

# Word Spotting und Texterkennung - PHOC-Net

## Notizen

Maximilian Springenberg

TU-Dortmund

## 0 Fragen

---

Aufbau BA: Attribut CNNs für die Einzelworterkennung

1. Einleitung
2. Grundlagen
  - CNN (kurz)
  - MLP (kurz)
  - PHOCNet (GLMs nicht notwendig, TTP schon)
3. verwandte Arbeiten
  - LSTM
4. Methodik
  - Anwendung auf Worterkennung: ?
5. Experimente
  - Distanz Metriken (Cosinus Distanz)
  - Lexikon / Geschätzte Attributvektoren
  - CCA / Probabilistisches Modell
  - Datensätze: RIMS, IAM, Washington
  - Metriken: Wort/ Character-Fehlerrate
6. Fazit

Umfang Exposee: 2-3 Seiten

# 1 PHOCNet

---

## 1.1 Jurnal

### 1.1.1 Begriffe

- Word-Spotting: Rückgabelisten basierter Ansatz.
- QbE: Bild von einem Word als Query
- QbS: Text des Wortes (String) als Query
- PHOC: Pyramidal Histogram of Characters (encodiert als Fisher Vector)
- Attribute SVMs: konkret in der Jurnal-Version Menge an SVMs, die jeweils ein Attribut des PHOC schätzen
- Temporal Pyramid Pooling layer: ?
- Framework für Fehlerfunktion: GLM (Generalized Linear Models)
- Attribute CNNs: ?
- VGG16: siehe Very Deep Convolutional Networks For Large-Scale Image Recognition

### 1.1.2 Ansatz