Deep Systems

Artificial intelligence for real problems: deepsystems.ai.

Nov 22 · 7 min read

⚔️ Big Challenge in Deep Learning: Training Data

Hello World. Today we are going to cover one of the most central problem in Deep Learning—training data problem.

We at DeepSystems apply deep learning to various real-world tasks. Here are some of them: self-driving cars, receipt recognition, road defects detection, interactive movie recommendations and so on.

And most of our time we spend not on building neural networks, but on dealing with training data. Deep Learning needs lots of data and sometimes it takes an hour to annotate just one image! We always thought: is there a way to speed up our work? Well, we found it.

We are proud to announce new awesome feature of Supervisely: AI powered annotation tool to segment objects on images much faster.
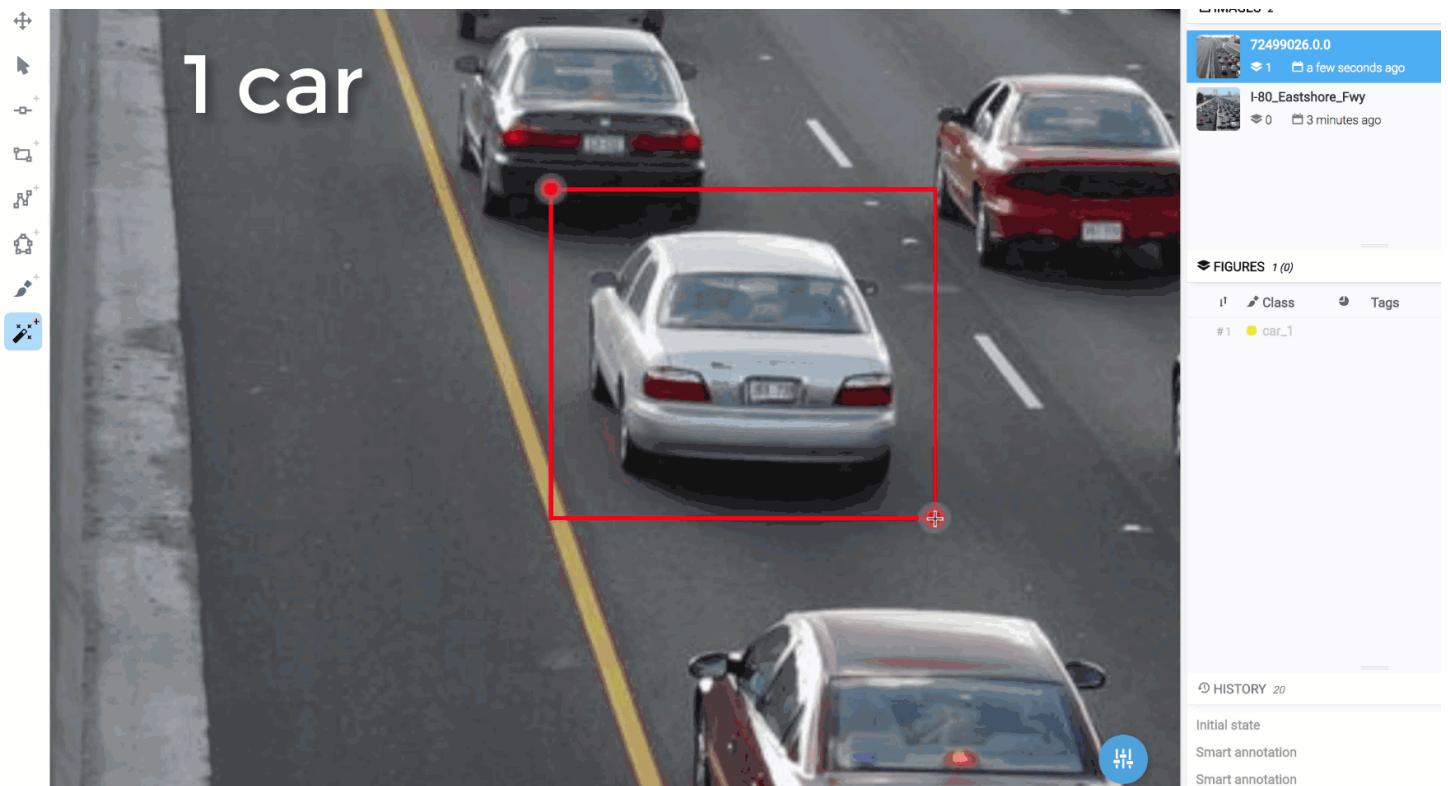
Here we will focus on computer vision, but similar thoughts can be applied to a large diversity of data: text, audio, sensor data, medical data and so on.

