Network In Network (NIN)

Min Lin et al.

① Problem Description

Abstract feature map + Classification

CNN: GLM (linear) + FC (dense)

maxout network: maximum pooling to make a piecewise linear approximator

② Problem Solution

① MLP Layer: universal function approximator

model various distributions of latent concepts.

② GAP: take the average of each feature map

① interpreted as categories confidence maps.

② prevents overfitting ③ more robust

③ Conceptual Understanding

① mlpconv Layer (multilayer perceptron)

$$f^1_{i,j,k_1} = \max\left(w_{k_1}^{1T} x_{i,j} + b_{k_1}, 0\right).$$

$$\vdots$$

$$f^n_{i,j,k_n} = \max\left(w_{k_n}^{nT} f^{n-1}_{i,j} + b_{k_n}, 0\right).$$

② GAP: Global Average Pooling.

(feature map) / (catagories)

CNNs: input → Conv. + Pool. → Conv. + Pool. → FC → output

(feature map) / (average) (catagories)

NIN: input → MLP → MLP → GAP → softmax → output

fine-tune: local receptive field size and weight decay. (use dropout)

Experiments: CIFAR-10, CIFAR-100, SVHN, MNIST

Improment: Perform object detection via NIN.