

Übungsblatt: 9

Bearbeitung am 2. Juli

Aufgabe 1: CTC-Algorithmus

Gegeben sind Bilder von Autoschildern, die eine verschiedene Anzahl von 29 Großbuchstaben, 10 Ziffern und das Leerzeichen zeigen können (OCR-Task).

- (a) Entwerfen Sie in Pseudo-Sprache ein Neuronales Netz zur OCR, das aus Conv- und Pooling-Layern besteht und ein Bi-Direktionales LSTM („BiLSTM n -nodes“) besitzt. Reshappen Sie Channels und Höhe zu einer Dimension, um die Eingabe des BiLSTM zu generieren. Als Aktivierungsfunktion der Ausgabe mittels eines FC-Layers (angewandt pro horizontaler Position) dient ein Softmax-Layer um Wahrscheinlichkeiten zu generieren. Geben Sie für jedes Layer die Ausgabedimensionen an.

Lösung:

Input	(W x H x C)
Conv2D 3x3, 20, Padding=Same	(W x H x 20)
MaxPool2D 2x2	(W/2 x H/2 x 20)
Conv2D 3x3, 40, Padding=Same	(W/2 x H/2 x 40)
Reshape W/2, H/2 * 40	(W/2 x (H/2 * 40))
BiLSTM 50	(W/2 x 100)
FC 40	(W/2 x 40)
Softmax	(W/2 x 40)

Hinweise: Padding Same ist nicht erforderlich. Die FC-Dimension ergibt sich, da es die letzte Schicht ist, aus der Alphabetgröße. Alternativ kann direkt ein BiLSTM mit 40 Nodes verwendet werden.

- (b) Welche Form von GT-Daten erwartet die CTC-Loss-Funktion?

Lösung: Paare an Eingabe-Bild und Label-Sequenz ($U \leq T$)

- (c) Gegeben Sei die Label-Sequenz „NEW AR 565“ und folgende Wahrscheinlichkeitstabelle (Ausgabe des Softmax):

Berechnen Sie die Forward- und Backward-Variablen, sowie die Gradienten-Updates für die gegebenen Daten. Welche Dimension haben α und β jeweils?

Achten Sie bei der Berechnung auf Paper, dass etliche Variablen Null sind! Alternativ können die das Excel-Sheet (siehe Wue-Campus) nehmen, um die Lösungen Algorithmus zu bestimmen.

Welche Sequenz gibt der Greedy-Decoder aus?

$P(L,T)$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
-	0.9	0.8	0.3	0.6	0.8	0.9	0.1	0.1	0.99	0.1	0.1	0.8	0.4	0.1	0.3	0.9	0.4	0.9
A	0	0	0	0	0	0	0	0	0	0	0.9	0.2	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0.2	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0.1	0	0	0	0	0	0	0	0
I	0.1	0	0	0	0	0.1	0	0	0	0	0	0	0.3	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0.01	0	0	0	0	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0.2	0.7	0.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0	0	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V	0	0	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0
W	0	0	0	0	0	0	0.4	0.9	0	0	0	0	0	0	0	0	0	0
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Z	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0	0.6	0.1
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.1	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0.9	0	0	0	0

Lösung: Siehe Excel-Tabelle