A Deep Learning Approach to Camera Pose Estimation

Federico Izzo Francesco Bozzo

University of Trento

January 12, 2022

Introduction

With this work we are going to present:

- the exploration of multiple dataset generation techniques;
- the COLMAP reconstruction of Povo 1 second floor;
- the development of relative and absolute pose estimation models;
- the fine-tuning of absolute pose estimation models;
- the post-processing of the model outputs;
- the model deployment using a FastAPI web-server.

The MeNet Model for Relative Pose Estimation

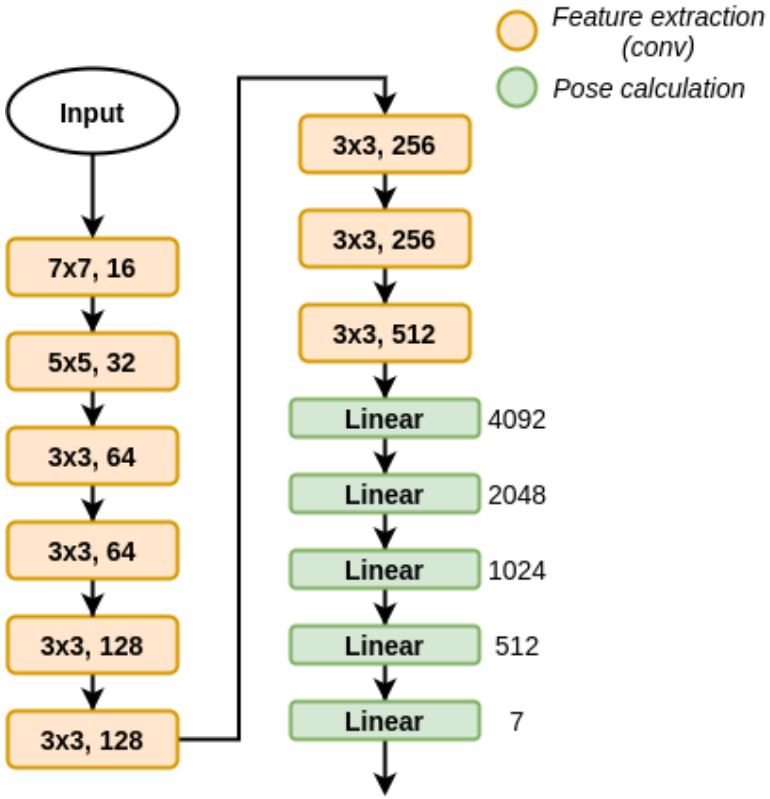

Figure: MeNet model architecture.

The **MeNet** model is targeted for **relative pose estimation**. The input of the network consists in a stack of two images: the goal is to estimate the relative pose of the second image with respect to the first one.

The PoseNet Model for Absolute Pose Estimation

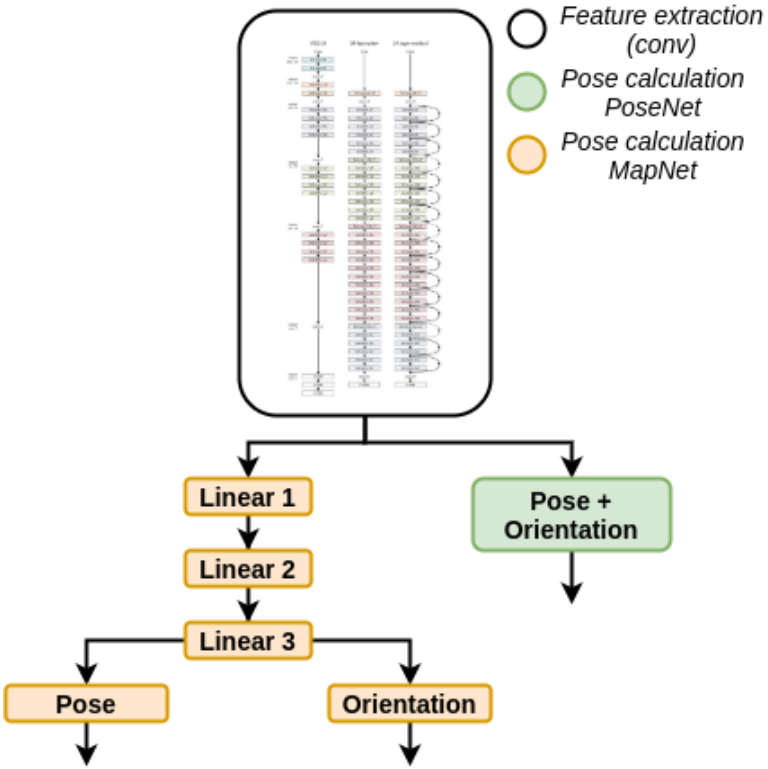

Figure: PoseNet model architecture.

The **PoseNet** model for absolute pose estimation is made up by two components:

- feature extraction through a sequence of convolutional layers (*backend*);
- pose regression on the extracted features using linear layers.

The MapNet Model for Absolute Pose Estimation

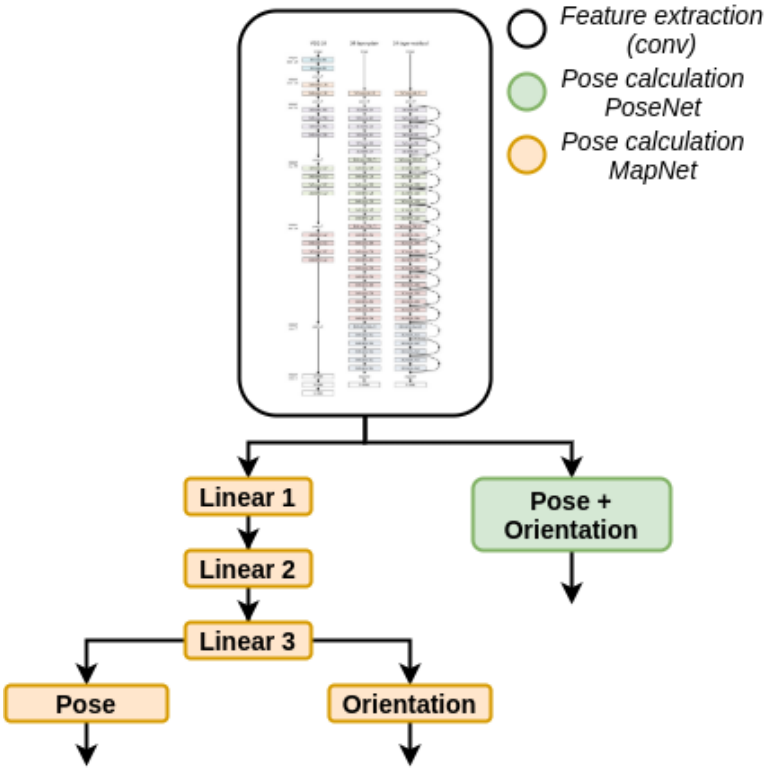

Figure: MapNet model architecture.

The **MapNet** model for absolute pose estimation represents an evolution of the PoseNet model with improvements:

- increase the number of final linear layers;
- penalize both absolute and relative errors in the loss.

Conclusions

To summarize, the final results presented in this work are:

- the exploration of multiple dataset generation techniques;
- the COLMAP reconstruction of Povo 1 second floor;
- the development of relative and absolute pose estimation models;
- the fine-tuning of absolute pose estimation models;
- the post-processing of the model outputs;
- the model deployment using a FastAPI web-server.