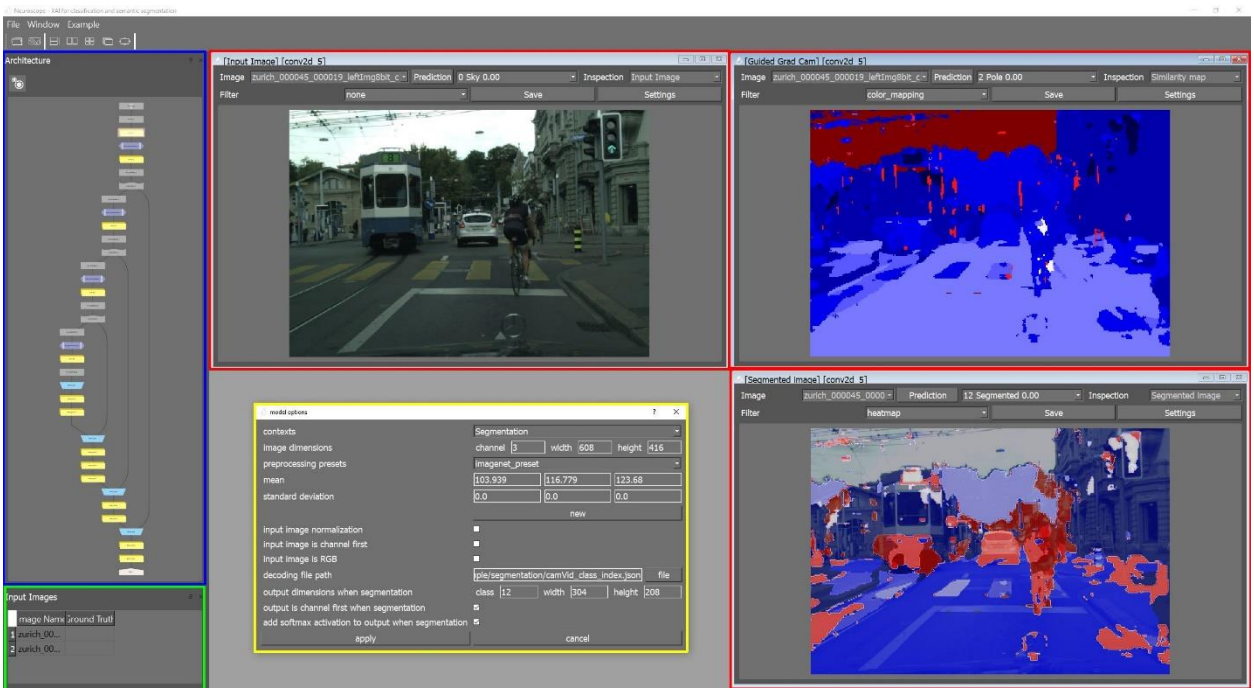
Quick Start Guide – Neuroscope version 1.0

German Research Centre for Artificial Intelligence (DFKI)

Last updated on 10.02.21, Contact: christian.schorr@dfki.de

Trust in AI predictions is a crucial point for a widespread acceptance of new technologies, especially in sensitive areas like autonomous driving. The need for tools explaining AI for deep learning of images is thus eminent. Our proposed toolbox Neuroscope addresses this demand by offering state-of-the-art visualization algorithms for image classification and newly adapted methods for semantic segmentation of CNNs. With its easy to use GUI, it provides visualization on all layers of a CNN. Due to its open model-view-controller architecture, networks generated and trained with Keras and PyTorch are processable, with an interface allowing extension to additional frameworks. This quick start guide explains how to explore Neuroscope's capabilities using two included neural networks for image classification and semantic segmentation. For the analysis of custom and user-generated models please refer to the upcoming Neuroscope manual.

Graphical User Interface

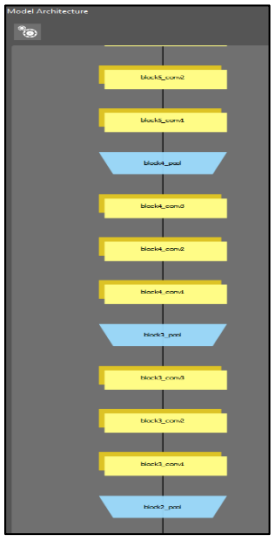


The graphical user interface of Neuroscope consists of four main parts: architecture window (blue), image list (green), model parameters (yellow) and inspections windows (red). To get started, click on *Example* in the tab and follow the following guide.

