Team Members: Jiaqi Liu, David Shaw

We built a model based on the architecture of a team that participated in the German Traffic Sign Recognition competition, but we also made some modification to the original design.

During preprocessing, all the training samples were down-sampled or up-sampled to 48x48 pixels. Global normalization was computed to center each input image around its mean value as well as enhancing edges. Local contrast normalization with Gaussian kernels from the original architecture was discarded.

Training images were split into 90% training and 10% validation sets.

The CNN architecture has the structure (Spatial Transformer Networks) - (Convolutional block) - (Spatial Transformer Networks) - (Convolutional block) - (Spatial Transformer Networks) - (Convolutional block) with RMSprop configuration alpha = 0.99, epsilon = 1e-8, weight decay = 0, and learning rate = 1e-5. The spatial transformation networks consist of localization network, grid generator, and sampler. They aim to perform a geometric transformation on an input map so that CNNs can be spatially invariant to the input data in a computationally efficient manner.

The layers are as follows:

Layer	Type	# Maps & neurons	Kernel
0	Input	3 m. of 48x48 n.	
1	Convolutional	200 m. of 46x46 n.	7x7
2	ReLU	200 m. of 46x46 n.	
3	Max-Pooling	200 m. of 23x23 n.	2x2
[redacted]			
5	Convolutional	250 m. of 24x24 n.	4x4
6	ReLU	250 m. of 24x24 n.	
7	Max-Pooling	250 m. of 12x12 n.	2x2
[redacted]			
9	Convolutional	350 m. of 13x13 n.	4x4
10	ReLU	350 m. of 13x13 n.	
11	Max-Pooling	350 m. of 6x6 n.	2x2
[redacted]			
13	Fully connected	400 neurons	1x1
14	ReLU	400 neurons	
15	Fully connected	43 neurons	1x1
16	Soft-max	43 neurons	

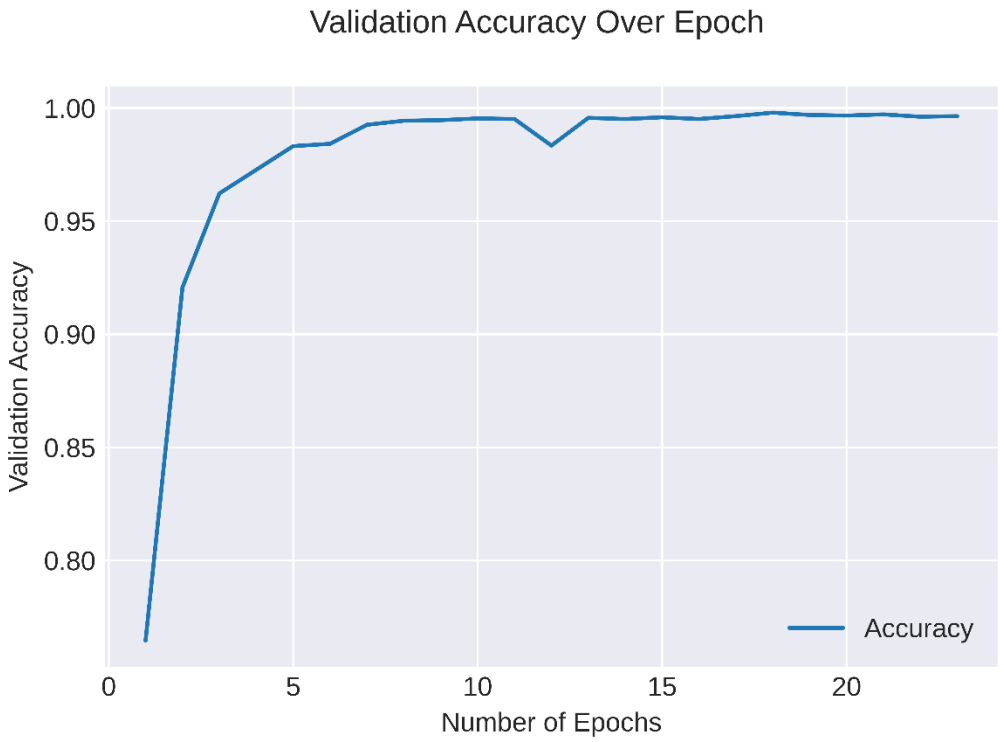
Table 1: Main CNN architecture without spatial transformer modules.

Convolutional layers' stride is set to 1, zero padding set to 2, and max-pooling layers' stride set to 2 with their zero padding to 0.

The implementation of Spatial Transformer Networks (STN) is based on pytorch's tutorial:

https://pytorch.org/tutorials/intermediate/spatial_transformer_tutorial.html

Result:



Citation: Arcos-García, Álvaro, et al. "Deep Neural Network for Traffic Sign Recognition Systems: An Analysis of Spatial Transformers and Stochastic Optimisation Methods." *Neural Networks*, vol. 99, 31 Jan. 2018, pp. 158–165., doi:10.1016/j.neunet.2018.01.005.