EvoTegra GmbH
Kreuzwingert 11 • D – 55296 Gau-Bischofsheim

EVOTEGRA
CONCEPTS & INTEGRATION

AI Industrialization

Abstract: While all other prerequisites are fulfilled, today's methods for data acquisition not just prevent a wide adoption, they also inhibit leveraging the full potential of AI. This article discusses a novel approach to the data problem that we consider a cornerstone for mass deployment.

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.

The current status of AI

The advantages 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.

While AI can reach super-human accuracy levels it is sensible 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 so much in classic products that integrate AI (one AI solution for all use-cases) rather than 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.

The data problem

But while all other prerequisites are fulfilled, data still remains to be the key obstacle for a wide adoption of AI. 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.

To solve the data problem there are currently two common approaches:
1.) The client is labeling the data using a tool
2.) A labeling service provider

Our estimation is that a single person can label up to 500 images per day. The issues affecting the quality are the monotone workload as well as the subjectivity of the individuals. As a result this process requires an additional quality control to assure the results can be used in an AI system.

Labeling Service Provider require a ramp up phase where the data requirements are aligned and the interfaces of the provider have to be implemented. Depending on the provider this process can take days to month. But also during the labeling stage the provider might need to be supported in solving questions and corner cases. Quality, throughput and costs vary by the provider. As a result an independent additional quality control process is recommended.

The vast majority of today's potential projects will stop at this point as the implied costs and risks of data acquisition do not enable a valid business case.

Independent of the data strategy the inherent problem of both approaches is as simple as *change*. The initial class model is typically based on human experience also known as assumptions. So the probability that a class model requires adaptation during a project is high. At this point AI projects based on either one approach typically overrun the budget.

But human labeled data has an even more adverse effect.

The bitter lesson

Human perception is influenced among others by expectation, distraction, context, sleep level, mood or genetic disposition. Human perception therefore is highly subjective.

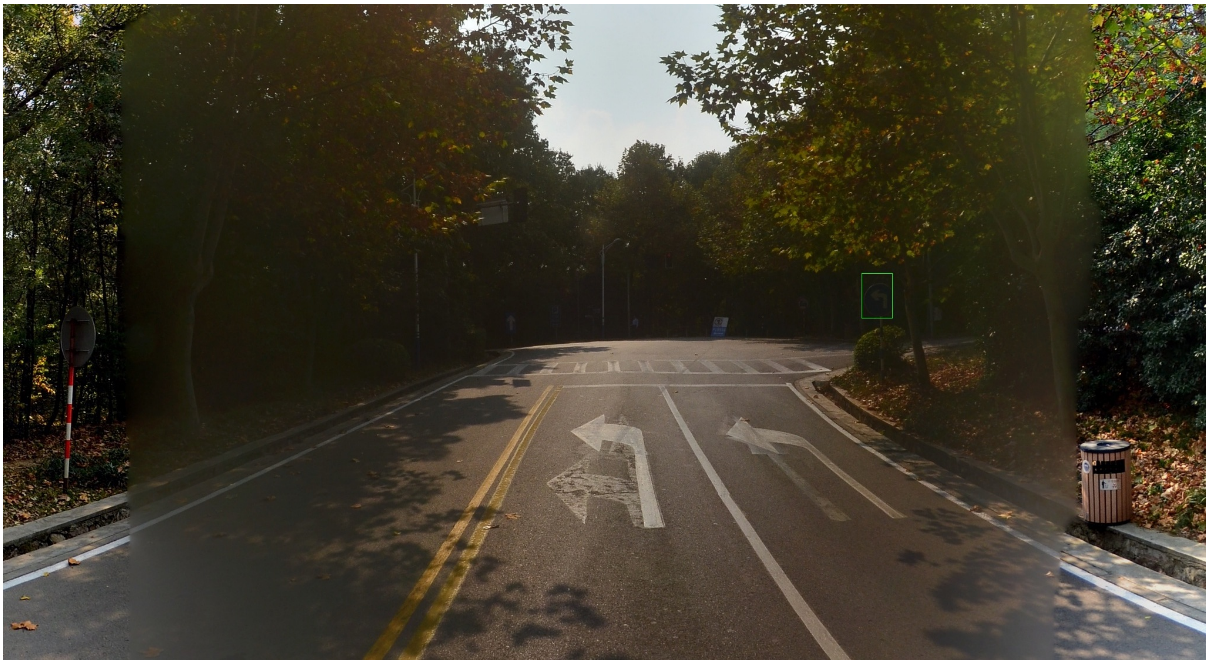
The "weak artificial intelligence" used today in practice is a self-optimizing mathematical approximation to an unknown complex function. Since today's artificial intelligence does not develop an awareness of what it learns, AI is completely objective.

