

Protokoll: Projekt – Machine Learning

Projektidee

- Video frame interpolation
 - Der Input sind zwei Frames, ein Startframe und ein Endframe.
Der Output soll ein/mehrere Frames sein, die dazwischen liegen.

2. Treffen 18.10.2021

Was haben wir geschafft?

- Einarbeit in Variational Autoencoders
- Betrachtung & Ausführung von Code für VAE mit MNIST-Datensatz

Was tun wir bis nächste Woche?

- Aufgaben
 - Modifikation des MNIST-Netzes mit RGB-Farben & höherer Auflösung
 - Johan & Quan
 - Erzeugung von Testdaten & Ausführung von MNIST-Code
 - Justus & Jannik

3. Treffen 25.10.2021

Was haben wir geschafft?

- Umstellung auf RGB-Bilder
 - Umstellung des MNIST-Template-Codes
- Anpassung der Netzarchitektur für eine höhere Auflösung
 - Nachvollziehen der Architektur und damit des "Template"-Quellcodes
- Nachvollziehen der Wertebereiche des "latent"-space
- Erfolgreiches Training des VAE auf Gesichts-Bilder (256x256x3 Emoji-Bilder)

Was tun wir bis nächste Woche?

- Aufgaben
 - Erste Schritte für das Interpolieren
 - "Mitteln" zweier Trainingsdaten-Latent-Space-Vektoren zum Erstellen eines "Übergang"-Bildes
 - Macht das gesamte Team zusammen um ein besseres Verständnis zu bekommen (Johan, Justus, Quan, Jannik)
 - Nachvollziehen von sinnvollen Wertebereichen / Methoden zur Mittelung zweier Bilder
 - Macht das gesamte Team zusammen um ein besseres Verständnis zu bekommen (Johan, Justus, Quan, Jannik)
 - Experimentieren mit der Fragestellung: "Ist es möglich eines der beiden Quell-Bilder zu einem höheren Einfluss zu gewichten?"

- Macht das gesamte Team zusammen um ein besseres Verständnis zu bekommen (Johan, Justus, Quan, Jannik)

4. Treffen 01.11.2021

Was haben wir geschafft?

- Code neu strukturiert
 - Möglichkeit zum Speichern und Laden der Gewichte
 - Sortierung nach Methoden / Skript
- Weitere Experimente mit verschiedenen Latentdimensions
 - Ergebnis: 50 schließt ähnlich gut ab wie 100
 - Normalverteilungsvektor lässt jedoch immer noch keine zufriedenstellenden Bilder generieren
- 3-Bilder-Datensatz gefunden für zukünftige Experimente in Richtung Animation
- Training über 200 Epochen führte zu ähnlichen Ergebnissen wie nach 50 Epochen.

Was tun wir bis nächste Woche?

- Quan: Skript schreiben zum extrahieren der 3-Bilder-Datensätze
- Johan: Experimente mit der Architektur des neuronalen Netzes
- Jannik & Justus: Sigma- / Mü-Experimente, welchen Einfluss haben Sigma und Mü auf unsere Ergebnisse?

5. Treffen 08.11.2021

Was haben wir geschafft?

- Sigma, My und Epsilon geprüft: Auf Sigma und My haben wir keinen Einfluss, da dies erlernbare Parameter des Netzes sind. Epsilon wird zufällig aus einer Normalverteilung gewählt. Nach ersten Experimenten, scheint ein leichtes Stauchen oder Strecken der Normalverteilung keinen nennenswerten Effekt auf das Netz zu haben.
- Vervielfachung der Parameter im Netz durch Architekturveränderungen: Hier scheint das Netz zwar etwas schneller zu lernen, jedoch schmiegt sich auch hier das Netz recht schnell an einen ähnlichen Loss an.
- Batchweises einlesen der Trainingsdaten ins Netz, um RAM-Überlast zu vermeiden.
- Extrahieren des drei-Bilder Datensatzes funktioniert nun.

Was tun wir bis nächste Woche?

- Johan: Alternativ Architektur von einem Artikel ausprobieren. Diese nutzt Batchnormalization und scheint recht gut zu funktionieren.
- Quan: Trainingszyklus überarbeiten mit Batchweise-Einlesen.
- Jannik & Justus: VQ-VAE anschauen. Könnte ein spannender Ansatz sein.
- Alle: Genauere Punkte festlegen: Was bringt uns nun näher an unser Ziel?

6. Treffen 15.11.2021

Was haben wir geschafft?

