

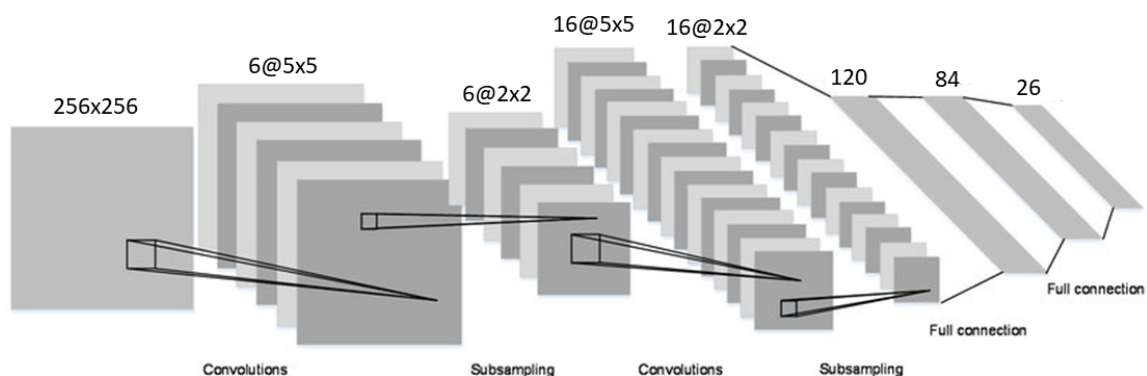
Übungsblatt: 10

Bearbeitung am 16. Juli

Aufgabe 1: Klassifikation

Gegeben sei das folgende Netz zur Klassifizierung von Großbuchstaben in Bildern der Größe 256x256 (Graustufen), wobei die Conv-Layer Valid-Padding haben und die Aktivierungsfunktion ReLU sei. Dabei können mehrere verschiedene Buchstaben auf einem Bild vorkommen. Jede Art von Buchstabe kann aber nur höchstens einmal auf einem Bild vorkommen.

Die Zahlen über den Schichten geben jeweils die Anzahl der Feature Maps und die Filtergröße bzw. die Anzahl an Knoten an.



- Geben Sie den Pseudo-Code und die Ausgabedimension aller Layer für die Implementierung des Netzwerks an.
- Geben Sie eine passende Lossfunktion (mit zugehöriger Formel für ein Trainingsbeispiel) für das Training an.
- Welche Formen der Regularisierung könnte man einbauen und welche Operationen der Datenaugmentierung sind hier sinnvoll? Warum/Warum nicht?
- Wie müssen das Netz und die Lossfunktion angepasst werden, damit man nur genau einen Buchstaben pro Bild klassifiziert?

Aufgabe 2: Sequenzklassifizierung

Im Folgenden soll das Netz aus Aufgabe 1 zur Sequenzerkennung von Buchstaben erweitert werden. Dafür nehmen wir vereinfacht an, dass nur die Buchstaben von A bis J vorkommen können. Es sollen Sequenzen von 0 bis 5 Buchstaben erkannt werden können.

- Gegeben sei die folgende Ausgabe des Netzes y_k^t . Dekodieren Sie die Sequenz mittels des CTC-Greedy-Decoders.

y_k^t	0	1	2	3	4
-	0.9	0.1	0.2	0.3	0.15
A	0	0.8	0.1	0	0
B	0	0	0.7	0.4	0
C	0	0	0	0	0
D	0	0	0	0	0.1
E	0	0	0	0.3	0.15
F	0	0	0	0	0.1
G	0.1	0.1	0	0	0.15
H	0	0	0	0	0.15
I	0	0	0	0	0.2
J	0	0	0	0	0

- (b) Die Ground-Truth für die Ausgabe ist eigentlich AI. Berechnen Sie die Forward- und Backward-Variablen α und β . Berechnen Sie damit die Gesamtwahrscheinlichkeit von AI.

Aufgabe 3: Bildsegmentierung und Objekterkennung

- (a) Wie unterscheiden sich Bildsegmentierung und Objekterkennung, d.h. wie wird klassifiziert, was sind die Ausgaben etc.?
- (b) Gegeben sei das bekannte FCN U-Net (Abbildung ??) mit der Variation, dass die Conv-Layer Padding = Same haben. Die Up-Convolutions sollen dabei die gleiche Filtergröße wie die normalen Convolutions haben. Geben Sie den Pseudocode (ohne Dimensionen) zu dieser Architektur an. Die Aktivierungsfunktion kann hier als Attribut der Conv-Layer aufgefasst werden. Als Ausgabe sollen 10 verschiedene, exklusive Klassen möglich sein.

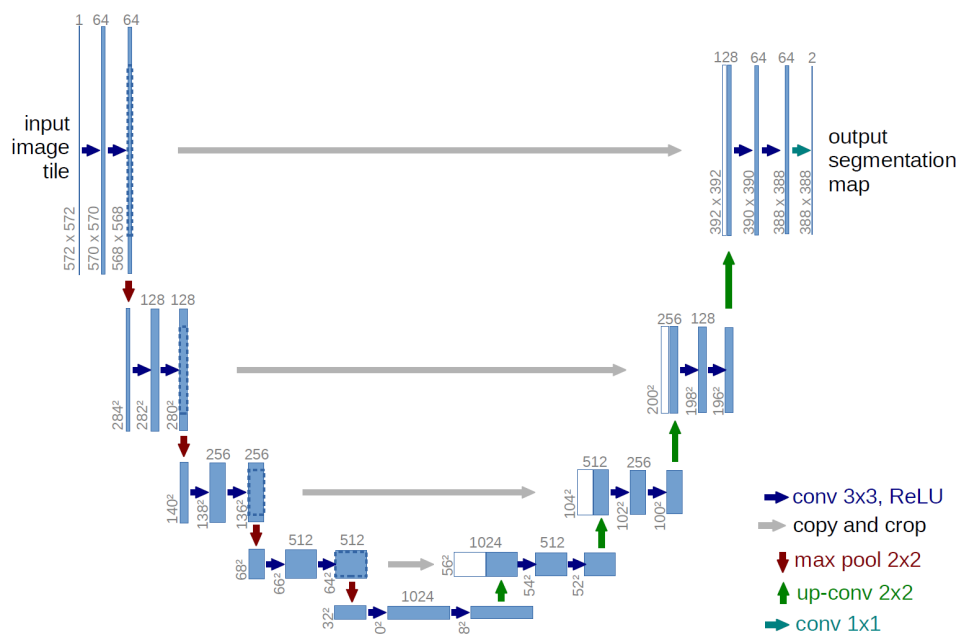


Abbildung 1: U-Net

- (c) Listen Sie die (Haupt-)Ansätze der Klassifikation in der Objekterkennung auf und nennen Sie jeweils eine Beispielarchitektur.

(d) Welche Arten von Segmentierung gibt es? Wodurch zeichnen sie sich aus?

Aufgabe 4: Medizinische Anwendung

Betrachten Sie die folgende Anwendung in der Medizin: Sie möchten die Diagnose von Ärzten unterstützen, indem Sie aus Röntgen-Scans das (gesamte) Lungenvolumen bzw. als ersten Schritt die Lungenfläche des Patienten bestimmen.

Beschreiben Sie, welche Art von Neuronalem Netz Sie einsetzen könnten und welche Art von Daten die Mediziner zum Training zur Verfügung stellen müssten.

