

Understanding Variational Autoencoders' Latent Representations of Remote Sensing Images

Hannes Stärk

Visual Computing, Bundeswehr University Munich

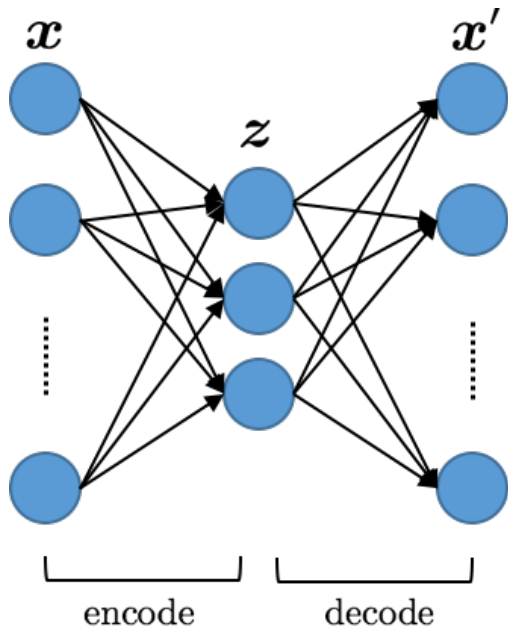
1. Gliederung

- ▶ Motivation
- ▶ Vorwissen
- ▶ Implementierung, Hardware, Software
- ▶ Architektur
- ▶ Latenter Raum
- ▶ Fazit, Wie kann es weiter gehen

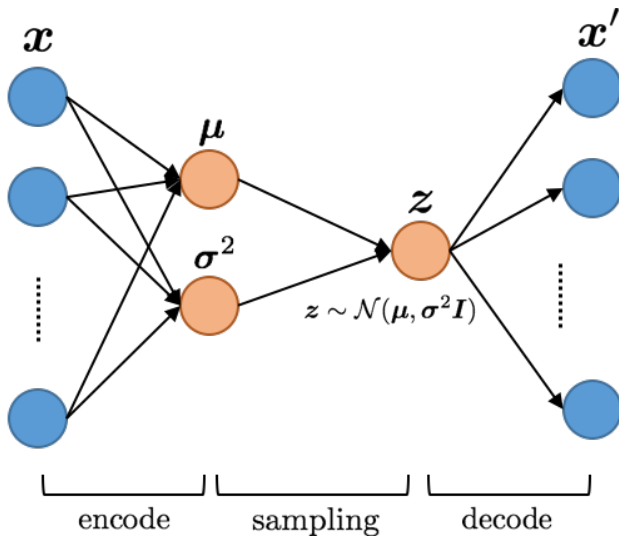
2. Motivation

- ▶ Trial-and-Error Multi-Task Architekturen
- ▶ Multi-Task Taxonomie
- ▶ Latente Informationen einzelner Schichten in Single-Task Modellen
- ▶ Latenten Raum eines Variational Autoencoders verstehen

3. Autoencoder



4. Variational Autoencoder



Credit: *GRID INC*

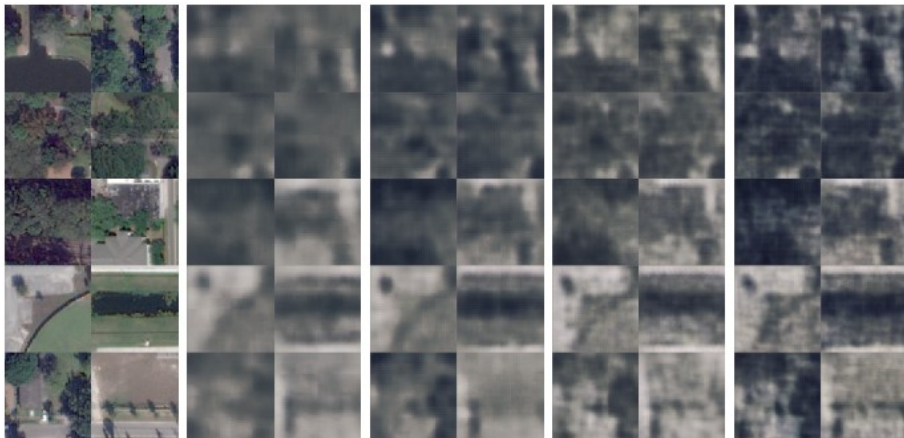
5. Implementierung, Hardware, Software

- ▶ Python
- ▶ Tensorflow
- ▶ Container der Uni Hannover
- ▶ Machine-Learning Rechner der UniBw

