

Maschinelles Sehen

Autoencoder (8 Points)

SoSe 2025

Ziel dieser Übung ist es, Dimensionalitätsreduktionsmethoden wie die Hauptkomponentenanalyse (PCA) als lineare Einbettung mit Autoencoder-Neuronalen Netzen zu vergleichen.

Aufgabe 1: PCA auf Gesichtsbildern (4 Points)

Für diese Aufgabe müssen Sie die Dimensionalitätsreduktion auf einem gegebenen Gesichtsdatensatz implementieren (siehe Principal Components Analysis (PCA), wie in der Vorlesung besprochen). Die Datei *pca.py* enthält weitere Details, um sie selbst zu implementieren (ohne Scikit-Lernen). Abbildung 1 zeigt ein Beispiel für eine Rekonstruktion aus einer kodierten Basis mit einer unterschiedlichen Anzahl von verwendeten Dimensionen. Der mittlere Rekonstruktionsfehler für die Testbilder sollte etwa (a) bei 10 Dimensionen bei ~ 16.18 , (b) bei 75 Dimensionen bei ~ 10.48 und (c) bei 150 Dimensionen bei ~ 9.36 liegen.

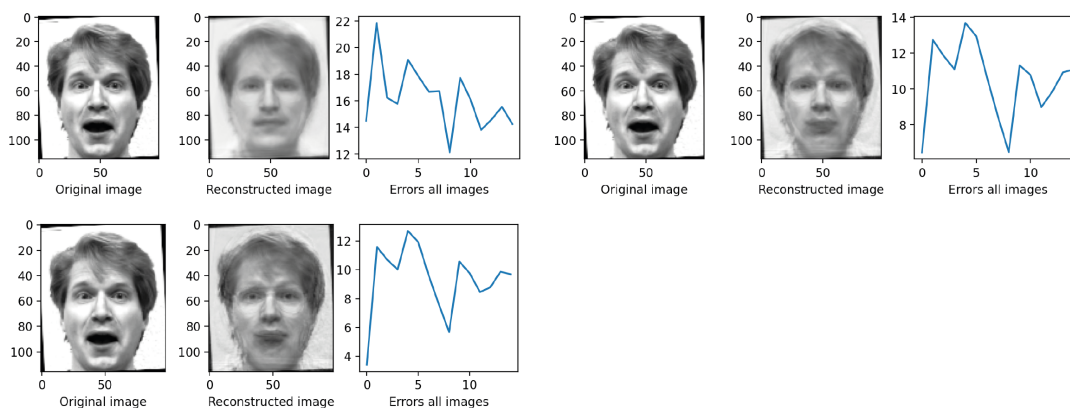


Abbildung 1: Beispiel eines originalen und rekonstruierten Bildes der Testdaten unter Verwendung einer Standard-PCA. Die Kurve zeigt den Rekonstruktionsfehler für alle 10 Testbilder. (a) verwendet 10 Dimensionen für die Rekonstruktion (b) verwendet 75 Dimensionen für die Rekonstruktion.

Aufgabe 2: Autoencoder (4 Points)

Die Datei *auto-encode.py* enthält Anweisungen für eine Autoencoder-Implementierung unter Verwendung eines sehr einfachen Fully-Connected Neuronalen Netzes. Bitte implementieren Sie dazu das Modell in PyTorch. Probieren Sie hier bitte unterschiedliche Varianten von Aktivierungsfunktionen, z.B. Sigmoid, Tanh, ReLU oder lassen Sie z.B. die Aktivierungsfunktion weg. Sie können auch die Anzahl der Schichten und Neuronen in den Schichten variieren. **Anmerkung:** Beide Methoden sollten ungefähr die gleichen Rekonstruktionsergebnisse erzielen. Abbildung 2 zeigt die Ergebnisse des Ansatzes - ein trainiertes neuronales Netzwerk, das nur eine lineare Abbildung verwendet und als Autoencoder-Netzwerk mit einem Encoder- und Decoderschritt arbeitet.

Abgabe: Bitte reichen Sie die resultierenden Plots und Ihren mittleren Rekonstruktionsfehler über alle Testbilder in einer Textdatei ein.



Abbildung 2: Beispiel eines originalen und rekonstruierten Bildes der Testdaten unter Verwendung eines einfachen linearen Autoencoders. Das Diagramm zeigt den Rekonstruktionsfehler für alle 10 Testbilder. (a) 10 Ausgabewerte im Encoder werden für die Rekonstruktion verwendet (mittlerer Fehler $\sim 21,04$) (b) 75 Ausgabewerte werden für die Rekonstruktion verwendet (mittlerer Fehler $\sim 14,87$) und bei (c) 150 Dimensionen bei $\sim 13,25$ liegen.

Die Abgabe soll via Moodle zum angegebenen Zeitpunkt erfolgen. Geben Sie bitte jeweils nur eine einzige .zip-Datei mit den Quellen Ihrer Lösung ab. Verspätete Abgaben werden mit einem Abschlag von 3 Punkten je angefangener Woche Verspätung belegt. **Bitte fügen Sie alle notwendigen Bilder ein, sodass jede Aufgabe direkt ausführbar ist. Die Bilddatenbank brauchen Sie nicht mit abgeben. Fügen Sie zusätzlich .png-Dateien für ihre erzeugten Ergebnisplots bei.** Verspätete Abgaben werden mit einem Abschlag von 3 Punkten je angefangener Woche Verspätung belegt.