EvoTegra GmbH
Kreuzwingert 11 • D – 55296 Gau-Bischofsheim

EVOTEGRA
CONCEPTS & INTEGRATION

Case study: AI in 20 days

Abstract: AI technology is ready for production use. However many questions regarding the specific use and the implementation are still open. With our case study, we show the result of a reference AI implementation, which can be applied to the majority of use cases. The total effort for the case study was limited to 20 days.

As a full stack solution provider for customer specific Deep Learning solutions, Evotegra provides the full range from consulting and data services to process integration. With the ability to provide low cost solutions at scale we target small and mid tier businesses that do not intend to build up their own AI competence.

What is AI?

The "weak artificial intelligence" used today in practice is a self-optimizing mathematical approximation to an unknown complex function. The optimization is based on observations, which are made available to artificial intelligence in the form of data during the training phase. During the subsequent application phase, a weak AI will no longer learn.
Since today's artificial intelligence does not develop awareness, AI in principle is completely objective.

Advantages and disadvantages of AI

The advantage of AI for object detection is the mix of unparalleled accuracy and speed even on most diverse objects in combination with the flexibility to learn basically anything a human can see.

The advantages of AI:

- High accuracy
- High speed
- Robust detection
- High adaptability (learning ability)

Disadvantages:

- Requires a lot of data
- Requires time to train (1-2 days)

The following object properties are suitable for recognition with AI:

- High variance
- Undefined spatial alignment

Anschrift:
EvoTegra GmbH
Kreuzwingert 11
D – 55296
Gau-Bischofsheim

Kontakt:
Tel.: +49 171 / 402 4242
E-Mail: info@evotegra.de
Homepage: www.evotegra.de

Gerichtsstand: Mainz
HR-Nr.: B 43519
USt.-ID Nr.: DE 217667197
Steuer-Nr.: 44/654/70280

Geschäftsführer:
Tobias Manthey

Bankverbindung:
Deutsche Bank 24 AG
IBAN DE46 5507 0024 0024 0820 00
BIC DEUTDEDBMAI

- Undefined lighting conditions
- Limited number of objects to recognize (<500)
- Little change in the set of objects to be detected

e.g. Road signs, Hazardous materials, persons, vehicles, bottles …

A contraindication is given if the set of objects frequently changes or a large number (> 1000) of different objects have to be detected.

The use of artificial intelligence

While AI can reach super-human accuracy levels it is sensitive to transfer. Any difference in the setup that is not covered in the training scenarios can affect the recognition rate. On the other hand adding more scenarios will increase the likelihood of false detections. While this might be acceptable for consumer-grade products, this characteristic is typically not suitable for industrial solutions. Therefore we do not believe much in classic products that integrate AI (one AI solution for all use-cases) compared to individual platform-based solutions that are highly optimized for a specific scenario.

Today's AI technology has matured over the last years to a "production ready" level. Networks can be reliably trained and executed, C++ integration and network optimization enables process integration. A wide range of available hardware from embedded systems to high-end data center solutions allows deployment of AI solutions in nearly any scenario.

To deploy AI solutions close to machines (edge-processing) is especially important to industry and image processing solutions. While the high bandwidth of data from a camera cannot be efficiently processed in the cloud also company policies might not allow cloud based processing of sensible production data.

Data

An AI solution is never better than the quality of the data used. For a common detection problem that includes 100 different classes we recommend in total between 100,000 and 1,000,000 samples depending on the scenario. With highly automated processes the necessary data can now be obtained in large quantities, short time in high quality at low cost. Other important factors during the project are the flexibility to adapt the scope without major additional effort as well as to ensure data quality.

Case study setup

The core advantage of artificial intelligence is its ability to learn and hence its adaptability to a wide range perception task. In order to ensure the transferability to a maximum number of potential use-cases in industry or transport, a particularly demanding task was chosen.

The requirements for the object recognition are:

- Small objects or long distance to the camera
- Uncontrolled lighting conditions
- Complex background
- Dynamic scenes
- Up to 50 m/s relative speed object to camera
- Mobile setup
- Maximum 30 watts power consumption
- No cloud
- Data from public space

The vast majority of potential use-cases have lower requirements.

We chose therefore to detect German traffic signs. As the case study assumes that there is no initial data, the data must be collected and prepared for AI training as part of the case study. In order to obtain a high quality dataset with high diversity, the number of images per trip and physical sign is limited to 5.

The total effort to conduct the case study was limited to 20 man-days.