Big picture: more data—smarter AI

Let me start with our small modification of very famous slide from Andrew Ng.
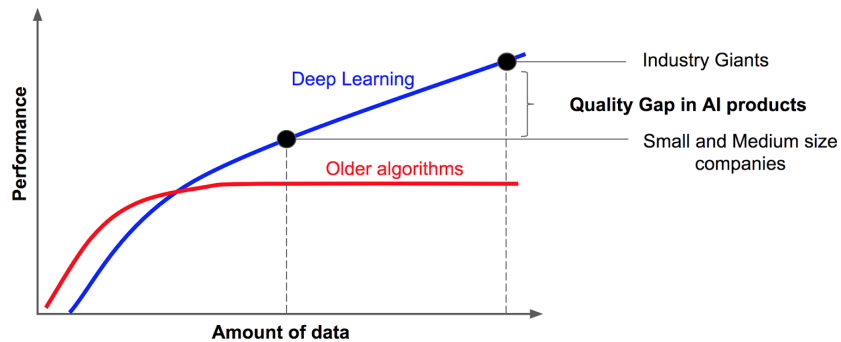
It is not a secret that deep learning outperforms other machine learning algorithms. Let's get some insights from this graph (some of them may seem obvious but …).

Conclusion 0: AI products need data.

**Conclusion 1**: the more data we have—the smarter AI will be.

Conclusion 2: industry giants have much more data than others.

Conclusion 3: quality gap in AI products is defined by amount of data.

So, network architecture can strongly influence the performance of an AI system, yet the amount of training data has biggest impact on performance. Companies which focused on data gathering provide better AI products and are hugely successful.

Common mistake: AI is all about building neural nets.

Let me show you a chart.

What people think AI is about   The reality
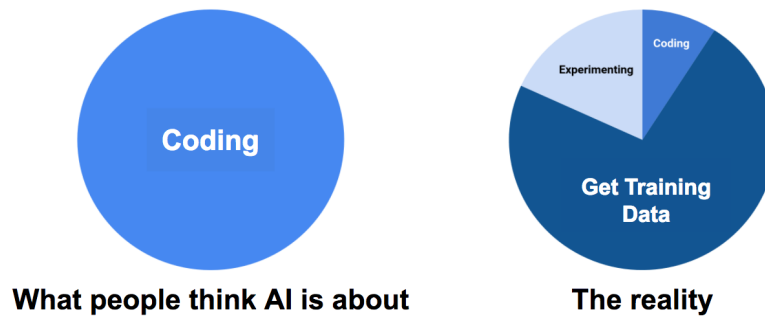
When people think about AI they think about the algorithms, but they should also think about the data. Algorithms are free: Google and other giants tend to share their state-of-the-art research with the world, but what they don't—they don't share data.

Lots of people have jumped on AI hype train and created awesome tools to build and train neural networks, but very few focus on training data. When companies try to apply AI they have all the tools to train neural networks but no tools to develop training data.

Andrew Ng Says Enough Papers, Let's Build AI Now!

Nice idea, we agree with him. There are a lot of papers and open-source implementations of state of the art neural network architectures that can cover almost all real-word problems. Imagine, you got new 1 billion dollars idea. The first question will not be: what kind of neural network will i use? Most probably it will be: Where can i get the data to build MVP?
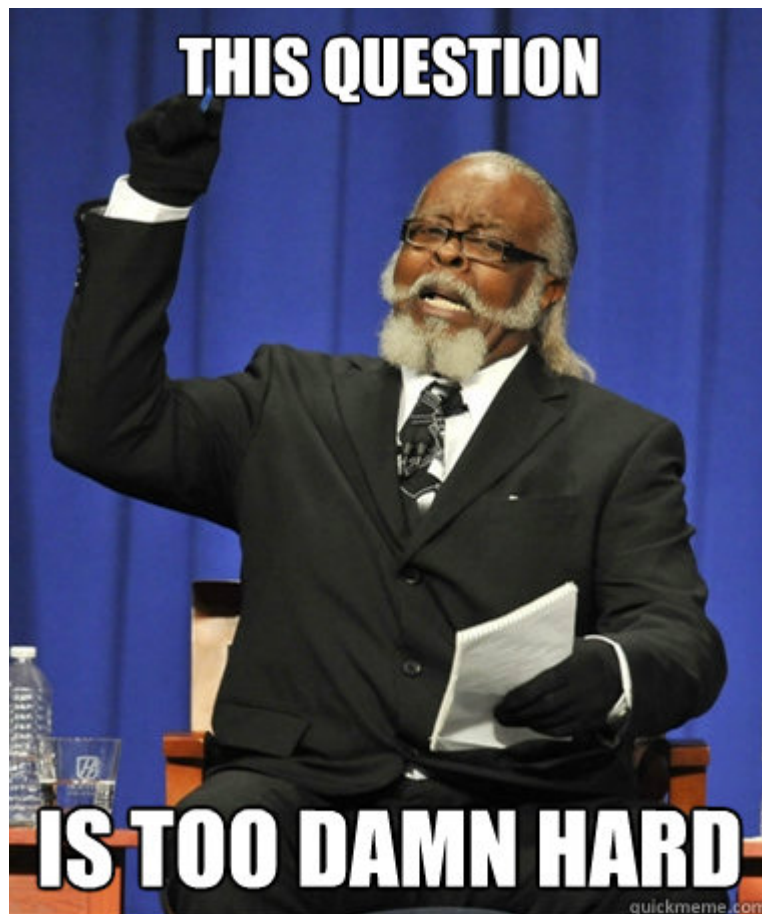
Sources of Training data. Let's find a silver bullet.

