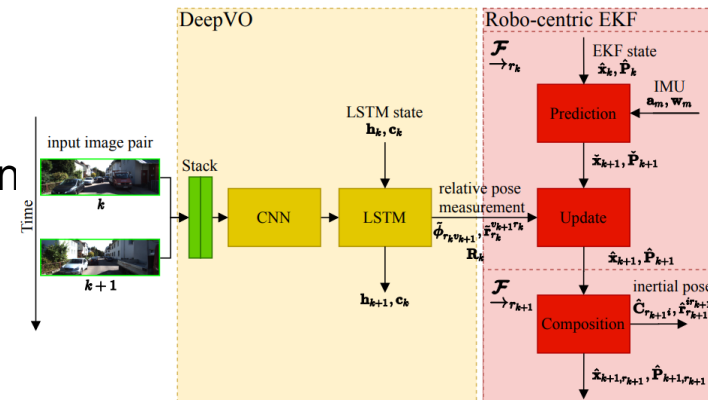


# VIO mittels robozentrischem Kalman Filter in einer Ende-zu-Ende Architektur

Visual Inertial Odometry (VIO) behandelt die Schätzung einer Roboterpose durch Kombination von visuellen Merkmalen und der Inertialodometrie des Roboters. Im Gegensatz zu klassischen Kalman Filtern ermöglicht die robozentrische Variante eine natürlichere Betrachtung von relativen Posenänderungen. Diese relativen Änderungen können im Bildmaterial mit aktuellen Deep Learning Verfahren geschätzt werden. Der Kalman Filter kann hierbei auch unmittelbar in die Netzarchitektur integriert werden, sodass ein Ende-zu-Ende Verfahren resultiert.



Chunshang Li. "Towards End-to-End Learning of Monocular Visual-Inertial Odometry with an Extended Kalman Filter"

## Aufgaben

- Herausarbeiten der Funktionsweise von robozentrischen Kalman Filtern
- Evaluierender Vergleich mit konventionellen Methoden

## Literatur

- Towards End-to-End Learning of Monocular VIO with an EKF (Li, C. 2020)
- Robocentric Visual-Inertial Odometry (Huai et. al, 2018)

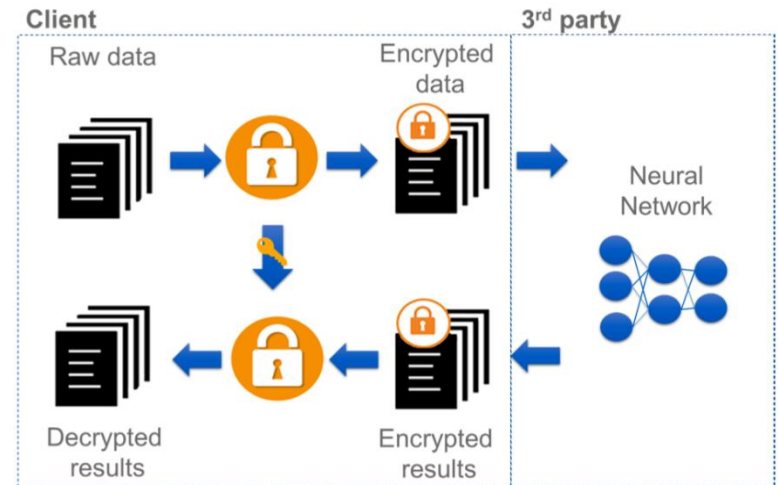
## Betreuer

Raphael Hagmanns (raphael.hagmanns@kit.edu)

# Privatsphäre wahrende medizinische Bildauswertung

Medizinische Daten, vor allem Bilddaten, werden zunehmend digitalisiert. Während früher optische Medien als Datenträger zum Austausch verwendet worden sind, geschieht dies heute bereits oft über eine direkte Vernetzung. Durch diese hohe Verfügbarkeit werden medizinische Bilddaten mehr und mehr von verschiedenen Forschenden für Deep Learning Algorithmen etc. verwendet.

