Gradient-Based Learning Applied to Document Recognition

主要方程:Yp=F(Zp,W),Zp代表第p个输入量,W代表系统中可调参数的集合

损失函数:Ep=f(Dp,Yp)

Gradient-Based Learning:

Wk=Wk-1-εδE(W)/δW 对可调参数Wk偏导

Stochastic gradient algorithm(on-line update)

比传统gradient descent快很多,但会上下抖动

同一个feature map中的所有元素公用同一套(W,b),目的是为了在图片的各个位置都能检测到同一种特征.

Sub-sampling (reducing the spatial resolution)也就是池化是为了减少输出受到倾斜和扭曲的影响.

Scalable Object Detection using Deep Neural Networks

A different line of work, closer to ours, is based on the idea that objects can be localized without having to know their class.