

DATA AUGMENTATION MIT VARIATIONAL AUTOENCODERN

Henri Iser

28. April 2021

Rheinische Friedrich-Wilhelms-Universität Bonn

1. Motivation
2. Variational Autoencoder
3. Datensätze
4. Generieren neuer Beispiele
5. Evaluation
6. Fazit

MOTIVATION

- Neuronale Netzwerke lösen komplexe Aufgaben
- Anforderung der großen Datenmengen oft nicht erfüllbar (*Few-Shot Learning*)
⇒ Data-Augmentation
- Generative Modelle effizient
- **Hier** Variational Autoencoder als Data Augmentation Ansatz

VARIATIONAL AUTOENCODER

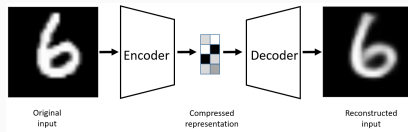


Abbildung: Encoder-Decoder Architektur¹

- Merkmalsvektoren stellen Repräsentation der Eingabe dar
- Merkmalsraum hat keine geometrische Struktur

¹ entnommen aus Lopez Pinaya u. a., 2019

VARIATIONAL AUTOENCODER

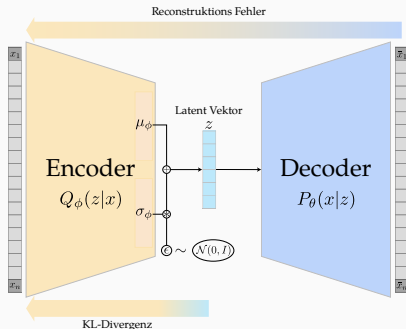


Abbildung: VAE Architektur

Rekonstruktions Fehler

$$\mathcal{L}_{NLL} = -\mathbb{E}_{z \sim Q_\phi(z|x)} [\log P_\theta(x|z)]$$

KL-Divergenz

$$\mathcal{L}_{KL} = \mathcal{D}_{KL} [Q_\phi(x|z) || \mathcal{N}(0, I)]$$

$$\mathcal{L}_{VAE} = \underbrace{\mathbb{E}_z [\log P_\theta(x|z)]}_{\text{Rekonstruktionsfehler}} - \beta \cdot \underbrace{\mathbb{E}_z [\mathcal{D}_{KL} [Q_\phi(z|x) \parallel \mathcal{N}(0, I)]]}_{\text{KL-Divergenz}} \quad (1)$$

außerdem

$$\beta_{norm} = \beta \cdot \frac{d}{N}, \quad (2)$$

mit Latent-Space Dimension d

¹ vorgeschlagen von Higgins u. a., 2017

- Merkmalsraum \rightarrow Latent-Space
- Zwischenräume stellen Interpolationen dar
- Merkmale probabilistisch abgebildet

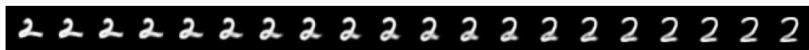


Abbildung: Interpolation innerhalb der Klasse 2

DATENSÄTZE

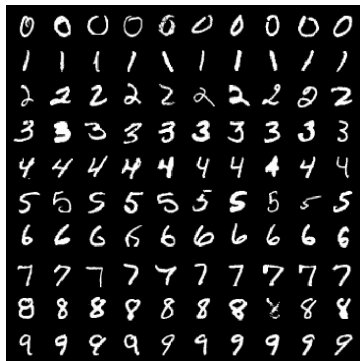


Abbildung: MNIST Datensatz



Abbildung: CelebA Datensatz

| Datensatz | # Attribute | kontinuierlich | # Klassen | # Beispiele | Balancing |
|-------------|-------------|----------------|-----------|-------------|-----------|
| card | 15 | 0.40 | 2 | 690 | 0.99 |
| diabetes | 8 | 1.00 | 2 | 768 | 0.93 |
| geneN | 60 | 0.00 | 3 | 3175 | 0.93 |
| glass | 9 | 1.00 | 6 | 214 | 0.84 |
| horse-colic | 20 | 0.70 | 3 | 364 | 0.84 |
| thyroid | 21 | 0.29 | 3 | 7200 | 0.28 |

