

# Network In Network

## Comment

## Abstract

- Utilizes micro neural network & global average pooling
- Easier to interpret and less prone to overfitting

## 1. Introduction

### Statistical meaning of *latent*

As opposed to *observable variable*, it is not directly observed but are rather inferred through mathematical model

- mlpconv layer - MLP consisting of multiple FC layers with nonlinear activation functions
- Via ***a global average pooling layer***, output the spatial average of the feature maps from the last mlpconv layer as the confidence of categories.
- FC layers are prone to overfitting and heavily depend on dropout regularization, whereas a global average pooling layer is itself a structural regularizer and prevents overfitting.

## 2. Convolutional Neural Networks

- Maxout network imposes the prior that instances of a latent concept lie within a convex set in the input space, which does not necessarily hold.
- Sliding a micro network over the input

## 3. Network In Network

### 3.1 MLP Convolution Layers

- Replaces the GLM to convolve over the input

$$\begin{aligned} f_{i,j,k_1}^1 &= \max(w_{k_1}^1{}^T x_{i,j} + b_{k_1}, 0) \\ &\vdots \\ f_{i,j,k_n}^n &= \max(w_{k_n}^n{}^T f_{i,j}^{n-1} + b_{k_n}, 0). \end{aligned}$$

- In case of maxout,

$$f_{i,j,k}^n = \max_m (w_{k_m}^T x_{i,j})$$

- Mlpconv layer differs from maxout layer in that the convex function approximator is replaced by a universal function approximator.

### 3.2 Global Average Pooling

- Replaces FC layers in CNN
- No parameters to optimize thus overfitting is avoided at that layer.
- More robust to spatial translations

### 3.3 Network In Network Structure

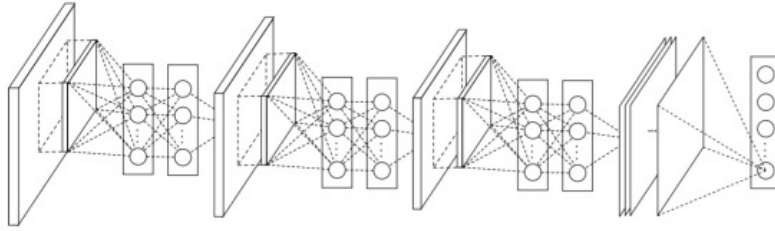


Figure 2: The overall structure of Network In Network. In this paper the NINs include the stacking of three mlpconv layers and one global average pooling layer.

Figure 1: enter image description here

## 4. Experiment

## 5. Conclusions