

Pooling Layers in CNNs

Erarbeitet von

Dr. Ludmila Himmelspach

Lernziele	1
Inhalt	2
Einstieg.....	2
Pooling Layer	2
Take-Home Message	3
Quellen	3
Weiterführendes Material.....	3
Disclaimer	4

Lernziele

- Du kannst erklären, warum Pooling Layer in CNNs eingesetzt werden
- Du kannst die Funktionsweise eines Pooling Layers erklären
- Du kannst Parameter nennen, die in Pooling Layern gewählt werden

Inhalt

Einstieg

Du weißt inzwischen, dass ein Convolutional Neural Network aus mehreren aufeinanderfolgenden Convolutional Layern besteht. In diesem Video lernst Du einen weiteren Baustein eines typischen CNNs kennen: den sogenannten *Pooling Layer*.

Pooling Layer

Bei der Verarbeitung von größeren und komplexeren Bildern hat man gemerkt, dass CNNs wegen einer großen Anzahl an Optimierungsparametern sehr rechen- und speicherlastig sind. Außerdem neigen CNNs zum Overfitting, weil sie zu viele Details aus Bildern extrahieren. Deswegen hat man die Idee gehabt, die Größe der Feature Maps in Convolutional Layern zu verkleinern bzw. die Information in Feature Maps zu aggregieren, in dem aus einer kleinen Nachbarschaft von Neuronen entweder ein Mittelwert gebildet oder ein Maximalwert ausgewählt wird. Diese Aufgabe wird von sogenannten *Pooling Layern* übernommen, die zwischen Convolutional Layern eingesetzt werden.

Quelle [1]

In der Abbildung siehst Du die Funktionsweise eines *Max-Pooling Layers*, der in CNNs oft seine Anwendung findet.

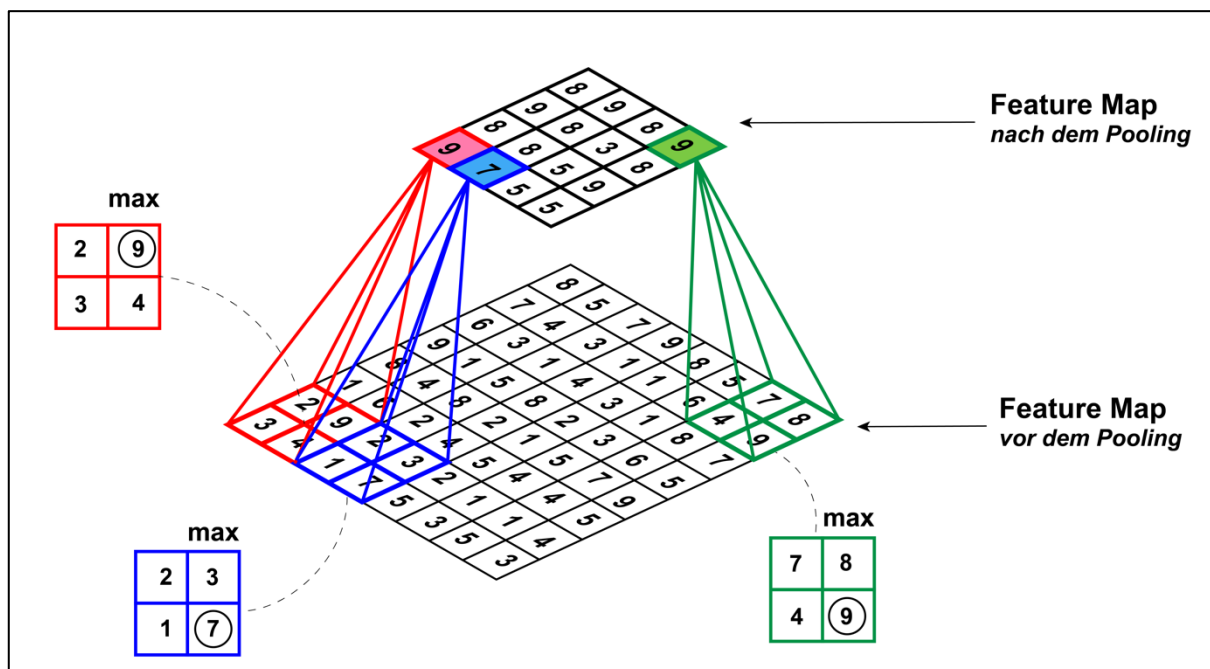


Abbildung 1: Max-Pooling Layer (Wahrnehmungsfeld der Größe 2 x 2, Schrittweite 2).

Unten befindet sich eine Feature Map eines Convolutional Layers und oben ist die resultierende Feature Map nach der Anwendung des Max-Poolings. In Pooling Layern gibt es

keine Parameter, die trainiert werden müssen. Du kannst lediglich die Art der Aggregation, die Größe des Wahrnehmungsfeldes und die Schrittweite bzw. den Grad der Überlappung der Wahrnehmungsfelder bestimmen. In diesem Max-Pooling Layer wurde das Wahrnehmungsfeld der Größe 2×2 und die Schrittweite 2 gewählt. Deswegen besitzt die resultierende Feature Map nur die halbe Höhe und die halbe Breite der ursprünglichen Feature Map. Wie der Name schon sagt, wird durch den Einsatz eines Max-Pooling Layers nur der größte Wert aus jedem Wahrnehmungsfeld in die nächste Schicht übergeben. Alle anderen Werte werden dabei verworfen.

Die Pooling Layer sind offensichtlich destruktiv. Um den Verlust der Information zu kompensieren wird oft in darauffolgenden Schichten die Filteranzahl und somit die Anzahl der Feature Maps verdoppelt. Wenn man sich zum Beispiel im ersten Convolutional Layer auf 64 Filter festgelegt hat, so erhöht man die Anzahl der Filter nach dem Pooling im zweiten Convolutional Layer auf 128. Diese Vorgehensweise hat sich in der Praxis als gut erwiesen, weil die Erhöhung der Filter bzw. der Feature Maps in oberen Schichten mehr Möglichkeiten ergibt, die einfacheren Muster zu komplexeren Mustern zu kombinieren.

Take-Home Message

In diesem Video hast Du die Funktionsweise eines weiteren Bauelements von CNNs, den sogenannten *Pooling Layers*, kennengelernt. Du hast erfahren, dass Pooling Layer zur Verringerung der Optimierungsparameter in Convolutional Layern führen und somit neben der Rechenlast auch das Risiko für Overfitting senken.

Quellen

- Quelle [1] Géron, A. (2023). *Praxiseinstieg Machine Learning mit Scikit-Learn, Keras und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme* (aktualisierte und erweiterte Auflage.). O'Reilly Verlag.

Weiterführendes Material

Chollet, F. (2018). *Deep Learning mit Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek*. MITP-Verlags GmbH & Co. KG.

Disclaimer

Transkript zu dem Video „Bildklassifikation und Bildsegmentierung: Pooling Layers in CNNs“, Dr. Ludmila Himmelspach.

Dieses Transkript wurde im Rahmen des Projekts ai4all des Heine Center for Artificial Intelligence and Data Science (HeiCAD) an der Heinrich-Heine-Universität Düsseldorf unter der Creative Commons Lizenz [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.