$Convolutional Neural Networks, CNN$

$convolution$

?$pooling non-$

$linear activation function$

$loss function$

$loss$

?$back-$

$propagation algorithm$

$1\times$

$13\times$

$35\times$

5

$stride$

$1\times$

1

$Sigmoid?Tanh?ReLU?$

$Sigmoid$

$Sigmoid$

$$f(x) = \frac{1}{1 + e^{-x}}$$

(1)

??$Sigmoid(0, 1)$

$Sigmoid$

$(0, 1)$

?

$Sigmoid$

$Tanh$

$Tanh$

$Sigmoid(0, 1)$

$sigmoid$

$Tanh$

$$tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$

(2)

$Tanh$??

$tanh x$

$tanh$

$ReLU$

$ReLU?$

$ReLU$

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

(3)

$ReLU x \geq$

0??

$x >$

$0 ReLU$

$ReLU x <$

0

$Parametric Rectified Linear Unit(PReLU)?$

$ReLU$

$average-$

$pooling max-$

$pooling$

$p-p-$

$norm p$

$1\times$

1

l_2

$$L_{loss} = -\frac{1}{N} \sum_{i=1}^{N}(y_i \log \hat{y_i} + (1 - y_i) \log (1 - \hat{y_i}))$$

(4)

L_{loss}

L_{loss}

y_i

$\hat{y_i}$

$\hat{y_i}$

$\hat{y_i}$

$Stochastic Gradient Descent,$

$$w = w - \eta \cdot dw$$

(5)

dw

η

$dw =$

0

?$momentum$

$$v = \mu \cdot v - \eta \cdot dw$$

(6)

μ

μ

v

η

dw

$root mean square prop?$