

– Supplementary Document –
**AutoSNAP: Automatically Learning Neural
Architectures for Instrument Pose Estimation**

David Kügler^{*,1,2}, Marc Uecker^{*,1}, Arjan Kuijper^{1,3}, and Anirban
Mukhopadhyay¹

¹ Department of Computer Science, TU Darmstadt, Darmstadt, Germany

² German Center for Neuro-degenerative Diseases (DZNE), Bonn, Germany

³ Fraunhofer IGD, Darmstadt, Germany

A. Qualitative Results

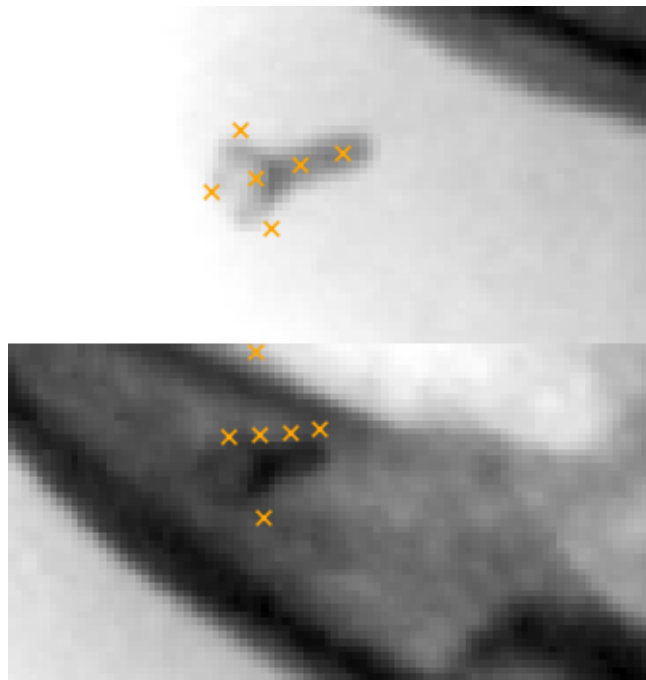


Fig.1: Qualitative Results on Dataset C: representative result (top), failure case (bottom). We also provide results of a full evaluation run on our website: <https://autosnap.david-kuegler.de>.

* equal contribution

B. Network Architectures

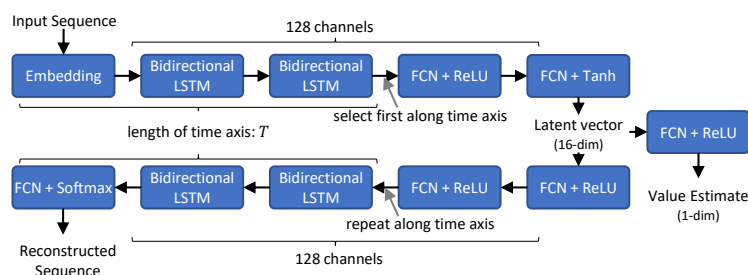


Fig. 2: Autoencoder Architecture. LSTMs are repeated T times (maximum sequence length). Input and Output can be flexible length (termination symbol).

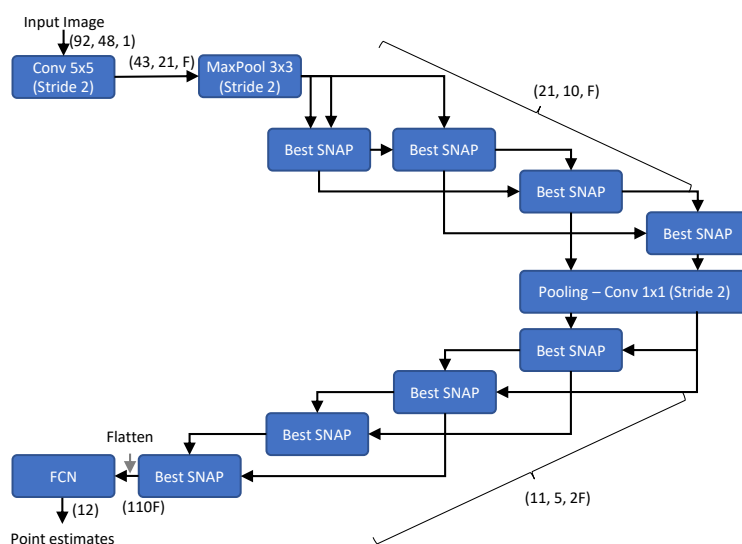


Fig. 3: Detailed SNAPNet Architecture. F corresponds to the number of filters (SNAPNet-A: 24; SNAPNet-B: 56), Fig. 4 (paper) shows the “Best SNAP” block