Let's consider some available options.

- **Open-sourced datasets.** The value of deep NN is in the data it is used to train it. Most available data in computer vision research are tailored to the problem of a specific research group and it is often that new researchers need to collect additional data to solve their own problems. That's why it is not a solution in most cases.

- **Artificial data**. For some tasks like OCR or text detection it is ok. But many examples (face detection, medical images, …) nicely illustrate that it is very hard to generate or even impossible. The common practice is to use artificial data in combination with real annotated images when it is possible.

- **Web**. It is hard to automatically collect high quality training data. Most likely that human should correct and filter it.

- **Order image annotation from someone**. There are some companies that provide such services. And, yes, we are no exception. But strong drawback is that you can not iterate fast. Usually even data scientist is not sure about how to annotate. General pipeline is to make iterative research: annotate small portion of images -> build NN -> check the results. Each new experiment will influence the next one.

- **Annotate images by hands**. Only you understand your task. Domain expertise is crucial. Medical image annotation is a good example: only doctor knows where tumor is. We understand that this process is time consuming, but if you want custom AI—there is no other ways.

So, as we can see, there is no silver bullet. And most common scenario is to create own task specific training data, generate artificial data and merge them with public datasets if it is possible.

The key is that for your custom task you have to create own unique dataset, fo sho.

Let's utilize Deep learning to build Deep Learning.

What? The idea is the following. Deep learning approaches are data hungry and their performance is strongly correlated with the amount of available training data.

Let me show you how annotation process is hard. Here is the raw numbers of how much time annotation process can take. Let's consider Cityscapes dataset (useful for self-driving cars). Fine pixel-level annotation of a single image from cityscapes required more than 1.5 hours on average. They annotated 5000 images. With a simple math we can calculate, that they spent about 5000 * 1.5 = 7500 hours. Consider, 1 hour = $10 (close to minimum wage in USA). Thus, only annotation for such dataset costs around $75K (not including additional costs).

It is also surprising that only single self-driving company has 1000 in-house workers that do image annotation. And that's the tip of the iceberg.

Imagine how much time and money companies and individuals spend for image annotation. It is unbelievable. This is the huge obstacle to

progress in AI. We have to do annotation for our own task, but it can last forever 😰.

Can neural networks help us to make it faster? Think about that. We are not the first who tried to answer the question.
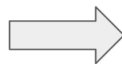
Field of semi-automated annotation of object instances has a long history. There are many classical methods to speed up annotation like Superpixels, Watershed, GrabCut. In last few years researchers try to utilize deep learning for this task (link1, link2, link3). Classical methods work bad and have many hyperparameters to search for every image, it is hard to generalize them and correct their results. Latest Deep Learning based approaches are works much more better, but in most cases they are not opensourced, it is hard to implement, reproduce results and integrate to any available annotation platform.

But we are the first who make AI powered annotation tools available for everyone. We designed our own NN architecture that has concepts similar to three links above. It also has one big advantage: our NN is class-agnostic. This mean that it can segment pedestrians, cars, potholes on the road surface, tumors on medical images, indoor scenes, food ingredients, objects from satellite and other cool stuff.
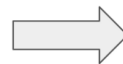
So, How does it work?

How to use AI powered segmentation tool

You just have to crop the interested object and neural network will segment it. It is very important that you can interact with it. You can click inside and outside object to correct mistakes.

Unlike semantic segmentation that partitions an image into multiple regions of pre-defined semantic categories, our interactive image segmentation aims at extracting the object of interest based on user inputs.

The primary goal of interactive segmentation is to improve overall user experience by extracting the object accurately with minimal user effort. Thus we significantly speed up the annotation process. There are some examples below. See for yourself.

Self-driving cars

This embedded content is from a site that does not comply with the Do Not Track (DNT) setting now enabled on your browser.

Please note, if you click through and view it anyway, you may be tracked by the website hosting the embed.

Learn More about Medium's DNT policy

SHOW EMBED

As you can see from our 45 seconds test AI powered annotation tool allows to annotate each image in a few clicks. While it is needed 57 clicks with polygonal tool to annotate single car.

Segmentation of food ingredients

This example demonstrates that with polygonal tool it is really hard and slow precisely segment objects of irregular shapes and with not straight edges. We would like to emphasise that clicks inside and outside object are much "cheaper" with respect to clicks on the edges.

This was our first attempt. Of course, there are cases when smart annotation works bad. But we are going to constantly improve the quality and make simple way for domain adaptation: to customize tool for specific task inside Supervisely without coding.

Conclusion

Data is a key in deep learning. It was time consuming and very expensive. But we and deep learning community actively try to solve training data problem. First steps are already done, results are promising, let's keep going.

We deliberately missed the topic about unsupervised learning. It is a very promising field of study, but today supervised learning dominates in real world applications.

In the next posts we will try to cover all possible use cases to help you understand that it is suitable for your tasks.

Sounds exciting? Go check us at https://supervise.ly. We are in public beta and it is free :-) If you will have any technical or general questions, feel free to ask in our Slack.

If you found this article useful, then let's help others too. More people will see it if you give it some 👏.