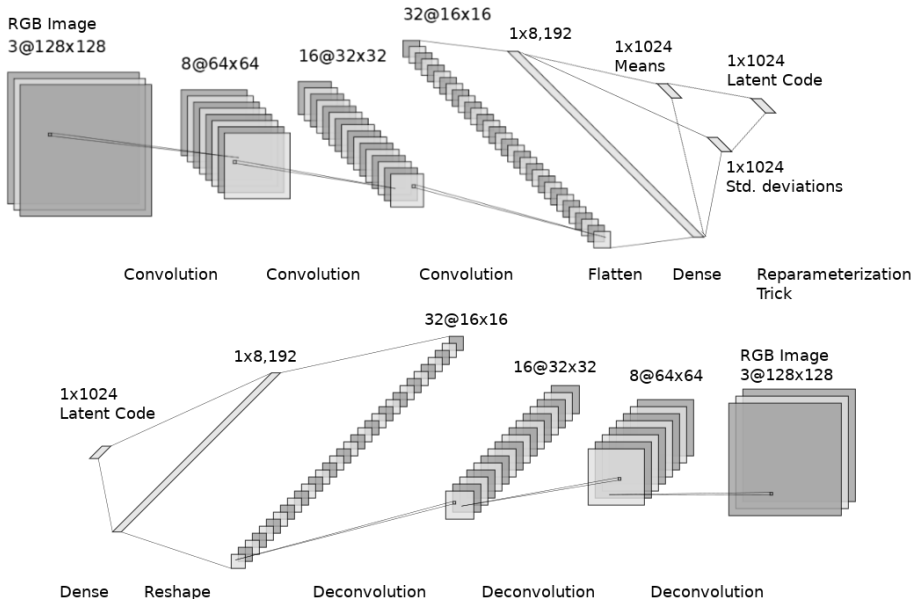
6. Architektur Experimente

- ▶ Anzahl von convolutional Schichten
- ▶ Anzahl von Filtern
- ▶ Kernel Größe
- ▶ Max/Average Pooling

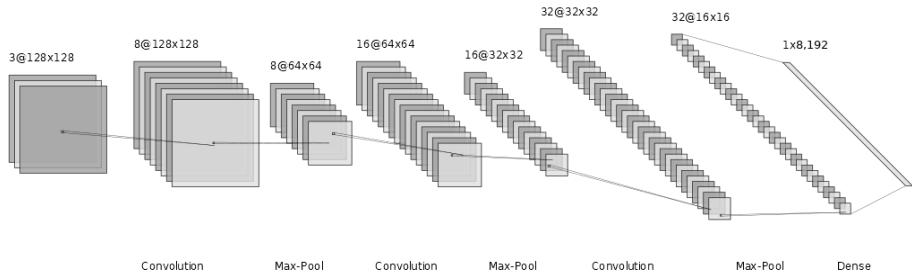
7. Ergebnisse: Schichtanzahl



8. Beste Architektur



9. Pooling Architektur



10. Training mit 178.000 Bildern



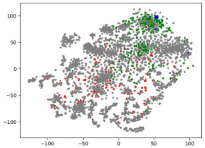
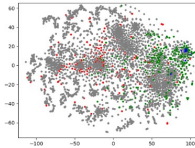
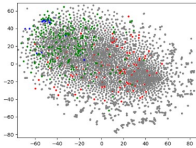
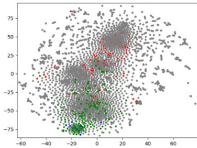
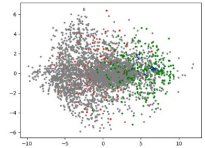
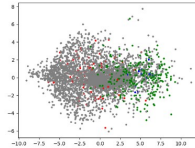
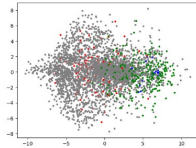
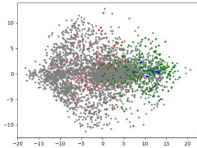
11. Latenter Raum Experimente

