Fundamentals of CNNs II

Course:
INFO-6152 Deep Learning with Tensorflow & Keras 2

Developed by:
Mohammad Noorchenarboo

September 22, 2025

Current Section

The Pooling Layer

In medical imaging or weather prediction, we often deal with huge images containing millions of pixels. **Question:** How can we reduce the size of these images while still keeping the most relevant features for deep learning models?

Concept: Pooling

Pooling is a **downsampling operation** applied to feature maps.
For a region R of size $m \times n$, the pooling function P produces one output value:

$$F'(i,j) = P\big(\{F(u,v) \mid (u,v) \in R(i,j)\}\big)$$

Common pooling types:

- **Max pooling:** $P(R) = \max(R)$
- **Average pooling:** $P(R) = \frac{1}{|R|} \sum_{(u,v)\in R} F(u,v)$

Output size formula:

$$H_{out} = \frac{H - m + 2p}{s} + 1, \quad W_{out} = \frac{W - n + 2p}{s} + 1$$

The Pooling Layer

TensorFlow: Max Pooling Layer

```
import tensorflow as tf

# MaxPooling after a Conv layer
model = tf.keras.Sequential([
    tf.keras.layers.Conv2D(filters=32, kernel_size=(3,3),
                           activation='relu', input_shape=(28,28,1)),
    tf.keras.layers.MaxPooling2D(pool_size=(2,2), strides=2)
])

model.summary()
# Expected: Feature map reduced in size by pooling
```

The Pooling Layer

Real-World and Numerical Examples

- **Medical imaging:** reduces resolution but keeps essential tumor boundaries.
- **Speech recognition:** compresses spectrograms while retaining sound patterns.
- **Weather prediction:** shrinks satellite images while preserving storm shapes.
- **Numerical example:**

  Input 4×4 matrix: $\begin{bmatrix} 1 & 3 & 2 & 4 \\ 5 & 6 & 1 & 2 \\ 2 & 4 & 0 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix}$

  Apply 2×2 max pooling with stride 2: $\begin{bmatrix} 6 & 4 \\ 4 & 4 \end{bmatrix}$

The Pooling Layer

Common Mistakes

- Choosing pooling windows that are too large $\rightarrow$ lose too much detail.
- Stacking many pooling layers $\rightarrow$ excessive information loss.

Caveats

- Pooling reduces dimensions but also discards some spatial precision.
- Modern CNNs often replace pooling with strided convolutions for better feature preservation.

Get more info $\rightarrow$ DeepLearning.AI: Pooling Explained

Summary: The Pooling Layer

Concept	Key Point
Pooling definition	Downsampling feature maps by summarizing local regions
Max pooling	Selects maximum value from region
Average pooling	Computes average value from region
Output size formula	$H_{out} = \frac{H-m+2p}{s} + 1, \quad W_{out} = \frac{W-n+2p}{s} + 1$
Benefits	Reduces dimensions, prevents overfitting, increases efficiency
Caveats	May lose spatial detail; can be replaced by strided convolutions

Current Section

1. The Pooling Layer
2. Global Pooling

Global Pooling

In medical imaging, after extracting features from an X-ray using deep CNNs, we often need to reduce the entire feature map into a single value per channel to make a classification decision (e.g., healthy vs. cancer). **Question:** How can we shrink a whole feature map into one number while preserving key information?

Concept: Global Pooling

Global Pooling reduces each entire feature map into a single value.
Types:

- **Global Average Pooling (GAP):**

$$y_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_c(i,j)$$

- **Global Max Pooling (GMP):**

$$y_c = \max_{i,j} F_c(i,j)$$

Here $F_c(i,j)$ is the activation at location (i,j) in channel c.
Key property: Output dimension depends only on the number of channels, not spatial size.

Global Pooling

TensorFlow: Global Average Pooling Example

```python
import tensorflow as tf

model = tf.keras.Sequential([
    tf.keras.layers.Conv2D(32, (3,3), activation='relu',
                           input_shape=(28,28,1)),
    tf.keras.layers.GlobalAveragePooling2D()
])

model.summary()
# Expected: Output shape is (None, 32)
# Each of 32 feature maps reduced to 1 number
```

Real-World Examples

- **Image classification:** replaces fully connected layers, reducing parameters.
- **Medical imaging:** each channel summarizes one diagnostic pattern.
- **MobileNets / EfficientNet:** rely on GAP to achieve lightweight architectures.

Global Pooling

Common Mistakes

- Confusing global pooling with normal pooling (global pooling always outputs 1×1 per channel).
- Using global pooling when fine spatial detail is needed (e.g., segmentation).

Caveats

- Removes all spatial information — cannot recover positions of features.
- Works best for classification but not for tasks like detection or segmentation.

Global Pooling as Taking a Final Exam Grade

Imagine a student's grades across many assignments (pixels). - **Global Average Pooling:** compute the average grade — overall performance. - **Global Max Pooling:** keep the highest grade achieved — best performance highlight.

Get more info → Keras: GlobalAveragePooling2D

Summary: Global Pooling

Concept	Key Point
Global pooling definition	Reduces entire feature map to one value per channel
Global Average Pooling (GAP)	Computes average of all spatial values
Global Max Pooling (GMP)	Takes maximum value across spatial map
Output size	Depends only on channels (e.g., C feature maps $\rightarrow$ vector of length C)
Benefits	Fewer parameters, prevents overfitting, used in lightweight models
Caveats	Loses spatial information, unsuitable for detection/segmentation