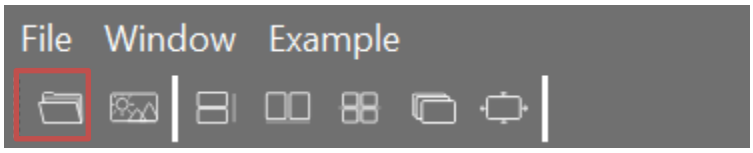
Examples

We provide two example models for Neuroscope. One is a pretrained VGG16 for classification and the other a pretrained U-Net for image segmentation. They can be accessed via the *Example* tab. The necessary model parameters and corresponding test images from autonomous driving are automatically selected and loaded.

1. Architecture Visualization

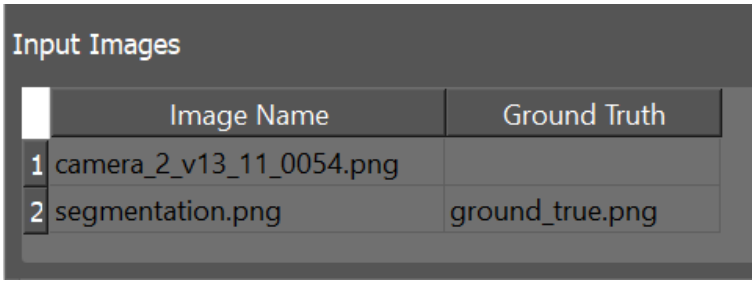


Clicking on the model import button in the toolbar opens a dialog to select the model you want to analyze. The architecture of the selected model is then visualized in the Model Architecture Window.



It displays all layers of the chosen neural network and distinguishes between different layer types using both color coding and different shapes. This view allows to choose layers of interest for subsequent inspection. Following the layers from input to output in conjunction with suitable inspection views illustrates the development of a model's internal behavior and thus helps to understand the predictions better.

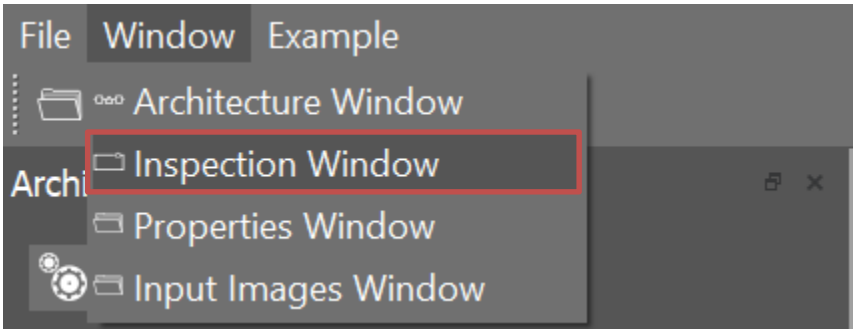
2. Image Loading



Clicking the image importing button in the toolbar allows to select one or multiple input images. The chosen images are shown in the Input Images window. For each of them, there may also be uploaded a corresponding ground truth image. This allows to compute confusion matrices and to evaluate the prediction according to a given metric, which can be chosen in the inspection window.

3. Model Inspection

The heart of Neuroscope is the inspection window. It is reached by clicking on *Window->Inspection Window* in the tab. Multiple inspection windows of the same image may also be opened side by side to compare different visualization methods.



A selected image from the image list can then be analyzed using a variety of visualization methods depending on the type of model – classification (C) or segmentation (S):

- Input image (C,S),
- Activation maps (C,S)
- Saliency map (C,S)
- Guided back propagation (C,S)
- Grad-CAM (C,S)
- Guided Grad-CAM (C,S)
- Grad-CAM++ (C)
- Score map (C,S)
- Segmented score map SSM (S)
- Guided segmented score map GSSM (S)
- Similarity map (S)
- Fusion score map FSM (S)
- Confusion Matrix (S, in conjunction with provided ground truth).

For a detailed explanation of the available methods please refer to the upcoming publication “Neuroscope - An explainable AI toolbox for semantic segmentation and classification of deep neural nets”.

When more than one inspection window is open, the buttons marked below can be used to arrange them top on top (a), side by side (b), in a matrix (c), in a stack (d) or centered (e).