard, & Brox, 2016) proposed the Up-Convolutional network, the architecture is shown in Figure 5. Specifically, The contractive network layer parameters are initialized using the VGG classification network. The expansive network layers are consisted of successive convolution layers after a upsampling operation. One characteristic of the proposed expansive part is the inclusion of dropout after the first refinement layer to avoid overfitting. The output of the network are scores for each of the learned categories.

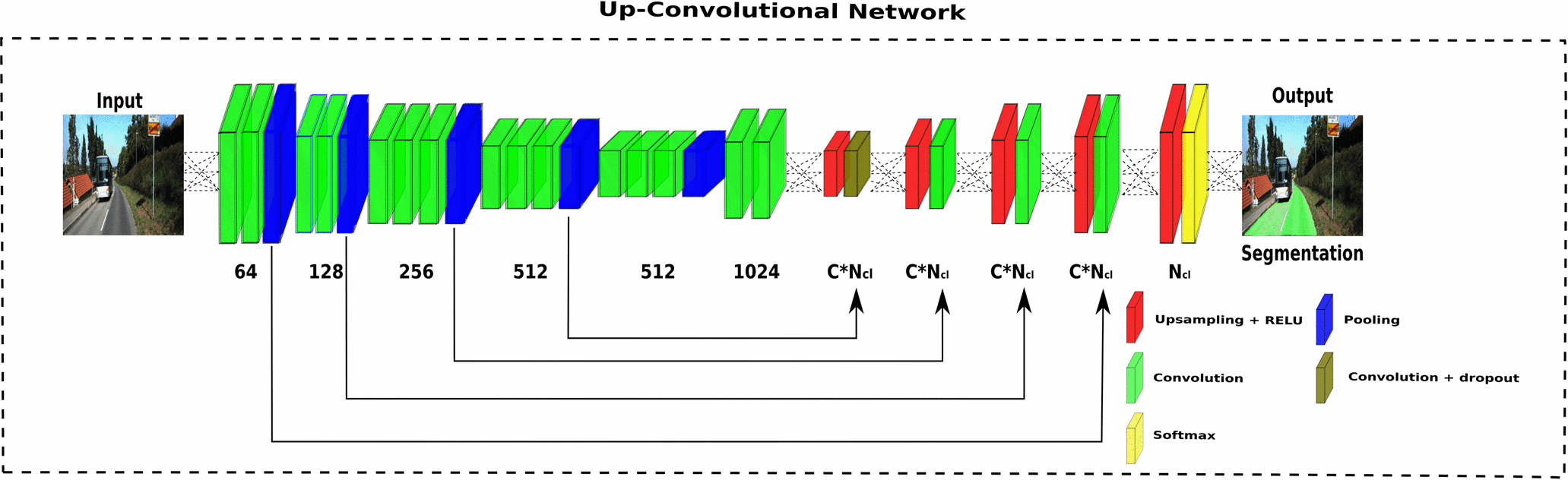


Figure 5 UP-Convolutional Network, where Ncl stands for number of classes and C for the scalar factor of filter augmentation.

At the same time, some studies (Alvarez, Gevers, LeCun, & Lopez, 2012; Fritsch, Kühnl, & Geiger, 2013; Kuhnl, Kummert, & Fritsch, 2012) can also be used to compare our methods although they don’t use KITTI dataset directly.

**REFERENCE**

Alvarez, J. M., Gevers, T., LeCun, Y., & Lopez, A. M. (2012). *Road Scene Segmentation from a Single Image*, Berlin, Heidelberg.

Brust, C.-A., Sickert, S., Simon, M., Rodner, E., & Denzler, J. J. a. p. a. (2015). Convolutional patch networks with spatial prior for road detection and urban scene understanding.

Fritsch, J., Kühnl, T., & Geiger, A. (2013, 6-9 Oct. 2013). *A new performance measure and evaluation benchmark for road detection algorithms.* Paper presented at the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013).

Kuhnl, T., Kummert, F., & Fritsch, J. (2012). Spatial Ray Features for Real-Time Ego-Lane Extraction. *2012 15th International Ieee Conference on Intelligent Transportation Systems (Itsc)*, 288-293.

Levi, D., Garnett, N., Fetaya, E., & Herzlyia, I. (2015). *StixelNet: A Deep Convolutional Network for Obstacle Detection and Road Segmentation.* Paper presented at the BMVC.

Mendes, C. C. T., Frémont, V., & Wolf, D. F. (2016). *Exploiting fully convolutional neural networks for fast road detection.* Paper presented at the Robotics and Automation (ICRA), 2016 IEEE International Conference on.

Oliveira, G. L., Burgard, W., & Brox, T. (2016). Efficient Deep Models for Monocular Road Segmentation. *2016 Ieee/Rsj International Conference on Intelligent Robots and Systems (Iros 2016)*, 4885-4891.