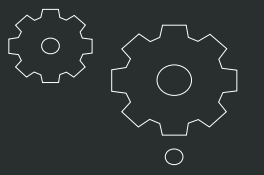
smart eye
since 1999

• smart eye technology

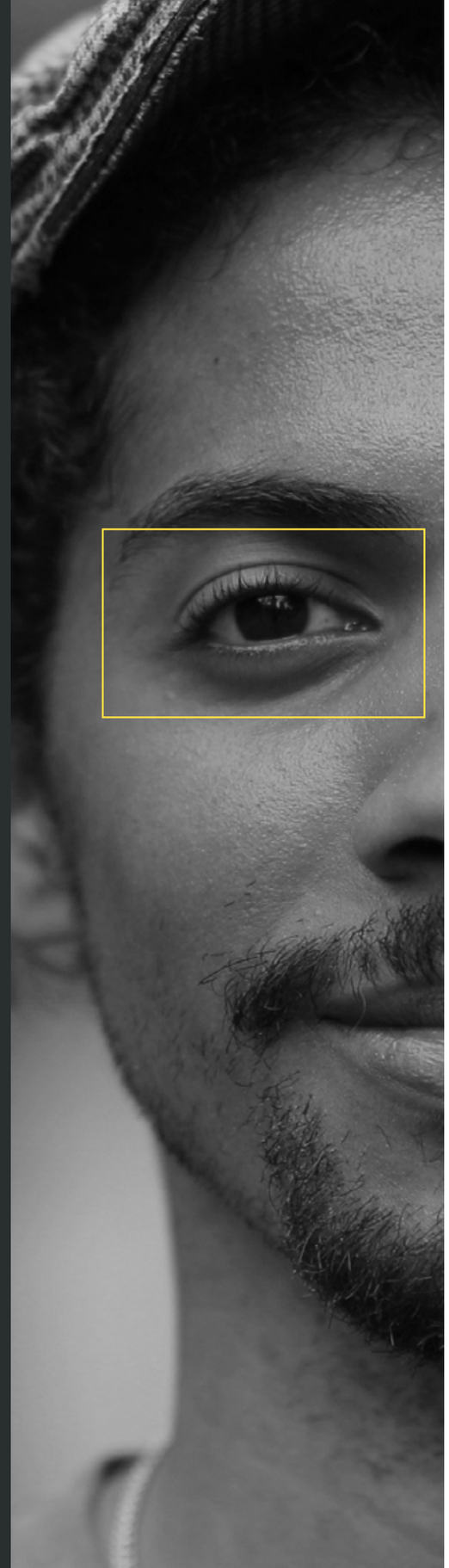
From eye movement to neural network

From eye movement to neural network

- An in-depth look to our technology



smart eye



Introduction

At Smart Eye, every eye movement we record and algorithm we write are intended to increase our understanding of the most complicated system in the world: the human being. To be able to take on this challenge, our software is fueled by AI-based technology that has taken years to develop and perfect.
In this deep dive into our technology, we will start by going back to the basics in order to explain the complex process of understanding human behavior.

The basic process

The first step is to provide our systems with data to analyze. To capture images of eye movements and facial features, we use entirely camera-based monitoring. Through the use of infrared light, the cameras are able to detect a reflection in each eye's cornea.

Once the images have been extracted, they are run through software that tracks the person's gaze and eye movements. Meanwhile, AI-supported algorithms interpret the details of the face captured – like the person's state of mind or mood. For every time the system identifies a face, it gets to know its unique characteristics a little better, enabling higher efficiency and more accurate results.

A guide to Smart Eye's

technology

Computer vision

To identify visual patterns and details in the captured images, Smart Eye uses computer vision. Computer vision is an umbrella term for all methods that use computer algorithms in order to decide what an image or video contains.