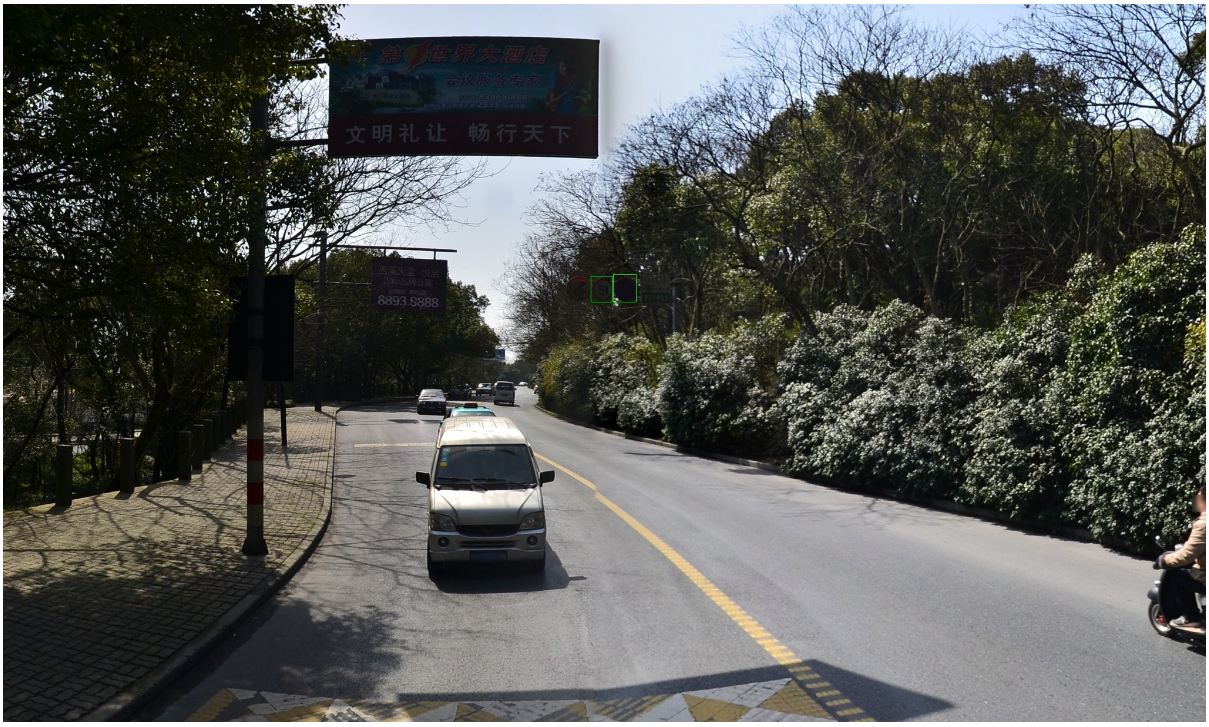
In March 2019 Rich Sutton, research scientist at Google DeepMind and a professor of computing science wrote the article called "The bitter lesson". In his article he describes how AI researchers are tempted to build human knowledge into their systems as a short-term solution to improve results, while this causes typically stagnation in the further development. What he calls the biggest lesson in 70 years of AI research is that computation is by far the most effective solution to improve AI.

