PhD projects R2 (2nd cohorte)

Convergence to Preferable Minima in Real-Time Nonlinear Optimization Problems
Multi Task Neural Networks in Autonomous Driving

Ivan Mykhailiuk, supervised by Prof. Dr. Christof Büskens - Lennart-Martin Evers, supervised by Prof. Dr. Christof Büskens

RTG 2224

Convergence to Preferable Minima

A typical real-time nonlinear optimization problem (**NLP**)

- is *highly nonlinear in constraints*, which means existence of **many local minima**;
- has *real-time constraints*, which requires to increase **computational efficiency**.

Global optimum is hard to find and often even undesirable.

State-of-the-art: nonlinear model predictive control (**MPC**)

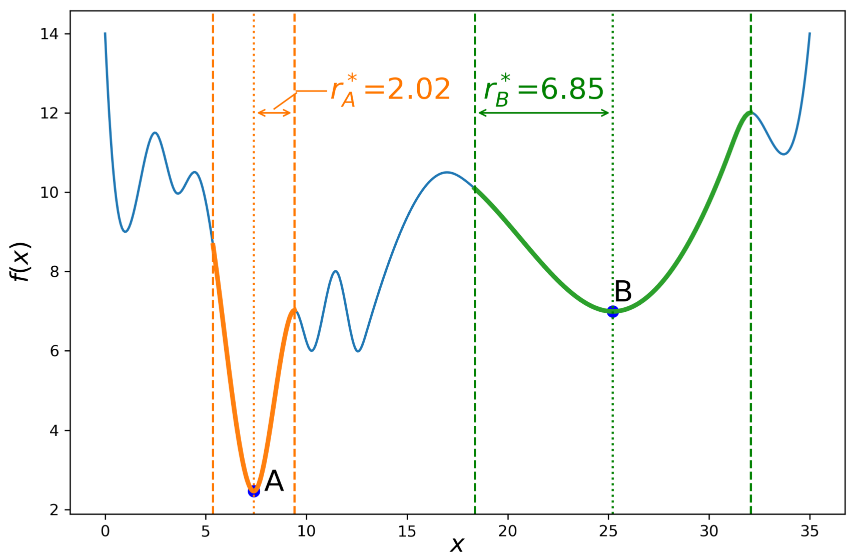
- sequence of **similar NLPs**;
- for each NLP – **few iterations** to *improve initial guess*;
- output of current NLP **initializes** next NLP.

Goal: develop a notion of **preferable** local **optimum**[1] which

- meets certain additional **criteria**, e.g. *attractability* and real-time *stability*;
- can be computed **offline** to initialize a *sequence of NLPs*.

Relevance: applications in optimal control of *autonomous driving* systems.

Example: Radius of Attraction



Define the "radius of attraction" of minimizer x^* w.r.t. the *gradient descent method* (GDM) as a solution of the problem

$$\begin{cases} \max\ r \quad \text{s.t.} \\ \int_{-r}^{r} \left(|f'(x+x^*)\,x| - f'(x+x^*)\,x\right) dx = 0; \quad r \geq 0. \end{cases} \tag{1}$$

In the **Figure** above, B has a larger radius of attraction w.r.t. GDM than the global minimizer A, hence, B is a good candidate to be defined as **preferable minimizer**.

Future work:

- generalize for functions of **several variables**;
- generalize for methods **other than GDM**;
- provide a **link to MPC** problems.

Multi Task Neural Networks

Deep convolutional neural networks can be used to solve different tasks in computer vision. These include the central problems in the perception of autonomous vehicles, such as

- **semantic segmentation**, i.e. assigning each pixel in an image a label;
- **object detection**, i.e. identifying the location, extent and kind of objects present in an image.

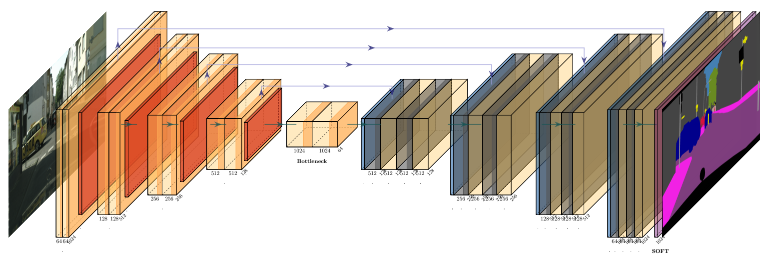
Modern neural network architectures such as Googles Deeplab[2] or FCOS[3] achieve state of the art results on both tasks individually. Research on tackling the problems *simultaneously* using at least partly *shared networks* is sparse however.

Goal: Unify task specific models:

- develop a *single encoder - multiple decoder* network architecture that achieves SOTA results in both segmentation and object detection;
- optimize network complexity to make models run in real time;
- develop and evaluate methods for training multi task networks, especially in the context of mixed datasets.

Relevance: sensor fusion and environment perception of autonomous driving systems.

Encoder-Decoder Architectures



A *neural network* $f\colon D \to O$ maps some *data*-space D to an *output*-space O. If the network can be decomposed into $f = d \circ e$, where $e\colon D \to L$ **encodes** the networks input to some *latent* space L, and $d\colon L \to O$ **decodes** variables from the latent space to our desired output space O, we say f is an encoder-decoder network.

- Most popular CNN architectures are encoder-decoder networks. In the visualization of U-NET[4] above, the structure is clearly visible.
- Using a single encoder and multiple, task specific decoders yields a network that performs multiple tasks simultaneously.

Future Work:

- investigate the influence of more decoder modules, such as lane detection or traffic sign recognition;
- evaluate the network on a real car in urban environments around Bremen.

[1] M. I., Schäfer, K., Flaßkamp, K., & Büskens, C. Preferable Minima in Nonlinear Optimization: Definition and Algorithmic Approaches. Proceedings of GAMM-2020 (in preparation).
[2] Chen, Liang-Chieh, et al. "Encoder-decoder with atrous separable convolution for semantic image segmentation." Proceedings of the European conference on computer vision (ECCV). 2018.
[3] Tian, Zhi, et al. "Fcos: Fully convolutional one-stage object detection." Proceedings of the IEEE International Conference on Computer Vision. 2019.
[4] Ronneberger, Olaf, et al. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.