Definition of the classes for detection

The scope for the recognition of traffic signs was the overall catalog of German Traffic signs including complex directional road signs, digital traffic signs as well as various types of traffic lights. As a result 297 different signs and markings were included in the catalog.

E.g.:

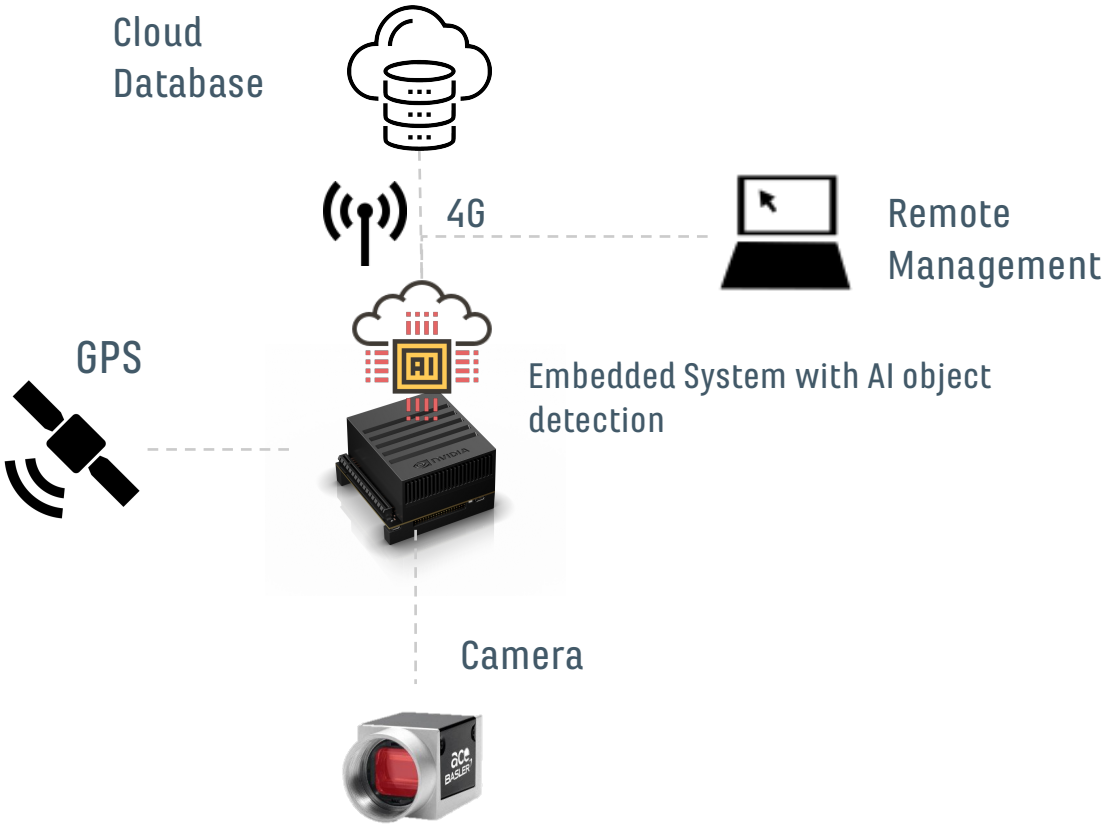
Effort: 1 day

Data collection

The goal of the case study was to fully capture the data as part of the case study. For data acquisition the fully automatic data collection system from Evotegra GmbH was used. Due to the high relative speed of the vehicle to the object with up to 180 km/h, the data has to be captured with at least 30 frames per second. In addition to images the system can optionally capture spatial data using a stereo camera system. By installing the system in a regular vehicle, neither safety nor usability must be affected. All changes must be non-permanent, fully reversible and it is required that removing the system from a vehicle must just take a few minutes.

The system consists of the following components:

- Embedded system NVIDIA Xavier AGX
- USB industrial camera with 1920x1200 and 40hz resolution
- High-speed LTE router with 100 GB monthly volume
- GPS
- Inverter or car charger
- Cloud web service
- Cloud SQL database



A pre-trained traffic sign recognition network was used on the embedded system. For some applications, either pre-trained networks or corresponding public datasets are available. If there are no pre-trained networks or data available, there will be an additional effort of 5-10 days.

By using the pre-trained network, the vast majority of camera images without traffic signs are already filtered in the vehicle. In order to execute neural networks in real-time in a vehicle, network optimization must be used. Images with potential signs are compressed, encrypted and a checksum is generated. As a result, the amount of data is reduced to a level where data can be transferred directly from the vehicle using the 4G/LTE network.