- ▶ 4000 Bilder => 4000 Encodings
- ▶ Dimension reduzieren
- ▶ Visualisieren

12. t-Stochastic-Neighbor-Embedding

- ▶ Machine-Learning Verfahren zur Dimensions Reduktion
- ▶ Besonders gut geeignet für einzelne Visualisierungen
- ▶ Fokus auf Kontext von Punkten zu ihren Nachbarn

13. Ergebnisse: Topographische Klassen



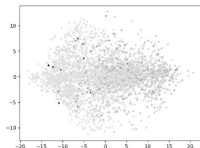
(a) Code size: 1,024

(b) Code size: 512

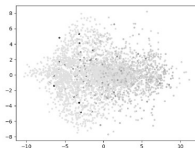
(c) Code size: 128

(d) Code size: 50

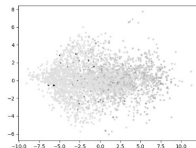
14. Ergebnisse: Höhen Daten



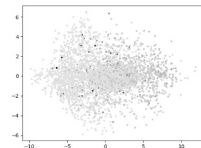
(a) Code size: 1,024



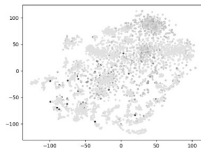
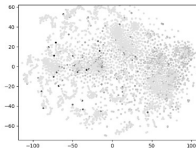
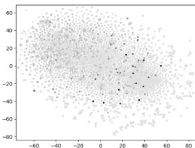
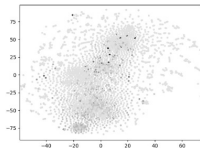
(b) Code size: 512