Tabelle: PROBEN1 Datensatz Sammlung

- numerische Attributsdaten

GENERIEREN NEUER BEISPIELE

SAMPLING IM LATENT-SPACE

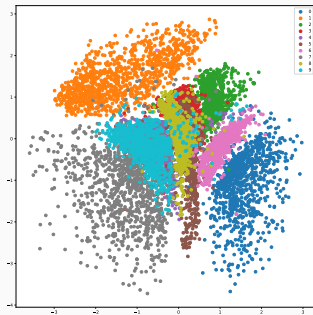


Abbildung: Latent-Space
MNIST

Generieren neuer Beispiele

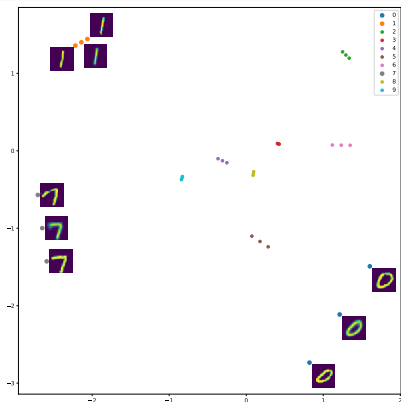


Abbildung: Interpolation im
Latent-Space

Henri Iser

13

- Sampling aus Normalverteilung

$$\hat{z}_i \sim \mathcal{N}(0, \alpha)$$

- Addition von Rauschen (Noise)

$$\hat{z}_i = z_i + \epsilon, \epsilon \sim \mathcal{N}(0, \alpha)$$

- Interpolation / Extrapolation zu k-Nächsten Nachbarn

$$\hat{z}_i = \pm \alpha \cdot [\mu_\phi(x_k) - \mu_\phi(x_i)] + \mu_\phi(x_i)$$

- Sampling aus Normalverteilung ← **Wie werden Labels zugeordnet?**

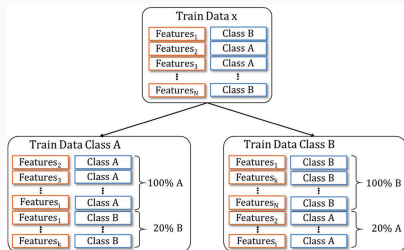
$$\hat{z}_i \sim \mathcal{N}(0, \alpha)$$

- Addition von Rauschen (Noise)

$$\hat{z}_i = z_i + \epsilon, \epsilon \sim \mathcal{N}(0, \alpha)$$

- Interpolation / Extrapolation zu k-Nächsten Nachbarn

$$\hat{z}_i = \pm \alpha \cdot [\mu_\phi(x_k) - \mu_\phi(x_i)] + \mu_\phi(x_i)$$



- Trainiere einen VAE je Klasse
- Aufteilung des Datensatzes
- (optional) Mix-In der anderen Klassen zu 20%

Abbildung: Aufteilung des Datensatzes¹

¹ entnommen aus Moreno-Barea, Jerez und Franco, 2020

Single-VAE

- selbst-überwacht Trainierbar
⇒ Nutzen großer Datenmengen
- Sampling hängt von Originalbeispielen ab

Multi-VAE

- Label Information benötigt
- Zufälliges Sampling möglich
- Erfasst Klassenmerkmale besser

GENERATIVE CLASSIFIER (GC)

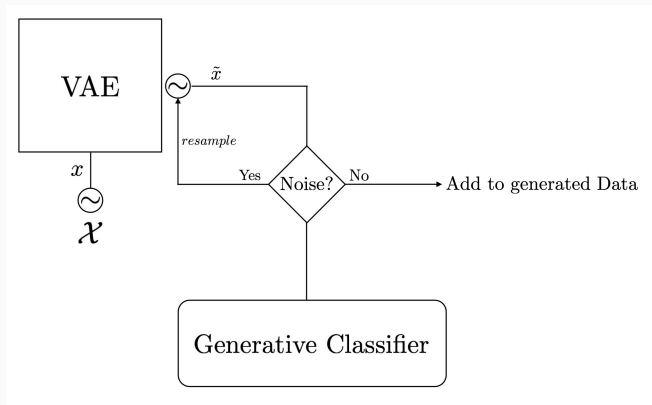


Abbildung: Konzept: Generative Classifier

- Erzeugen verrauschter Originalbeispiele ⇒ Für numerische Daten
- Verwenden eines frühen VAE Modells, um "schlechte" Rekonstruktionen zu simulieren ⇒ Für Bilddaten