Da medizinische Daten laut DSGVO ein besonders hohes Schutzniveau erfordern, ist es notwendig bei der Weitergabe von medizinischen Daten die Privatsphäre des Betroffenen zu wahren. Gleichzeitig soll der Nutzen der Daten erhalten bleiben um die Auswertungsqualität sicherzustellen.



Quelle: Vizitiu et. al: "Towards Privacy-Preserving Deep Learning based Medical Imaging Applications " *IEEE Instrumentation and Measurement Society* (2019)

## ■ Aufgaben:

- Literaturrecherche zu Grundlagen und Umsetzungen von Privatsphäre wahrnder Bildverarbeitung
- Beschreibung und Bewertung der Einsatzmöglichkeiten von Privatsphäre wahrnder Techniken für Medizinische Bildauswertung

## ■ Literatur:

- Vizitiu et. al: Towards Privacy-Preserving Deep Learning based Medical Imaging Applications
- Li et. Al: Privacy-preserving Federated Brain Tumour Segmentation

## ■ Betreuer:

- Arno Appenzeller (arno.appenzeller@kit.edu)

# Privacy Enhancing Visual Cryptography Schemes

In Videoüberwachungssystemen und biometrischen Anwendungen werden häufig Bilddaten mit Personenbezug erfasst und verarbeitet, wie etwa Gesichter oder Fingerabdrücke. Hierbei kann die informationelle Selbstbestimmung der Betroffenen verletzt werden. Mittels Visual Cryptography Schemes können Bilder in einzelne *Sheets* aufgeteilt werden, die separat betrachtet keine relevanten Informationen mehr beinhalten. Nur durch die gemeinsame Betrachtung der Sheets kann das originale Bild rekonstruiert werden. Dadurch können Datenschutzrisiken beim Verarbeiten von Bilddaten vermindert werden.

## Aufgaben

- Literaturrecherche zu Visual Cryptography Schemes und deren Einsatzmöglichkeiten
- Bewertung der Verfahren hinsichtlich Nutzen und Performance

## Betreuer

Paul Wagner (paul.wagner@kit.edu)

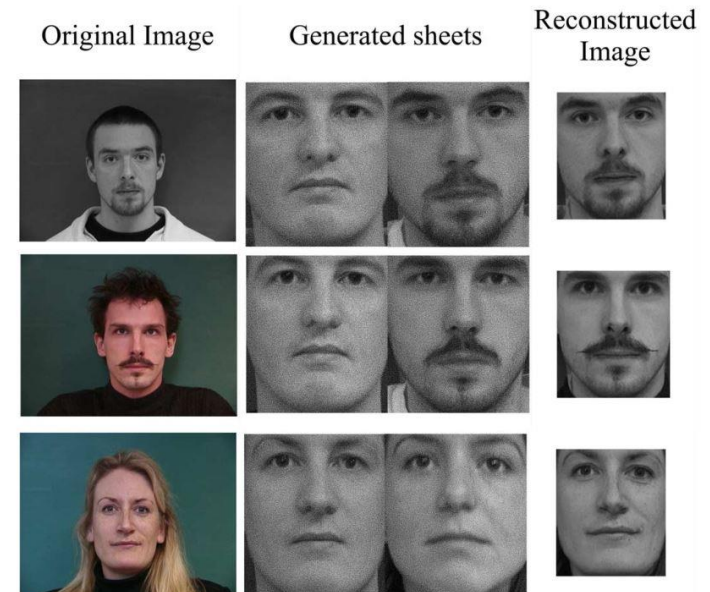


Fig. 15. Illustration of the proposed approach using images from the IMM database.

Ross, Arun, and Asem Othman. "Visual cryptography for biometric privacy." *IEEE transactions on information forensics and security* 6.1 (2010): 70-81.

# Flow-based Generative Models

The discipline of generative modeling has experienced enormous leaps in capabilities in recent years, mostly with likelihood-based methods and generative adversarial networks (GANs). Likelihood-based methods can be divided into three categories: autoregressive models, Variational autoencoders (VAEs), and Flow-based generative models. Among them, Flow-based generative models are conceptually attractive due to tractability of the exact log-likelihood, tractability of exact latent-variable inference, and parallelizability of both training and synthesis.