- Angleichen unserer Erkenntnisse aus den Arbeiten der Vorwoche

- Gemeinsame Überlegung einer neuen Strategie für die Zukunft des Projektes
- Recherche: Lesen einiger Paper, um eine neue Richtung zu finden
- Alle gucken sich das Paper "Learning Image Matching by Simply Watching Video" an

Was tun wir bis nächste Woche?

- Justus: Kontaktaufnahme mit den Autoren des oben genannten Papers, Architektur aus o.g. Papers in TF Keras umzusetzen
- Jannik: Architektur aus o.g. Paper in TF Keras umzusetzen
- Quan: Weitere Paper auschecken, im Besonderen Optical Flow Estimation
- Johan: Ähnlich wie Quan, vielleicht im Bereich der Depth Maps speziell schauen

7. Treffen 22.11.2021

Was haben wir geschafft?

- Entscheidung getroffen für eine Technologie
 - Paper: "Learning Image Matching by Simply Watching"
- Recherche zum Einlesen des Triplet-Datensatz, sodass der Zusammenhang der 3 Frame-Sequenz beachtet wird

Was tun wir bis nächste Woche?

- Justus & Quan: Preprocessing des Datensatzes & Input des Netzes testen durch manuellen Input
- Johan & Jannik: Visualisierung des Trainings vom Netz

8. Treffen 29.11.2021

Was haben wir geschafft?

- Einlesen und vorverarbeiten der Test- und Trainingsdaten
- Training für 1h mit dem Vimeo90k Datensatz
- Inkrementelle Speicherung der weights
- Darstellung des Losses
- Darstellung des Outputs als Gifs

Was tun wir bis nächste Woche?

- Netz noch länger trainieren
- Zwischenbild mit zwei weiter auseinander liegenden Bildern generieren
- Eine Serie von Zwischenbildern generieren
- Zusätzliche Informationen für die Loss Berechnung nutzen
- Die Möglichkeit evaluieren ein anderes vortrainiertes Netz mit in unsere Architektur zu integrieren

8. Treffen 06.12.2021

Was haben wir geschafft?

- Netz noch länger trainieren
- Eine Serie von Zwischenbildern generieren
- Die Möglichkeit evaluieren ein anderes vortrainiertes Netz mit in unsere Architektur zu integrieren

- Weiteres Vorgehen in der Veranstaltung diskutiert. Ergebnis: Wir werden ein neues Thema beginnen (höchstwahrscheinlich im Bereich NLP).
- Recherche zu möglichen Themengebieten im Bereich NLP

Was tun wir bis nächste Woche?

Justus, Quan & Jannik:

- Zwischenbild mit zwei weiter auseinander liegenden Bildern generieren
- Weitere Möglichkeiten für ein Folgeprojekt recherchieren/überlegen

Johan:

- Weitere Recherche zu state-of-the-art Technologien für die Video-Frame-Interpolation

9. Treffen 13.12.2021

Was haben wir geschafft?

- Einteilung der Unterthemen für die Dokumentation
- Nachfolgendes Projekt: Artistic Style Transfer auf Texten

Was tun wir bis nächste Woche?

- Alle: Schreiben der zugewiesenen Unterthemen

10. Treffen 03.01.2022

Was haben wir geschafft?

- Dokumentation zum Thema Video-Interpolation/CVAE abgeschlossen
- Recherche zu Transformer-Arbeitsweise
- Recherche zu Datensätzen der NLP
- Umdenken der Projektidee: Textstil-Transfer wahrscheinlich zu komplex (und kaum Daten zu finden)
 - Daher: Neues Projektziel: Zusammenfassen von Texten

Was tun wir bis nächste Woche?

- Recherche zum Training von Word-Embeddings zur Verwendung in Transformer-Modellen (Johan & Jannik)
 - Spezielles Training notwendig oder vortrainierte Modelle?
- Aufbauen/Trainieren eines Word-Embedding-Modells (Johan & Jannik)
- Einlesen/Aufbereiten der Text-Datensätze zur weiteren Verarbeitung (Justus & Quan)

11. Treffen 10.01.22

Was haben wir geschafft?

- Angefangen mit Nachbau der CBOW-Architektur
- parallel dazu recherche zu passendem Transformer-Modell für die Zusammenfassungen-Aufgabe.

Was tun wir bis nächste Woche?

- CBOW-Architektur Input dynamisch einlesen und trainieren
- Weitere Recherche zum Thema Textzusammenfassung im Hinblick auf Transformer.