⇒ Zusätzliche Hyperparameter

EVALUATION

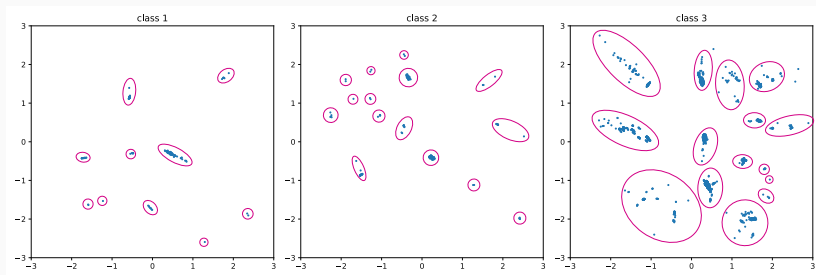


Abbildung: Clusterbildung bei diskreten Attributen (Datensatz *thyroid*)

PROBEN1 PERFORMANZ

| Datensatz | β | Baseline | VAE | VAE + GC |
|----------------|-------------------|----------|--------------|--------------|
| card | norm=0.133 | 0.871 | 0.873 | 0.869 |
| balance: Ja | 0.5 | 0.871 | 0.868 | 0.864 |
| Mix data: Nein | 1.0 | 0.871 | 0.876 | 0.871 |
| diabetes | norm=0.25 | 0.823 | 0.831 | 0.826 |
| balance: Ja | 0.5 | 0.823 | 0.830 | 0.822 |
| Mix data: Nein | 1.0 | 0.823 | 0.827 | 0.833 |
| geneN | norm=0.033 | 0.813 | 0.802 | 0.814 |
| balance: Nein | 0.5 | 0.813 | 0.818 | 0.813 |
| Mix data: Ja | 1.0 | 0.813 | 0.812 | 0.813 |
| glass | norm=0.222 | 0.985 | 0.977 | 0.985 |
| balance: Ja | 0.5 | 0.985 | 0.977 | 0.985 |
| Mix data: Nein | 1.0 | 0.985 | 0.984 | 0.985 |
| horse | norm=0.1 | 0.828 | 0.837 | 0.844 |
| balance: Ja | 0.5 | 0.828 | 0.833 | 0.835 |
| Mix data: Nein | 1.0 | 0.828 | 0.836 | 0.828 |
| thyroid | norm=0.095 | 0.954 | 0.952 | 0.956 |
| balance: Nein | 0.5 | 0.954 | 0.927 | 0.956 |
| Mix data: Ja | 1.0 | 0.954 | 0.925 | 0.954 |

- Sampling aus $\mathcal{N}(0, I)$
- ~ 6000 Schritte (VAE)
- Latent-Space Dimension
 $d = 2$

Tabelle: PROBEN1 Weighted-F1-Score

PROBEN1 FEW-SHOT SZENARIO

| # Beispiele pro Klasse | # Generierte Beispiele | Baseline | VAE | VAE + GC |
|------------------------|------------------------|--------------|--------------|--------------|
| 5 | 5 | 0.619 | 0.693 | 0.675 |
| 10 | 10 | 0.727 | 0.778 | 0.761 |
| 20 | 20 | 0.735 | 0.753 | 0.778 |
| 30 | 30 | 0.772 | 0.786 | 0.796 |
| 50 | 50 | 0.818 | 0.815 | 0.808 |
| 100 | 100 | 0.935 | 0.927 | 0.922 |