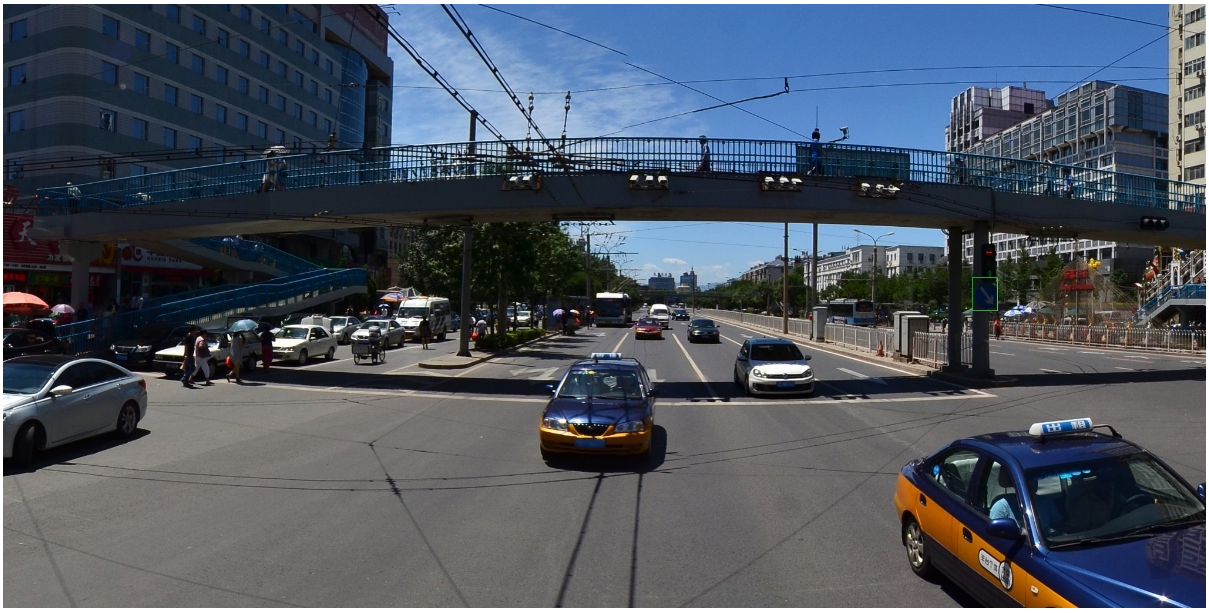
So to leverage the full potential of AI, human knowledge must not be tightly integrated into Artificial Intelligence. But this is in fact what all current solutions to the data problem do. The human knowledge comes in form of the definition of classes and human labeled data. To fuel an objective system with subjective data means in fact to provide objectively inconsistent data. This has negative effects on generalization and hence prevents today's Deep Learning solutions to exploit their full potential.

Subjective versus objective perception

To validate our assumption we applied our method on the Tsinghua-Tencent 100K dataset for traffic sign detection. We analyzed the negative images, i.e. the images declared to contain no sign. Our analysis found many traffic signs which are either difficult to see or in unusual locations, which leads to the assumption these images were labeled by humans. In the samples we picked, we show only front-facing signs and marked them with a bounding box.







The lesson learned

As a solution provider we depend on low cost, high quality data in large volumes and short time. Based on our experience with human labeled data, we had implemented these conclusions in our data processes over the last couple years. In our projects data is not labeled by humans, but in a highly automated process supervised by a person.

The effort to label data in general correlates to the total number of instances in the data and the number of classes. But before data can be automatically labeled, the data pipeline needs to be setup. Depending on the project phase and requirements highly automated labeling is using different techniques based on (un)supervised machine learning and even classical algorithms. The effort to run a data pipeline is initially higher and decreases over time.

In a reference scenario of:

- 1,000,000 total images
- 50% of images containing a single instance
- 100 different classes

