Image understanding with deep conventional network

Liangjie Cao

9 May 2018

Today I learn image understanding with deep conventional network.(Table. refTable) The paper says since the early 2000s, ConvNet has been applied with great success. Images can be labelled at the pixel level which will have application in technology. Mobileye and NVDIA(Fig. 1) has used such ConvNet-based methods in their upcoming vision systems for cars so long.

ConvNet were largely forsaben by the mainstream computer vision, and machine-learning communication until the ImageNet competition in 2012. This has caused most major technology companies, including Google, facebook, Microsoft, IBM, Yahoo, Twiiter and Adobe. ConvNets are easily amenable to efficient hardware implementations in chips or field programmable gate arrays. [1] A number of companies such as NVIDIA, Mobileye, Intel, Qualcomm and Samsung are developing ConvNet chips to enable real time vision applications in smartphones, cameras, robots and self-driving cars. [2]

The paper also says a stunning demonstration combines ConvNets and recurrent net modules for the generation of image captions (Fig. 2). Techniques to generate more training examples by deforming the existing ones. This success has brought about a revolution in computer vision. The following days, I will continue learning what this model really works.
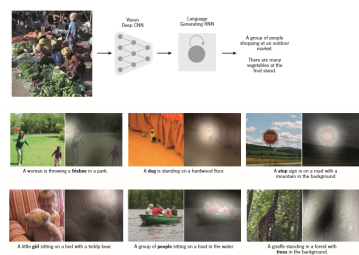
Figure 1: **NVIDA**

Figure 2: **From image to text**

1	convolutional layers
2	pooling layers
3	Relu layer

Table 1: **Layers used by ConvNets**

References

[1] Sackinger E. Bromley J. LeCun Y. & Jackel Boser, B. An analog neural network processor with programmable topology. *Solid State Circuits*, (26):2017–2025, 1991.

[2] Yoshua Bengio& Geoffrey Hinton Yann LeCun. Deep learning. *Nature*, 521(28):9, 2015.