Tabelle: *thyroid* Weighted-F1-Score

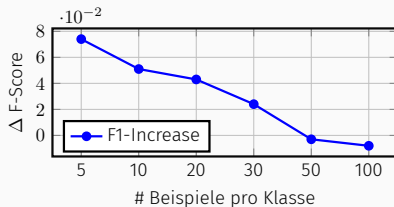


Abbildung: F1 Verbesserung

- Reduzierte Datensatzgröße
- Verbesserung nimmt mit mehr Daten ab

MNIST FEW-SHOT SZENARIO

Latent-Space Dimension $d = 50$, $\beta = 0.5$, keine β -Normalisierung

| # Beispiele pro Klasse | Baseline | Noise | Interpolation | Extrapolation | Interpolation + Noise | Extrapolation + Noise | Multi-VAE |
|------------------------|--------------|-------|---------------|---------------|-----------------------|-----------------------|--------------|
| 2 | 0.660 | 0.619 | 0.669 | 0.654 | 0.674 | 0.653 | 0.678 |
| 3 | 0.699 | 0.662 | 0.706 | 0.700 | 0.703 | 0.706 | 0.706 |
| 4 | 0.734 | 0.703 | 0.740 | 0.732 | 0.744 | 0.734 | 0.734 |
| 5 | 0.776 | 0.721 | 0.782 | 0.756 | 0.787 | 0.768 | 0.770 |
| 10 | 0.849 | 0.825 | 0.855 | 0.861 | 0.859 | 0.849 | 0.854 |
| 20 | 0.905 | 0.878 | 0.905 | 0.911 | 0.905 | 0.909 | 0.904 |
| 30 | 0.929 | 0.898 | 0.924 | 0.930 | 0.927 | 0.923 | 0.923 |
| 50 | 0.941 | 0.920 | 0.939 | 0.941 | 0.942 | 0.938 | 0.937 |
| 100 | 0.954 | 0.940 | 0.951 | 0.953 | 0.951 | 0.950 | 0.949 |
| 200 | 0.962 | 0.950 | 0.958 | 0.960 | 0.956 | 0.958 | 0.956 |
| 500 | 0.964 | 0.955 | 0.962 | 0.963 | 0.962 | 0.963 | 0.959 |
| 1000 | 0.965 | 0.952 | 0.963 | 0.964 | 0.964 | 0.963 | 0.960 |
| 2000 | 0.967 | 0.953 | 0.964 | 0.966 | 0.963 | 0.964 | 0.963 |

Tabelle: MNIST Few-Shot: Weighted-F1-Score

Latent-Space Dimension $d = 50$, $\beta = 0.5$, keine β -Normalisierung
Originalbeispiele = 5

| # Generierte Beispiele pro Klasse | Baseline | Noise | Interpolation + Noise | Extrapolation + Noise | Multi-VAE |
|--------------------------------------|----------|-------|--------------------------|--------------------------|--------------|
| 2 | 0.776 | 0.402 | 0.636 | 0.542 | 0.635 |
| 3 | 0.776 | 0.518 | 0.681 | 0.637 | 0.666 |
| 4 | 0.776 | 0.542 | 0.699 | 0.616 | 0.697 |
| 5 | 0.776 | 0.555 | 0.686 | 0.574 | 0.712 |
| 10 | 0.776 | 0.618 | 0.740 | 0.696 | 0.764 |
| 20 | 0.776 | 0.670 | 0.748 | 0.740 | 0.773 |
| 30 | 0.776 | 0.700 | 0.765 | 0.744 | 0.769 |
| 50 | 0.776 | 0.726 | 0.774 | 0.754 | 0.763 |
| 100 | 0.776 | 0.750 | 0.797 | 0.771 | 0.768 |
| 200 | 0.776 | 0.765 | 0.790 | 0.761 | 0.771 |
| 500 | 0.776 | 0.788 | 0.792 | 0.770 | 0.774 |
| 1000 | 0.776 | 0.775 | 0.788 | 0.773 | 0.767 |
| 2000 | 0.776 | 0.777 | 0.792 | 0.765 | 0.770 |

Tabelle: Weighted-F1-Score bei nur generierten Daten

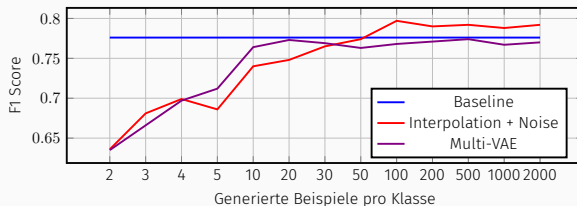


Abbildung: Weighted-F1-Score bei nur generierten Daten

MNIST FEW-SHOT: MEHR GENERIERTE DATEN

Latent-Space Dimension $d = 50$, $\beta = 0.5$, keine β -Normalisierung
Generierte Beispiele pro Klasse = 2000

| # Beispiele pro Klasse | Baseline | Noise | Interpolation + Noise | Extrapolation + Noise | Multi-VAE |
|------------------------|--------------|--------------|-----------------------|-----------------------|-----------|
| 2 | 0.660 | 0.679 | 0.668 | 0.638 | 0.634 |
| 3 | 0.699 | 0.716 | 0.715 | 0.711 | 0.687 |
| 4 | 0.734 | 0.748 | 0.747 | 0.740 | 0.724 |
| 5 | 0.776 | 0.779 | 0.795 | 0.772 | 0.768 |
| 10 | 0.849 | 0.851 | 0.868 | 0.871 | 0.839 |
| 20 | 0.905 | 0.894 | 0.907 | 0.919 | 0.897 |
| 30 | 0.929 | 0.905 | 0.922 | 0.936 | 0.916 |
| 50 | 0.941 | 0.914 | 0.935 | 0.945 | 0.925 |
| 100 | 0.954 | 0.923 | 0.946 | 0.951 | 0.936 |
| 200 | 0.962 | 0.929 | 0.952 | 0.960 | 0.934 |
| 500 | 0.964 | 0.939 | 0.957 | 0.960 | 0.924 |
| 1000 | 0.965 | 0.946 | 0.960 | 0.961 | 0.913 |
| 2000 | 0.967 | 0.953 | 0.963 | 0.964 | 0.905 |

