



explain ZFNet with layers and formulas



ZFNet, short for Zeiler and Fergus Network, is a convolutional neural network (CNN) architecture that was introduced in 2013 by Matthew Zeiler and Rob Fergus. It is an improvement over the AlexNet architecture and was designed to provide better visualization and understanding of the features learned by CNNs. ZFNet won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2013.

Key Improvements over AlexNet:

1. **Smaller Filter Size:** ZFNet uses smaller filter sizes in the first convolutional layer (7x7 instead of 11x11 in AlexNet). This allows the network to capture finer details in the input image.
2. **Stride:** ZFNet uses a smaller stride (2 instead of 4 in AlexNet) in the first convolutional layer, which helps in preserving more spatial information.

ZFNet Architecture:

The ZFNet architecture consists of 8 layers: 5 convolutional layers and 3 fully connected layers. Below is a detailed breakdown of the layers:

1. Input Layer:

- **Input Size:** 224x224x3 (RGB image)

2. Convolutional Layer 1 (Conv1):

- **Filter Size:** 7x7
- **Number of Filters:** 96
- **Stride:** 2
- **Padding:** Valid (no padding)
- **Activation Function:** ReLU (Rectified Linear Unit)
- **Output Size:** 110x110x96

Formula:

$$\text{Output Size} = \left\lfloor \frac{\text{Input Size} - \text{Filter Size} + 2 \times \text{Padding}}{\text{Stride}} \right\rfloor + 1$$
$$\text{Output Size} = \left\lfloor \frac{224 - 7 + 2 \times 0}{2} \right\rfloor + 1 = 110$$

3. Max Pooling Layer 1 (Pool1):

- **Pooling Size:** 3x3
- **Stride:** 2
- **Output Size:** 55x55x96

Formula:

$$\text{Output Size} = \left\lfloor \frac{\text{Input Size} - \text{Pooling Size}}{\text{Stride}} \right\rfloor + 1$$

$$\text{Output Size} = \left\lfloor \frac{110 - 3}{2} \right\rfloor + 1 = 55$$

4. Convolutional Layer 2 (Conv2):

- **Filter Size:** 5x5
- **Number of Filters:** 256
- **Stride:** 2
- **Padding:** Same (padding is added to make the output size the same as the input size)
- **Activation Function:** ReLU
- **Output Size:** 55x55x256

5. Max Pooling Layer 2 (Pool2):

- **Pooling Size:** 3x3
- **Stride:** 2
- **Output Size:** 27x27x256

6. Convolutional Layer 3 (Conv3):

- **Filter Size:** 3x3
- **Number of Filters:** 384
- **Stride:** 1
- **Padding:** Same
- **Activation Function:** ReLU
- **Output Size:** 27x27x384

7. Convolutional Layer 4 (Conv4):

- **Filter Size:** 3x3
- **Number of Filters:** 384
- **Stride:** 1

- **Padding:** Same
- **Activation Function:** ReLU
- **Output Size:** 27x27x384

8. Convolutional Layer 5 (Conv5):

- **Filter Size:** 3x3
- **Number of Filters:** 256
- **Stride:** 1
- **Padding:** Same
- **Activation Function:** ReLU
- **Output Size:** 27x27x256

9. Max Pooling Layer 3 (Pool3):

- **Pooling Size:** 3x3
- **Stride:** 2
- **Output Size:** 13x13x256

10. Fully Connected Layer 1 (FC1):

- **Number of Neurons:** 4096
- **Activation Function:** ReLU
- **Output Size:** 4096

11. Fully Connected Layer 2 (FC2):

- **Number of Neurons:** 4096
- **Activation Function:** ReLU
- **Output Size:** 4096

12. Fully Connected Layer 3 (FC3):

- **Number of Neurons:** 1000 (corresponding to the 1000 classes in ImageNet)
- **Activation Function:** Softmax
- **Output Size:** 1000

Key Formulas:

1. Convolution Operation:

$$\text{Output}(x, y) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \text{Input}(x + i, y + j) \times \text{Filter}(i, j) + \text{Bias}$$

1. ReLU Activation Function:

$$\text{ReLU}(x) = \max(0, x)$$

1. Softmax Activation Function:

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

where z_i is the input to the softmax function for class i , and K is the total number of classes.

Summary:

ZFNet is a CNN architecture that builds on AlexNet by using smaller filter sizes and strides in the initial layers to capture finer details. It consists of 5 convolutional layers followed by 3 fully connected layers. The network uses ReLU activation functions in the hidden layers and a softmax activation function in the output layer for classification. The architecture was designed to improve feature visualization and understanding, leading to better performance in image classification tasks.



 New chat

Message DeepSeek



NEW



AI-generated, for reference only