Source: [1]

## Aufgaben

- How do flow-based generative models work?
- What are the pros and cons of flow-based generative models?

## Literatur

- [1] Kingma, Diederik P. and Prafulla Dhariwal. “Glow: Generative Flow with Invertible 1x1 Convolutions.” NeurIPS (2018).
- [2] Dinh, Laurent et al. “NICE: Non-linear Independent Components Estimation.” CoRR abs/1410.8516 (2015).
- [3] Dinh, Laurent et al. “Density estimation using Real NVP.” ArXiv abs/1605.08803 (2017).

## Betreuer

Chengzhi Wu (chengzhi.wu@kit.edu)

# Inkrementelle Few-Shot Objektdetektion

In der Praxis werden immer wieder neue Anforderungen an Objektdetektoren gestellt. Anwender möchten neue Klassen detektieren können ohne ein aufwändiges, vollständiges Training durchzuführen. Diese Problemstellung greifen inkrementelle Verfahren auf, die aufgrund ihrer Architektur neue Klassen ohne oder mit minimalen Trainingsaufwand detektieren können.



Quelle [1]

## Aufgaben

- Literaturrecherche zu inkrementeller Few-Shot Objektdetektion
- Zusammenfassung und Vergleich von einschlägigen Verfahren

## Literatur

- [1] Juan-Manuel Perez-Rua, Xiatian Zhu, Timothy M. Hospedales, Tao Xiang. Incremental Few-Shot Object Detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

## Betreuer

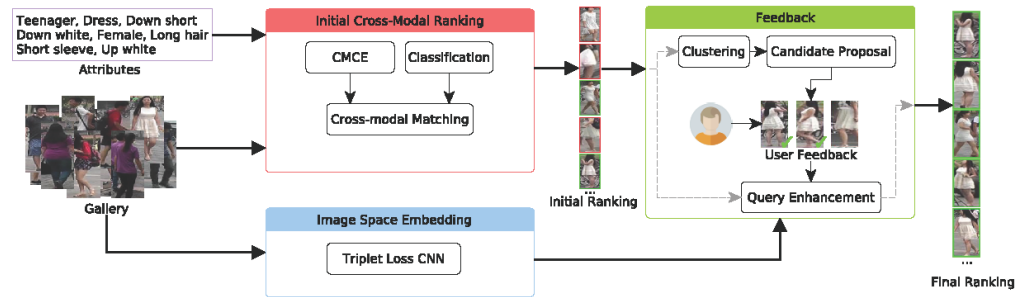
Stefan Wolf (stefan.wolf@iosb.fraunhofer.de)

# Relevance Feedback for Deep Learning-based Person Re-identification

Die Personenwiedererkennung versucht eine Person anhand eines Anfragebilds in einer großen Bilddatenbank wiederzufinden, indem eine nach Ähnlichkeit sortierte Rangliste erzeugt wird. Häufig enthält diese erste Ergebnisliste jedoch nur wenige korrekte Ergebnisse. Mithilfe von Feedback des Systemoperators kann die Ergebnisliste in einem zweiten Verarbeitungsschritt deutlich verbessert werden.

## Aufgaben

- Literaturrecherche zu Relevance Feedback
- Vergleich der Verfahren in der Literatur hinsichtlich ihrer Eignung für die Personenwiedererkennung



Quelle: [2]

## Literatur

- [1] Heng Xu, Jun-yi Wang and Lei Mao, "Relevance feedback for Content-based Image Retrieval using deep learning," 2017 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, 2017, pp. 629-633, doi: 10.1109/ICIVC.2017.7984632.
- [2] Specker, A., Schumann, A., & Beyerer, J. (2019, September). An Interactive Framework for Cross-modal Attribute-based Person Retrieval. In 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) (pp. 1-8). IEEE.