C. Extra Experiments

C.1 Performance of DARTS reimplementation

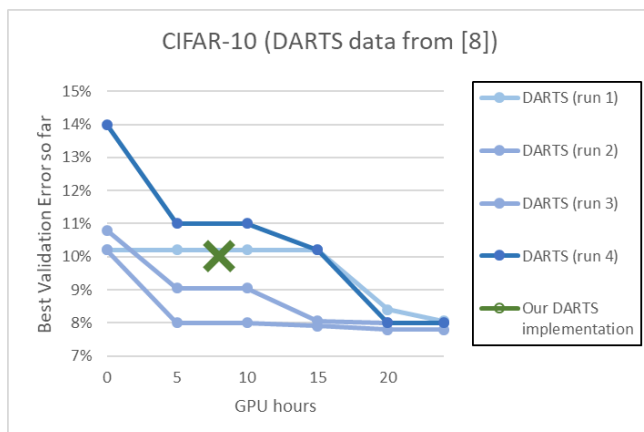


Fig. 4: Comparison of performance on CIFAR-10 of our DARTS implementation with the official DARTS implementation. DARTS (run 1-4) copied from [1].

C.2 Dependency of evaluation w.r.t. initial estimate

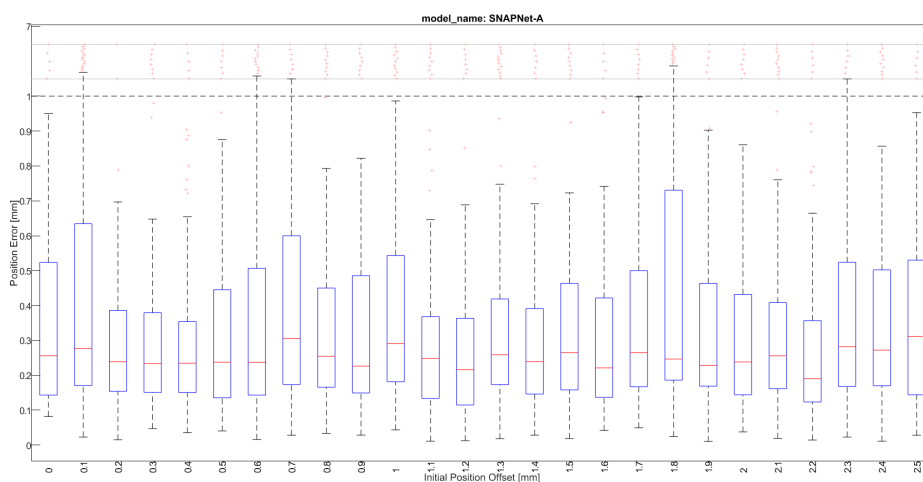


Fig. 5: Evaluation on Dataset C: position errors of SNAPNet-A with respect to initial position errors; no correlation with initial error identified after 3 iterations.

References

1. Liu, H., Simonyan, K., Yang, Y.: Darts: Differentiable architecture search. In: ICLR 2019 (2019), <https://arxiv.org/pdf/1806.09055>

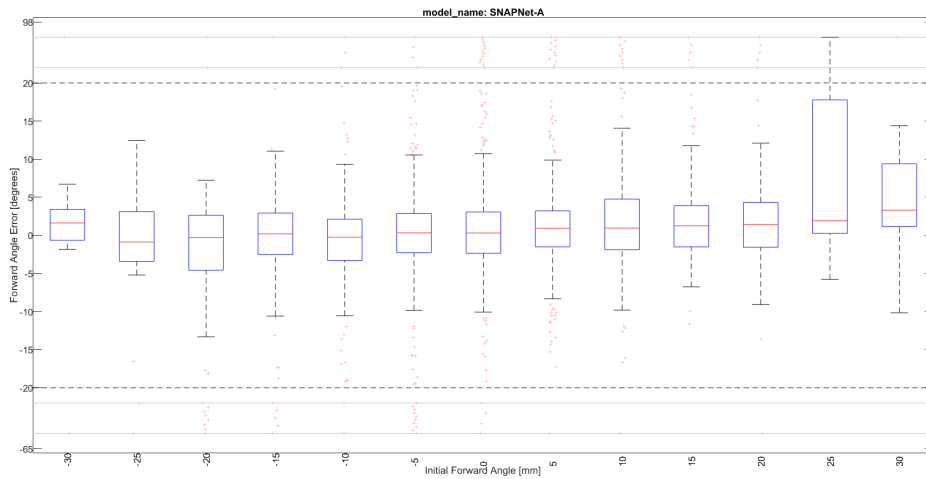


Fig. 6: Evaluation on Dataset C: angle errors of SNAPNet-A with respect to initial angle errors; no correlation with initial error identified after 3 iterations.