Methods based on computer vision can be divided into traditional computer vision methods and deep learning methods. The main difference between the two subcategories is that traditional computer vision methods largely require manual coding, while deep learning methods include systems that can learn independently when fed large amounts of example data. Additionally, traditional computer vision methods analyze the different components of an image individually, while deep learning methods analyze the entire image as a unit. We'll elaborate on which method Smart Eye uses and why, later on.

Machine learning

The foundation of many newer computer vision methods is AI-based machine learning. Machine learning refers to a system's ability to be trained by being exposed to datasets.

To make machine learning possible, the system needs to have access to very large amounts of data to use as learning examples. The machine then processes the information extracted from the examples through algorithms.
Peter Rundberg, technical expert at Smart Eye, explains the concept of machine learning:

- As humans, we can tell if a person seems tired, but it can be difficult to describe what we detect as the visual signs of drowsiness. That's what an AI-algorithm can be trained to do by being fed large amounts of recorded data of drivers looking both tired and alert.

Deep learning

In turn, deep learning methods can be described as a sub-category to machine learning. Deep learning methods are highly modern types of machine learning methods, based on AI-technology.

– Deep learning is a very popular area of research at the moment. The progress being made proves that these techniques can do things that were previously thought to be impossible, says Peter Rundberg.

At the core of deep learning technology are many-layered neural networks.

What is a neural network?

There are several different types of neural networks, but they all consist of small building blocks. Each of these building blocks has a very simple function and is practically useless on its own.

– These building blocks are what can be likened to a neuron in a brain, even though it obviously isn't the exact same thing as a biological neuron, says Peter Rundberg.

When large amounts of these basic building blocks, neurons, are connected to each other they form a neural network. Methods that use neural networks are based on how the neurons connect to each other, and the level of strength of these connections can be trained.

How to train a neural network

While the architecture of a neural network is designed by the programmer, the training of the network adjusts the level of strength of the connections between the different neurons. This is done by supplying the network with example images, training data, while simultaneously indicating to the network what the image portrays.

```
x = int(input())
y = int(input())
outputP = x*weights[0] + y*weights[1] + bias*weights[2]
if outputP > 0 : #activation function
  outputP = 1
else :
  outputP = 0
print(x, "or", y, "is : ", outputP)


outputP = 1/(1+numpy.exp(-outputP)) #sigmoid function


 input_data = X_train[i:min(X_train.shape[0],i+batch_size)]
labels = Y_train[i:min(X_train.shape[0],i+batch_size)]

NeuralNetwork(nn.Module):

    def __init__(self):
        super().__init__()
        self.fc1 = nn.Linear(32, 64)
        self.relu1 = nn.ReLU()
        self.fc2 = nn.Linear(64, 256)
        self.relu2 = nn.ReLU()
        self.out = nn.Linear(256, 1)
        self.final = nn.Sigmoid()

    def forward(self, x):
        op = self.fc1(x)
        op = self.relu1(op)
        op = self.fc2(op)
        op = self.relu2(op)
        op = self.out(op)
        y = self.final(op)
        return y
```

How to train a neural network

For example, let us pretend we are trying to train a neural network to distinguish between different types of animals. We might start with feeding the network an image of a dog, while letting the network know that the image depicts a dog. This will urge the network to identify the features in the image that are typical for a dog.

We then need to continue feeding the network with hundreds of thousands of images that are different from each other in every way but one: they all contain dogs. Over time, the network will learn to detect the details that characterize dogs.

The next step would be to continue providing the network with massive amounts of images containing other animals. These images also need to have a *correct annotation*, meaning the network needs to be informed of what animal is represented in the image. Once sufficiently trained, the network can be fed an image without a correct annotation - with no information about what the image contains – and determine which animal the image depicts on its own.