Tabelle: MNIST Few-Shot: Weighted-F1-Score

EINFLUSS DES β -FAKTORS

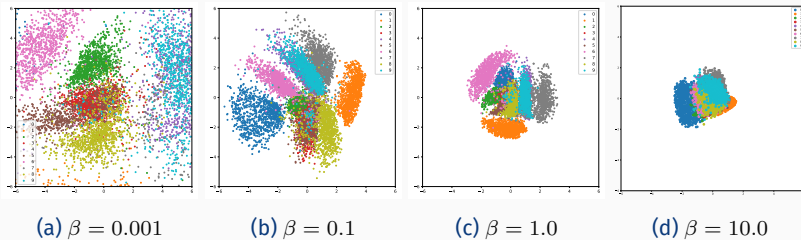
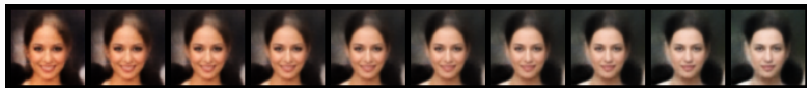


Abbildung: Latent-Space MNIST mit verschiedenen Werten für β

- Hohe β -Werte führen zu unscharfen Rekonstruktionen



(a) $\beta = 150$



(b) $\beta = 1$

Abbildung: Dimension Sweep CelebA; Trainiert mit β_{norm}

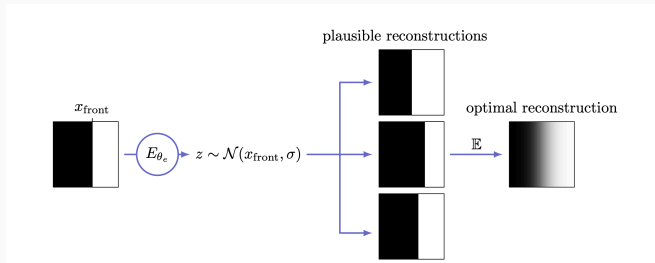


Abbildung: Unschärfe Auftreten beim VAE¹

¹ entnommen aus Plumerault, Borgne und Hudelot, 2020

FAZIT

Numerische Daten (PROBEN1):

- Auf numerischen Daten Hyperparameter sensitiv, besonders bei diskreten Attributen
- Geringe Verbesserung der Klassifikations Güte
- Im Vergleich zu Moreno-Barea, Jerez und Franco, 2020 (GAN) geringere Performanz
- Few-Shot kleine, konsistente Verbesserung

Bilddaten (MNIST, CelebA):

- Nützliche Latent-Space Struktur
- Dekorrelation unabhängiger Merkmale (β -VAE)
- Few-Shot: Konsistente Verbesserung
- Single-VAE > Multi-VAE durch selbst-überwachtes Lernen
- Starke Unschärfe Effekte



Higgins, Irina u.a. (2017). „B-Vae : Learning Basic Visual Concepts With a Constrained Variational Framework“. In: *Iclr 2017*.



Lopez Pinaya, Walter Hugo u.a. (2019). „Autoencoders“. In: *Machine Learning: Methods and Applications to Brain Disorders*, S. 193–208. doi: 10.1016/B978-0-12-815739-8.00011-0. arXiv: 2003.05991.



Moreno-Barea, Francisco J., José M. Jerez und Leonardo Franco (2020). „Improving classification accuracy using data augmentation on small data sets“. In: *Expert Systems with Applications* 161, S. 113696. ISSN: 09574174. doi: 10.1016/j.eswa.2020.113696. URL: <https://doi.org/10.1016/j.eswa.2020.113696>.



Plumerault, Antoine, Hervé Le Borgne und Céline Hudelot (2020). „AAVE: Adversarial Variational Auto Encoder“. In: *January*, S. 1–16. arXiv: 2012.11551. URL: <http://arxiv.org/abs/2012.11551>.

FRAGEN?