

Latihan IF3270
Senin 21 April 2025

1. Diberikan model CNN dengan arsitektur sebagai berikut (*complex layer terminology*) untuk klasifikasi gambar:

- input layer berukuran $32 \times 32 \times 3$,
- satu layer konvolusi dengan 5 kernel berukuran 7×7 , dengan padding =1 dan pergeseran 3 sel, tahap detector dengan ReLU, serta tahap maxpooling dengan kernel berukuran 4×4 , tanpa padding, dan pergeseran 2 sel.
- satu layer *fully connected (dense)* dengan jumlah neuron 10
- output layer 5 neuron dengan fungsi aktivasi softmax

Berikanlah jawaban dengan mencantumkan detail perhitungan.

a. (Nilai 5) Tentukanlah ukuran feature map yang dihasilkan dari layer konvolusi dengan mengisi tabel berikut ini. Detail perhitungan wajib diisi.

Tahap	Detail perhitungan	Ukuran Feature Map

b. (Nilai 7.5) Tentukanlah jumlah bobot (trainable parameter) pada model CNN tersebut (shared parameter = True). Output shape diisi dgn dimensi feature map (dalam format $v \times v \times d$) atau dimensi vektor (dalam format length)

Layer	Detail perhitungan	#Trainable Param	Output Shape
Konvolusi			
FC (hidden)			
FC (output)			
Total trainable param			-

Jawab:

1a.

Tahap	Detail perhitungan	Ukuran Feature Map
Konvolusi +Detector	$V=1+(W-F+2P)/S=1+(32-7+2*1)/3=10$	10*10*5
Pooling	$V=1+(W-F+2P)/S=1+(10-4+0)/2=4$	4*4*5

1b.

Layer	Detail perhitungan	#Trainable Param	Output Shape
Konvolusi	$5*(7*7*3+1)$	740	4*4*5
FC (hidden)	$(4*4*5+1)*10$	810	10
FC (output)	$(10+1)*5$	55	5
Total trainable param		1605	-

Kesalahan umum 1a:

- Lupa menambahkan suku 1 dalam formula $V=1+(W-F+2P)/S$
- Ukuran feature map: $V*V*\text{\#kernel}$
- Pooling mengubah dimensi, tetapi tidak mengubah kedalaman.
 $V*V*\text{\#kernel} \rightarrow V1*V1*\text{\#kernel}$

Kesalahan umum 1b:

- Kesalahan formula #trainable param
Number of weights with weight sharing:
 $= nbFeatureMaps * (filterLength * filterWidth * nbChannel) + 1 bias$
- Kesalahan menghitung ukuran flatten. FC (hidden) menerima input hasil flatten feature map layer convolution.
- Output shape untuk FC adalah panjang vektor.

Implementasi di keras:

```
from keras.models import Sequential
from keras import layers
model = Sequential()
```

```

model.add(layers.Conv2D(filters=5, kernel_size=(7,7), activation='relu',
input_shape=(32,32,3),strides=3))
model.add(layers.ZeroPadding2D(padding=1))
model.add(layers.MaxPooling2D(pool_size=(4,4),strides=2))
model.add(layers.Flatten())
model.add(layers.Dense(units=10, activation='relu'))
model.add(layers.Dense(units=5, activation = 'softmax'))

```

Catatan: Padding pada Conv2D tidak support custom padding.

```
model.summary()
```

Model: "sequential_7"

Layer (type)	Output Shape	Param #
conv2d_7 (Conv2D)	(None, 9, 9, 5)	740
zero_padding2d_17 (ZeroPadding2D)	(None, 11, 11, 5)	0
max_pooling2d_4 (MaxPooling2D)	(None, 4, 4, 5)	0
flatten_4 (Flatten)	(None, 80)	0
dense_8 (Dense)	(None, 10)	810
dense_9 (Dense)	(None, 5)	55

Total params: 1,605 (6.27 KB)

Trainable params: 1,605 (6.27 KB)

Non-trainable params: 0 (0.00 B)

2. (Nilai 7.5) Lakukanlah forward propagation untuk mendapatkan feature map dari suatu layer konvolusi jika diberikan data input (\mathbf{x}) berukuran $3 \times 3 \times 2$ pada suatu layer konvolusi yang memiliki 1 kernel berukuran 2×2 . Stride yang digunakan 1, dan tanpa menggunakan padding. Tahap detector dengan ReLU, dan tahap pooling dengan fungsi maksimum dengan kernel berukuran 2×2 , tanpa padding, dan pergeseran 2 sel.

Berikanlah secara eksplisit feature map per tahap dari layer konvolusi tersebut.

Data input \mathbf{x}

1	0	1
0	0	1
1	1	0

1	0	1
0	1	0
1	0	1

Kernel (bias 0.5)

1	0
0	-1

0	-1
1	0

Jawab:

Convolution Stage: Example

Example 2: input size $3 \times 3 \times 2$ ($W=3$); 3 filters $2 \times 2 \times 2$ ($K=3$, $F=2$); moving stride $S=1$; input padding $P=0$:

$$V = ([3 - 2 + 2 \times 0] / 1) + 1 = 2$$

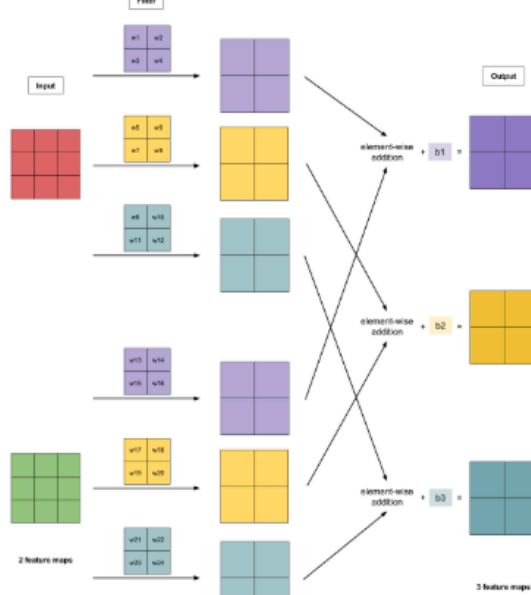
Output volume: $2 \times 2 \times 3$

Example 3: input size $32 \times 32 \times 3$; 10 filters of size $5 \times 5 \times 3$; single stride; and no zero padding $\rightarrow W=32$, $K=10$, $F=5$, $P=0$, $S=1$:

$$V = ([32 - 5 + 0] / 1) + 1 = 28$$

Output volume: $28 \times 28 \times 10$.

Convolution of 2-channel image with 2×2 filter to output 3 channels



<https://towardsdatascience.com/counting-no-of-parameters-in-deep-learning-models-by-hand-8f171624188947#5137>

Tahap konvolusi

$1+0+0+0+0+0+0+0=1+0.5=1.5$	$0+0+0+-1+0+-1+1+0=-1+0.5=-0.5$
$0+0+0+-1+0+-1+1+0=-1+0.5=-0.5$	$0+0+0+0+0+0+0+0=0+0.5=0.5$

Tahap detector

1.5	0
0	0.5

Tahap maxpooling

1.5
