

Proseminar „Convolutional Neural Networks - Methoden und Anwendungen“

Bilderkennung mit Vgg

Justin Schartner

14. Mai 2022

Struktur

- 1** Einleitung
- 2 Allgemeines
- 3 Architektur
- 4 Training
- 5 Beispiel
- 6 Bewertung
- 7 Ausblick

Problemstellung

Gibt es CNN-Architekturen, welche Bilder noch besser, als schon bekannte Architekturen klassifizieren können?

Lösung

Durch die Erhöhung der Tiefe eines CNN verspricht man sich genauere Aussagen über Bilder machen zu können.

Ergebnis

Die VGG-Architektur hat bewiesen, dass die Tiefe eines CNN, eine ausschlaggebende Komponente hinsichtlich der Bilder-Klassifizierung ist.

Gliederung

- Was ist VGG? **Allgemeines**
- Wie funktioniert VGG, was macht es besonders? **Architektur**
- Wie trainiert man ein VGG-net, was ist wichtig? **Training**
- Wie implementiert man ein VGG-net? **Beispiel**
- Wie gut ist VGG? **Bewertung**

Struktur

- 1 Einleitung
- 2 Allgemeines**
- 3 Architektur
- 4 Training
- 5 Beispiel
- 6 Bewertung
- 7 Ausblick

Idee

Steigerung der Genauigkeit durch:

- Steigerung der Tiefe, des CNNs
- Schachtelung von Convolutional-Layer-Blöcken
- Einsatz von kleinen 3x3-Filtern und einer Stride von 1

Struktur

- 1 Einleitung
- 2 Allgemeines
- 3 Architektur**
- 4 Training
- 5 Beispiel
- 6 Bewertung
- 7 Ausblick

Architekturarten

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 x 224 RGB-image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Architektur

Convolutional Layers

- receptive field: 3×3 , 1×1
- activation: ReLU, stride: 1, padding: 1, channels: 64, 128, 512, 512

Pooling Layer

- Max-Pool
- field: 2×2 , stride: 2

Fully-Connected Layers

- activation: ReLU
- channels: 4096, 4096, 1000

Softmax Layer

Convolutional Layers



Struktur

- 1 Einleitung
- 2 Allgemeines
- 3 Architektur
- 4 Training**
- 5 Beispiel
- 6 Bewertung
- 7 Ausblick

Struktur

- 1 Einleitung
- 2 Allgemeines
- 3 Architektur
- 4 Training
- 5 Beispiel**
- 6 Bewertung
- 7 Ausblick

Vgg initialisieren

Vgg trainieren

Vgg benutzen

Struktur

- 1 Einleitung
- 2 Allgemeines
- 3 Architektur
- 4 Training
- 5 Beispiel
- 6 Bewertung**
- 7 Ausblick

Komplexität

Vergleich

Struktur

- 1 Einleitung
- 2 Allgemeines
- 3 Architektur
- 4 Training
- 5 Beispiel
- 6 Bewertung
- 7 Ausblick**

Ausblick