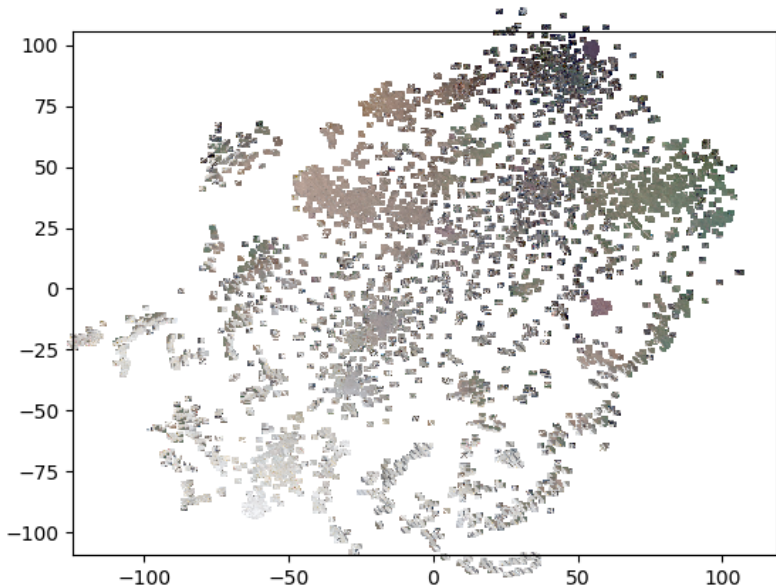
(c) Code size: 128



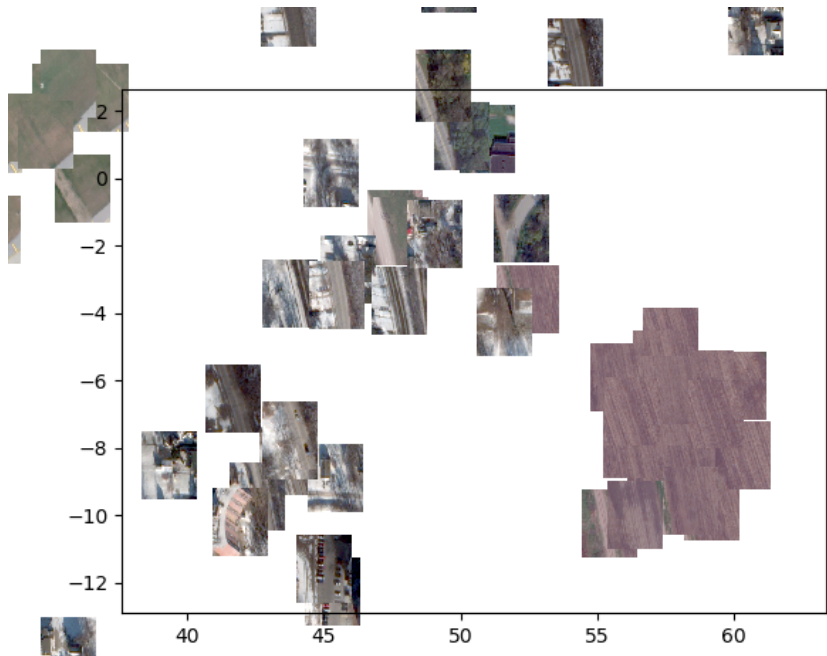
(d) Code size: 50



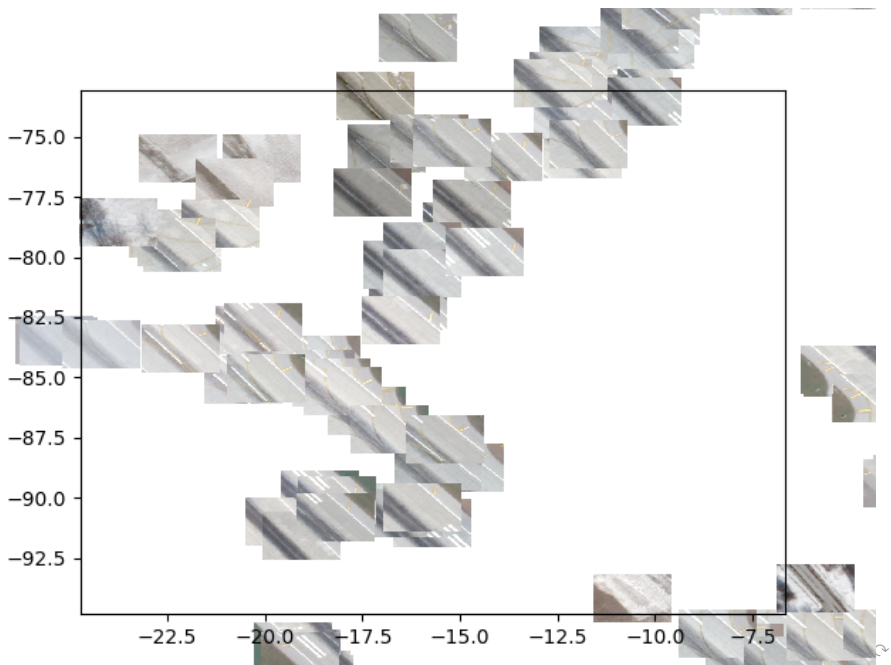
15. Ergebnisse: Eingabebilder



16. Richtungscluster



17. Straßencuster



18. Fazit und wie es weiter gehen kann

- ▶ Der VAE lernt nach komplizierten Features zu clustern
- ▶ Mit t-SNE kann man die Cluster gut visualisieren
- ▶ Gute Rekonstruktionen vs. Guter Latenter Raum?
- ▶ Präzisere Methoden um Cluster zu Features zuzuordnen
- ▶ Als nächstes bei anderen Tasks als bei Autoencodern testen

19. Rekonstruktionen und Enkodierungsgröße



(a) 1024

(b) 512

(c) 128

(d) 50