Targeted data acquisition, strong encryption, integrity checks and identification of the communication partner ensure responsible handling of the data and ensure compliance with the requirements of the GDPR. A future improvement is the removal of all privacy related features already in the vehicle, which will be integrated into the next version of the data collection system.

Since the vehicle was just commonly used and the system is fully automated, no effort was charged for the data collection in the study. The duration of the data collection therefore depends on the number of vehicles and their usage. As part of the study, the data was recorded in a vehicle over a period of 2 months with a distance of about 5000km in different regions at different times of day including night.

64068 pictures with a resolution of 1920x1200 pixels in a volume of 69 GB were recorded.

Effort: 1 day

Data processing

Training an AI system requires images as well as the information of the location of the object within the image. This labeling process is usually carried out manually or by service providers and hence is usually very time consuming or expensive. Therefore in order to minimize the cost and effort highly automated data labeling was used. Compared to a manual processing, automation can increase the productivity of a person by up to 4000%. Next to quantity, automated data process improves the quality and helps to identify incomplete or erroneous assumptions regarding the data.

For example it showed during the data analysis that more types of beacons exist than initially expected. These were therefore added to the catalog of classes supported by the system.

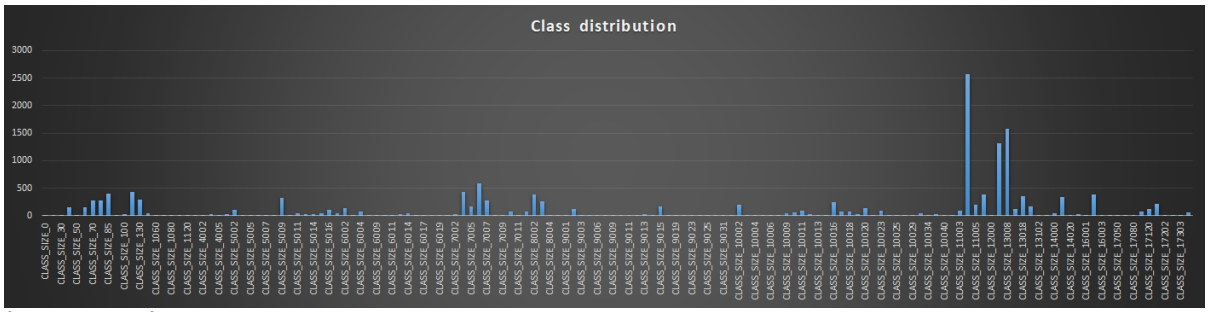
Beacon (605)

Beacon (605) Variant 1

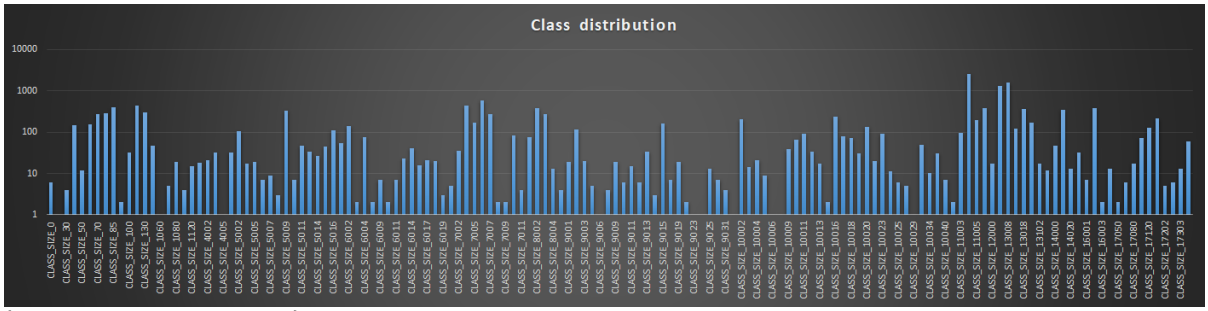
Beacon (?) Variant 2

In general it is safe to assume that the initial assumptions on the classes to be detected by the system have to be supplemented and adapted several times in the course of an AI project.

A total of 202940 data labels in 172 categories were obtained from the 64068 images.



