

Konzept Master-Konferenzseminar

Sebastian Schmidt

Thema: Erkennung japanischer Schriftzeichen mittels maschinellen Lernens

- (1) Anmerkung: Der Name des Themas wurde von „Erkennung japanischer Schriftzeichen mittels Konvolutionaler Neuronaler Netze“ geändert, da wahrscheinlich ebenfalls Vision Transformer getestet werden, bei welchen es sich nicht um Konvolutionale Neuronale Netze handelt.

Abstract

Das japanische Schriftsystem ist mit drei verschiedenen Alphabeten (Hiragana, Katakana, Kanji) eines der komplexesten Schriftsysteme der Welt. Eine Person muss dabei im Japanischen ca. 2000 Schriftzeichen erlernen, um eine Zeitung vollständig verstehen zu können. In dem Paper „Recognizing Handwritten Japanese Characters Using Deep Convolutional Neural Networks“ von Charlie Tsai wurde eine Erkennung mittels der VGGNet Netzen, speziellen Architekturen Konvolutionaler Neuronaler Netze, erprobt. Im Ansatz wurde dabei sowohl die separate Erkennung aller Alphabete als auch die gemeinsame aller möglichen Schriftzeichen erprobt. Dabei konnten Testgenauigkeiten von bis zu 99,53% erreicht werden. In diesem Paper sollen die Ergebnisse nachvollzogen und mit weiteren Mitteln die Qualität des Modells überprüft werden. Dies soll den Einsatz und klaren Vergleich weiterer Netzwerkarchitekturen wie beispielsweise der Vision Transformer oder ResNet ermöglichen. Zusätzlich sollen gegebenenfalls nach erfolgreicher Recherche weitere moderne Netzwerkarchitekturen gesichtet und mit dem implementierten Mitteln überprüft werden. Auch ein Einsatz weiterer Datensätze zur Verallgemeinerung der Daten auf nicht handgeschriebene Zeichen und einer Unterstützung des Lernprozesses wäre denkbar. Zuletzt sei bei erfolgreicher Umsetzung der vorherigen Punkte eine weitere Verallgemeinerung denkbar bei welcher durch die Extraktion einzelner Zeichen aus Textpassagen eine Erkennung ganzer japanischer Texte ermöglicht wird.

- (2) Anmerkung: Der Abstract wurde leicht angepasst, um den geänderten Titel zu berücksichtigen.

Gliederung

1. Einleitung
 - Kurze Informationen über das japanische Schriftsystem
 - Warum wollen wir japanische Schriftzeichen automatisch erkennen?
 - Auf die Hauptquelle eingehen und die Ergebnisse dieser kurz darstellen.
2. Grundlagen
 1. Konvolutionale Neuronale Netze
 - Kurzer Anriss der grundlegenden Idee hinter Konvolutionalen Neuronalen Netzen
 2. Vision Transformer
 - Kurze Erklärung der Konzepte hinter Transformer
 - Kurze Erklärung der Vision Transformer
3. Übersicht der Daten
 - Informationen zu den verwendeten Daten (Anzahl, Imbalance, ggf. weitere Besonderheiten)

4. Eingesetzte Modelle

1. VGG Netze

- Aufbau der konvolutionalen VGG Netze (VGG16/ VGG19)

2. ResNet

- Aufbau des ResNet zur Bildklassifizierung

3. Vision Transformer

- Aufbau des eingesetzten Vision Transformer

5. Darstellung der Ergebnisse

- Darstellung der Ergebnisse mit Accuracy und durchschnittlichen F1-Scores
- Konnten die Ergebnisse der Hauptquelle nachvollzogen oder verbessert werden?
- ggf. auf Besonderheiten eingehen und Unerwartetes darstellen

6. Fazit

- Ergebnisse zusammenfassen
- Kurze Bewertung zu Ergebnissen mit Bezug auf Hauptquelle geben
- Ausblicke über weitere mögliche Schritte (Neue Modelle, Einsatzmöglichkeiten der Modelle)

- (3) Anmerkung: Die eingesetzten Modelle können sich im Verlauf der nächsten Zeit noch ändern. Es ist mir daher nicht möglich zu versichern, ob Kapitel 4 tatsächlich so aufgebaut wird.

Hauptquelle

Tsai, Charlie. "Recognizing Handwritten Japanese Characters Using Deep Convolutional Neural Networks." (2016). http://cs231n.stanford.edu/reports/2016/pdfs/262_Report.pdf