– And that is exactly what we do at Smart Eye, says Peter Rundberg. *Only we might not primarily want to identify different types of animals in an image, but rather want the output to tell us in which direction a person is looking.*

When training a network into correctly analyzing a person's facial features, the process is roughly the same. But instead of images of animals, the system needs to be fed large amounts of images containing faces of different genders, ages and ethnicities.
These images also need to be correctly annotated – the system needs to be told that the image contains a human face and what expression it displays. And that is how Smart Eye uses deep learning methods to understand human behavior.

Combining traditional computer vision with modern deep learning

Real life situations can be unpredictable. We do not believe there is a "one size fits all"-solution for every step in the complex process of analyzing human behavior. Therefore, our approach is to apply both traditional computer vision methods and modern deep learning, depending on the situation.

– We want to use the best methods for the problem at hand, explains Peter Rundberg. That means not throwing away the older, more traditional methods and just going for the deep learning methods. That would be to disregard all the great research that has made the more traditional computer vision methods very efficient.

One good reason to combine the traditional computer vision methods with the modern deep learning methods is that the two approaches reach their conclusions through completely different processes. This allows for one method to be used in order to get a "second opinion" on the other's output. If both systems generate the same answer, we can be confident it is the correct result. But two differing outputs could indicate a potential problem with the image.
This could imply that the image needs to be further analyzed, or at least indicate an error, which according to Peter Rundberg is equally important.

– In the automotive industry especially, it's very important to be able to say "we don't know", instead of inaccurately saying a person is looking to the right, for example.
An incorrect output could be pretty dangerous when you're driving a car. In that case, it's better to say that we can't be sure about what exactly the driver is doing. Then the car will have to react in a different way.

The possibilities of deep learning

There is no doubt deep learning technology has taken the tech industry by storm. Deep learning is an unusually active field of research, both academically and industrially. The methods are developing rapidly, and any company that wants to stay on top of modern technology needs to be caught up with the new, modern AI-based methods.

To Peter Rundberg, it has become clear that both traditional computer vision methods and deep learning methods have their unique strengths.

– *The advantage of using methods based on deep learning is that they are often very good at extracting information from images that can be very hard to analyze using traditional computer vision.*

For example, deep learning methods are currently unprecedented in terms of detecting details in images of poor quality. Since traditional computer vision methods focuses on features like sharp edges and contrast in the images, it can be very difficult to analyze an image that is too dark or too blurry using those methods. A neural network, on the other hand, can be trained to see through the bad quality and distinguish the information behind it. This makes deep learning methods more robust and stable than traditional computer vision methods. Since deep learning systems analyze the entire image at the same time, they are less sensitive to incomplete images where some components are concealed.

Conclusion

For 20 years, Smart Eye has been working in close collaboration with the automotive industry –an industry that demands for our technology to keep an exceptional standard. They rely on us to keep their costumers safe, and we take our job very seriously.

To ensure safe transportation for drivers, passengers and pedestrians alike, our systems need to be almost 100% accurate in their output. They need to be able to offer precise results in all environments and light conditions, and they have to be able to see through sunglasses or other face coverings. This would not be possible without our advanced deep learning methods.

The software also needs to rank high in availability. Installing complex computer systems in vehicles can get extremely expensive and take up a lot of resources. This requires us to develop adaptable systems that are cost efficient and use processor power conservatively.

But in order to keep up with the fast-paced developments of the automotive industry, we are always looking for new solutions to solve both existing and future problems. And Peter Rundberg believes the future will bring technological innovations that could make way for a new generation of AI-technology.

– I think AI-researchers will make progress in the development of technologies that aren't completely based on deep learning. It's possible you could create other types of neural networks that can take in more information, more history and make better critical decisions in the future.

Smart Eye is committed to delivering the most advanced nonintrusive 3D head and eye tracking system in the world. We strive to establish a standard of reliability and availability which is unparalleled in the industry. We are equally committed to accommodating even the most complex applications and demanding field of view requirements from a remote perspective, while still maintaining superior accuracy.

Follow us:

Linkedin Twitter Facebook

Contact: www.smarteye.ai