(Class distribution)



(Class distribution logarithmic scale)

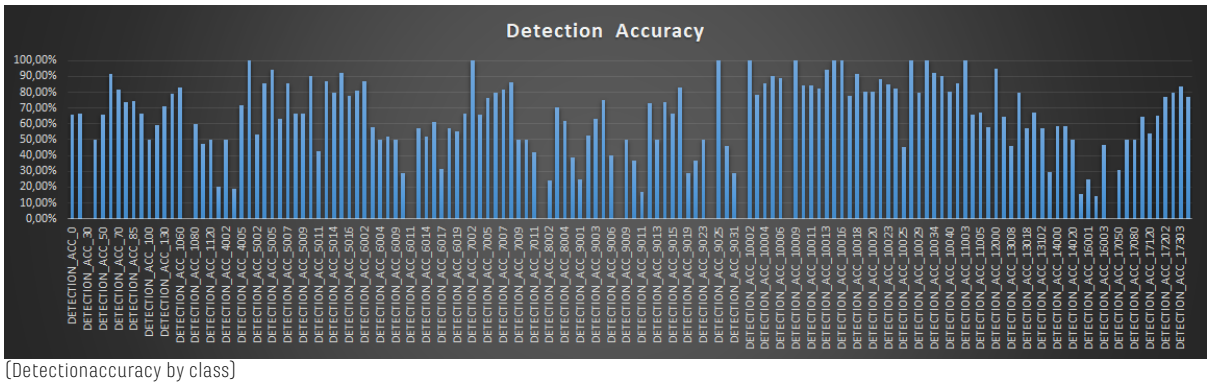
In a demanding scenario such as traffic sign recognition, we recommend using about 10,000 images per category for training. In simpler use cases 1000 images per category can be sufficient. As apparent in the class distribution, no class could reach the recommended mark in the study.

Effort: 15 days

Training

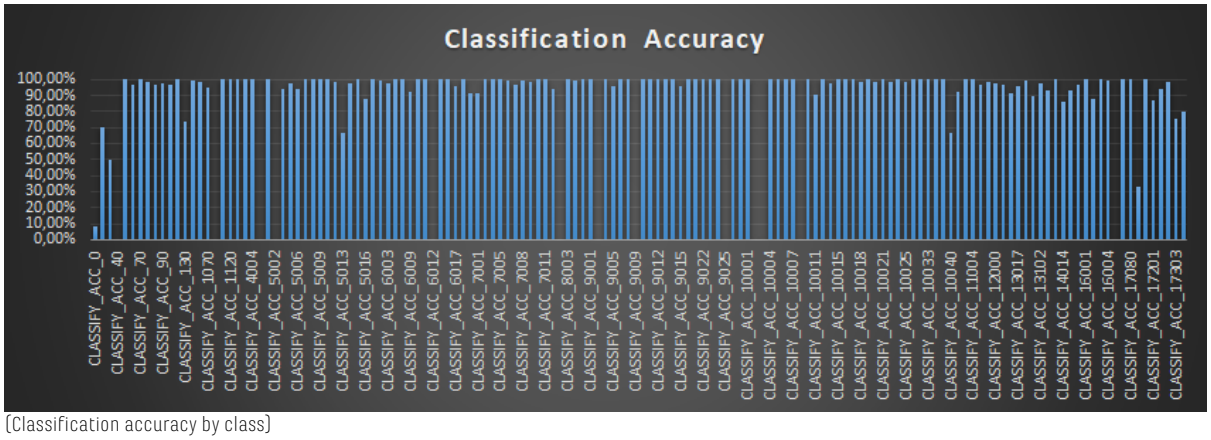
Due to the uneven data distribution only 172 classes could be used for the training and some classes had to be merged.

After the division into training and test data, around 20 different networks for object recognition and classification with different parameters were trained. The duration of a training run is usually 1-2 days. Training and testing are highly automated and can be parallelized.



(Detectionaccuracy by class)

The average detection accuracy reaches an F score of 0.7. The F score measure combines precision and recall by a mean. Precision is the ratio of detected signs to all detections while recall defines the ratio of detected signs to all signs.



(Classification accuracy by class)

The average classification accuracy is 92%.

Especially the recognition accuracy depends as usual by the data distribution.

Effort: 2 man-days

Optimization and deployment

Networks must usually be optimized in order to execute in real time on embedded systems. By merging layers and simplifying operations the execution can be accelerated by up to 500%. For detection of small or distant objects such as traffic signs, a high resolution for the network execution is required.

The Evotegra AI platform is used to execute the networks. The modular platform allows different types of networks to be executed from different deep learning frameworks. The platform already integrates different methods for object tracking and classification and has different cloud, data and device interfaces. Fully implemented in C++, the platform can be integrated into virtually any environment.