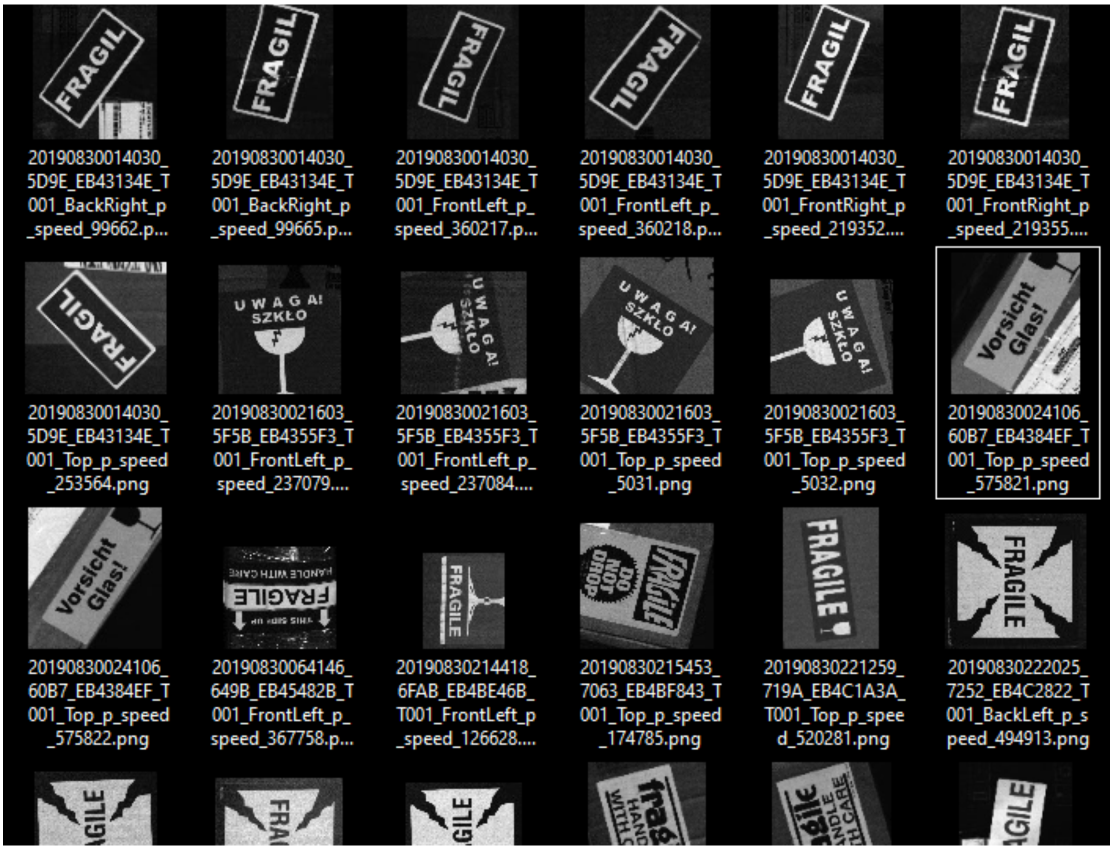
a single person can supervise the labeling of 10,000 to 20,000 images per day *on average*. Compared to human labeling this is a productivity increase of up to 4000%. The bulk size for a single run is typically 100,000 images.

But as important highly automated labeling supports changes of the class definitions, which significantly reduces project risk. And all these advantages come at a quality level of the data, which is not just significantly better, but is also constantly monitored throughout the project.

As a result we now fuel an objective system with objective data managed by human intuition. Our observations show that generalization now reaches levels that enable AI to find the errors made in the human assumptions by finding missing classes and categories. This enables AI to not just detect images based on their content, but also on their meaning.

Human intuition and ethics complement perfectly with the productivity and objectivity of AI. Since AI quickly uncovers false or incomplete human assumptions, the learning is mutual. And since human intuition cannot be replaced by artificial intelligence in the medium term, we generally recommend the use of AI as an assistance system.

The Rosetta Stone of AI



This is an example of our automated labeling process in the detection of "Fragile" sticker in what we call the Rosette Stone of AI. Due to the lack of standardization nearly infinite variations of this sticker exists across various languages and symbols. So despite the fact many of these labels lack presence in the training data all above images show sticker meaning "Fragile" and therefore are considered true positive. The probability that a group of individuals would consistently label the images in a similar fashion is basically non-existent.

Due to the fact that this is not a singular but a general observation during the process the class definition will most evidently change during a project multiple times. Only highly automated labeling can support this method of evolutionary learning.

Conclusion

To understand AI it is necessary to see the world through the eyes of an AI. Objective systems fueled with objective data unlock the hidden potential and boost AI augmented workers with up to 4000% more productivity. Paving the way to AI mass adoption evolutionary learning makes data available in short time, large quantities at low costs for most specific use-cases.

That's not a vision. That is our daily business.