## Betreuer

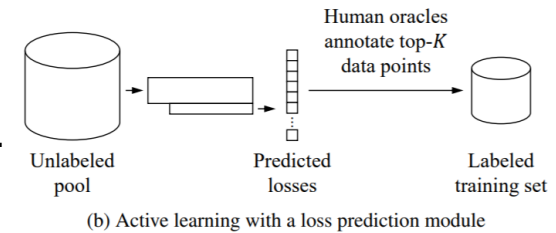
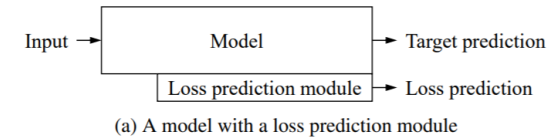
Andreas Specker (andreas.specker@iosb.fraunhofer.de)

# Active Learning for Deep Learning

Automatically detecting and understanding pedestrians in real-world uncooperative scenarios is very challenging.

A major issue is limited annotated data and annotation budget for supervised learning. In practice, large amount of data is available, however it is unrealistic to annotate all of it.

Active Learning is a building block to choose the data to be annotated.



Source [4]

## Assignment

- How do Active Learning / Uncertainty Sampling works?
- Pro and cons of the studied models for use in production / human pose estimation

## Literature

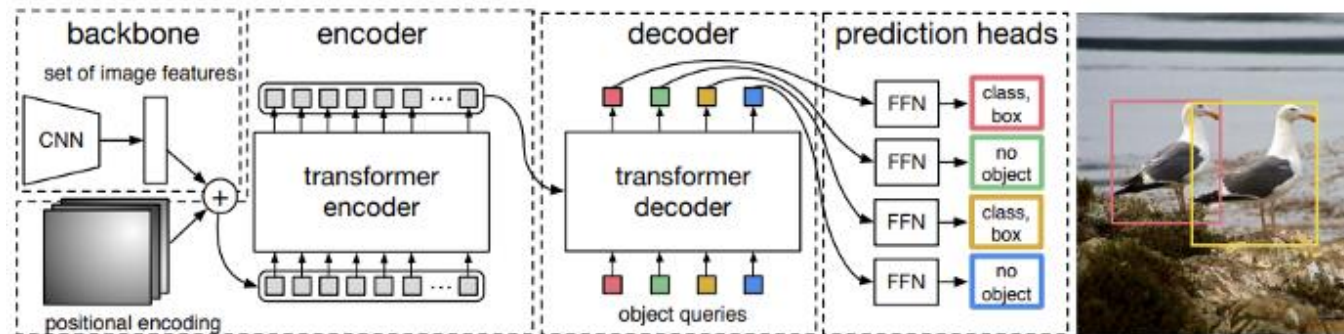
- [1] Gal, Yarin, and Zoubin Ghahramani. "Dropout as a bayesian approximation: Representing model uncertainty in deep learning." ICML 2016.
- [2] Houlsby, Neil, et al. "Bayesian active learning for classification and preference learning." arXiv preprint arXiv:1112.5745 (2011)
- [3] Gal, Yarin, Riashat Islam, and Zoubin Ghahramani. "Deep bayesian active learning with image data." ICML 2017.
- [4] Yoo, Donggeun, and In So Kweon. "Learning loss for active learning." CVPR 2019.

## Supervisor

Mickael Cormier (mickael.cormier@iosb.fraunhofer.de)

# Transformers Networks in Computer Vision

Transformer models are primarily used in the field of natural language processing (NLP) where they have shown great success. Recently the computer vision community tried to adapt these models for vision tasks. As a result, Transformer models have already been successfully used in image recognition, object detection, segmentation, image generation, video understanding and more.



## ■ Aufgaben:

- Literaturrecherche zum Einsatz von Transformer Netzwerken in Computer Vision Anwendungen
- Bewertung der Transformer-Netze hinsichtlich der Möglichkeiten und Limitierungen, die diese in Computer-Vision Anwendungen bieten

## ■ Literatur:

- Khan et. al: Transformers in Vision: A Survey. 2021
- Carion et al., End-to-End Object Detection with Transformers. 2020.

## ■ Betreuer:

- Benedikt Fischer (benedikt.fischer@kit.edu)