40-50 frames per second can be analyzed on the embedded system at a resolution of 960x600 pixels.

Effort: 1 day

Results

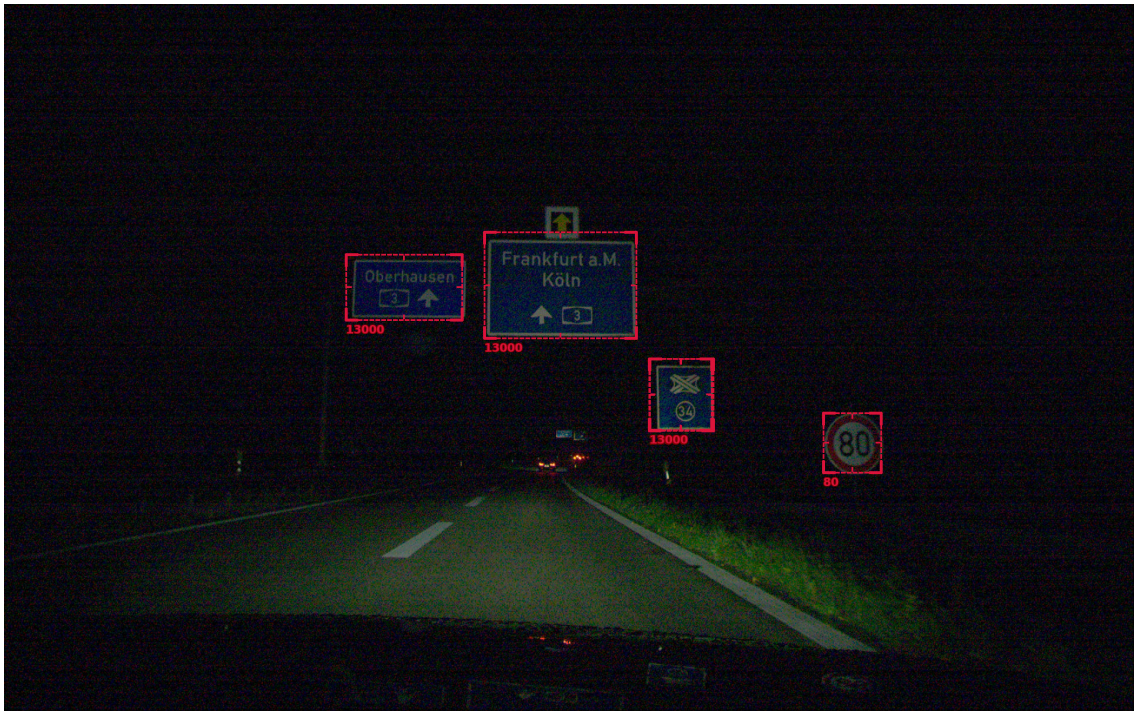
Despite the limited amount of data, the recognition of most classes is already satisfactory to good. The detection works reliably even at night and is therefore robust against rough light conditions and noisy images. The execution speed in combination with the resolution enables reliable in-car detection even at high speeds.

Since the results are essentially improved by adding more data, the success can be estimated early in AI projects. In combination with data change management and quality control, the project risk can be minimized. Additional data can be captured by fully automated systems in a targeted manner and at minimum cost.

With the presented processes, data records even in the order of millions can be collected and managed in a relatively short time. With highly automated processes and flexible platforms, customer-specific AI solutions can be developed at a cost that for the first time enable highly optimized solutions on individual machine level.

AI cannot replace human intuition in the medium term. In fact human intuition blends perfectly with the objectivity and performance of AI. Therefore the full potential can be leveraged by the smart combination of these complementary types of intelligence. As a result AI offers a demanding job to employees while quality can be better objectified and productivity increased by up to 4000%.

AI is therefore suitable today for a large variety of visual inspection tasks, e.g. for completeness checks or quality control. For the first time cameras can be used as a universal sensor, e.g. in the visual monitoring of instruments, switches and processes. In order to guarantee the personal rights of the employees, the processing of the images can be done directly on the machine. As a result, images do not have to be permanently stored or transferred.


Oberhausen
3
13000
Frankfurt a.M.
Köln
3
13000
34
13000
80
80


100
100
100
100
13000
Seitenstreifen
befahren
12000
13000
Seitenstreifen
befahren
12000


120
120
120
120