PROF. DR. MARTIN STEINEBACH
OPEN THESIS TOPICS

Our research group is always interested in motivated students who want to do their Bachelor's or Master's thesis. In the following we provide open thesis topics we can currently offer.

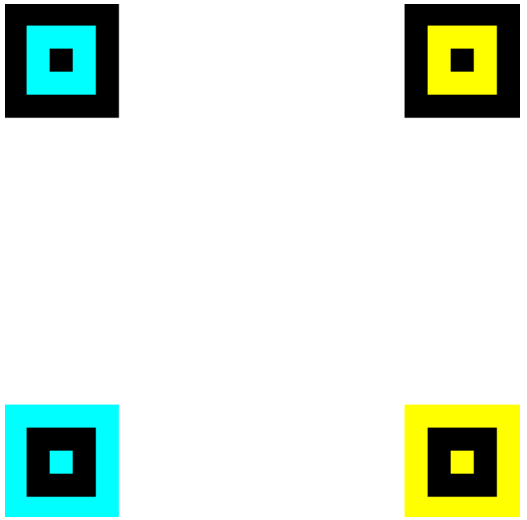
In case you are interested in these thesis subjects – or have your own thesis proposals – please don't hesitate to contact us:

E-Mail: Martin.Steinebach@sit.fraunhofer.de
WWW: https://www.sit.fraunhofer.de/de/mediasecurity/
https://www.sit.fraunhofer.de/de/itforensics/

Version August 2, 2024

OBJECT RECOGNITION THROUGH VISIBLE COLOR PATTERN

- Plug and Play solution for "discovering" offline objects in images and videos
- Develop a robust method that uses visible color patterns to make objects discoverable for digital devices (video, image).
- In the Thesis
  - n ≥ 4 different color patterns are to be created, which can be robustly recognized by neural /deep learning
  - Post-processing (dewarping, rotation, etc.) of the recognized space enclosed by the color patterns for further use (OCR or similar)
  - A good starting point is for example



Martin Steinebach / Betreuer*in: JH
Martin.Steinebach@sit.fraunhofer.de,

AUDIO DEEPFAKE GENERATION

- The quality of audio deepfakes is rapidly improving in conjunction with the advances in artificial intelligence. Audio deepfake generation methods can be classified into two main categories: voice conversion (VC) and text-to-speech synthesis (TTS).
- The goals of this thesis are:
  - To analyze state-of-the-art generation methods (VC and/or TTS) and to identify potential avenues for improvement.
  - To provide an implementation of a generation system that can enhance the quality of the generated audio recording and increase its similarity to the target speaker

- Earlier work regarding audio deepfake generation:
  - https://www.mdpi.com/2076-3417/13/5/3100
  - https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10096399

Martin Steinebach / Betreuer: KS
Martin.Steinebach@sit.fraunhofer.de,

AUDIO DEEPFAKE DETECTION

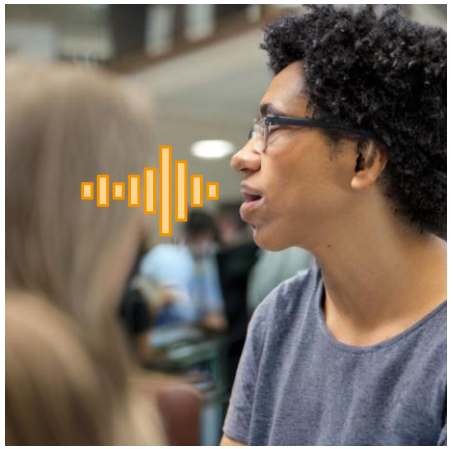
- As the amount of AI-generated audio content on social media rapidly increases, there is a growing necessity to develop detectors that can distinguish between AI-generated and genuine audio.
- The goals of this thesis are:
  - To analyze state-of-the-art detection methods and to identify potential avenues for improvement.
  - To provide an implementation of a classification system that can improve the results on e.g. the in-the-wild dataset

- Earlier work regarding audio deepfake detection:
  - https://www.mdpi.com/1999-4893/15/5/155
  - https://dl.acm.org/doi/pdf/10.1145/3658664.3659662

Martin Steinebach / Betreuer: KS
Martin.Steinebach@sit.fraunhofer.de,

ROBUSTNESS OF AUDIO DEEPFAKE DETECTION METHODS

- New synthesis methods can be used to alter the spoken information contained in audio recordings and video soundtracks
- Current state-of-the-art detection methods can achieve high accuracies but their robustness against common audio post-processing (e.g., compression, band-pass filtering), operations have not been investigated
- The goals of this thesis are:
  - To analyze the robustness of state-of-the-art detection methods on post-processed audio deepfakes
  - To provide an implementation of a classification system that can improve the robustness
- Earlier work regarding robust audio deepfake detection:
  - https://www.semanticscholar.org/reader/58cb55954dc1ccf88d6d4932067041406c18bfda



Stable Diffusion

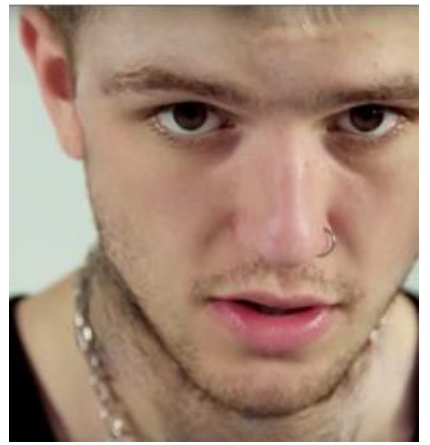
ATHENE
Nationales Forschungszentrum für angewandte Cybersicherheit

Martin Steinebach / Betreuer: KS
Martin.Steinebach@sit.fraunhofer.de,

Fraunhofer SIT

DETECTING SYNTHETIC IMAGE CONTENT CREATED BY „INPAINTING“

- Detail in digital images can be enhanced or created using ML-based image synthesis methods in terms of „inpainting“. Such enhancements can be applied for malicious purposes such as forging digital evidence or distributing fake news
- Current methods for identifying inpainting feature varying detection performance with respect to different image inpainting algorithms
- The goals of this thesis are:
  - To analyze the shortcomings of current detection algorithms for “inpainting”
  - To improve the forgery detection performance by exploiting common characteristics of a selection of synthesis algorithms
- Earlier work regarding the detection of splicing boundaries:
  - https://openaccess.thecvf.com/content_CVPR_2020/papers/Li_Face_X-Ray_for_More_General_Face_Forgery_Detection_CVPR_2020_paper.pdf



https://www.reddit.com/r/LilPeep/comments/j4c1uz/tattooless_peep_using_the_nvidia_image_inpainting/

ATHENE
Nationales Forschungszentrum für angewandte Cybersicherheit

Martin Steinebach / Betreuer: RF, SZ
Martin.Steinebach@sit.fraunhofer.de,

Fraunhofer SIT

DETECTING SYNTHETIC IMAGES CREATED BY „FULL IMAGE SYNTHESIS“

- Digital images can be created using ML-based image synthesis. In contrast to „inpainting“ *full* images can be synthesized from scratch in a photorealistic appearance. Such enhancements can be applied for malicious purposes such as making up digital evidence or fake news
- Current methods for “full image synthesis” identification feature varying detection performance with respect to different image synthesis algorithms
- The goals of this thesis are:
  - To analyze the shortcomings of current detection algorithms for “full image synthesis”
  - To implement a detection method exploiting characteristics of a selection of synthesis algorithms to improve the forgery detection performance
- Earlier works regarding the annotation of synthetically generated images:
  - https://arxiv.org/pdf/2211.00680v1.pdf

openai.com

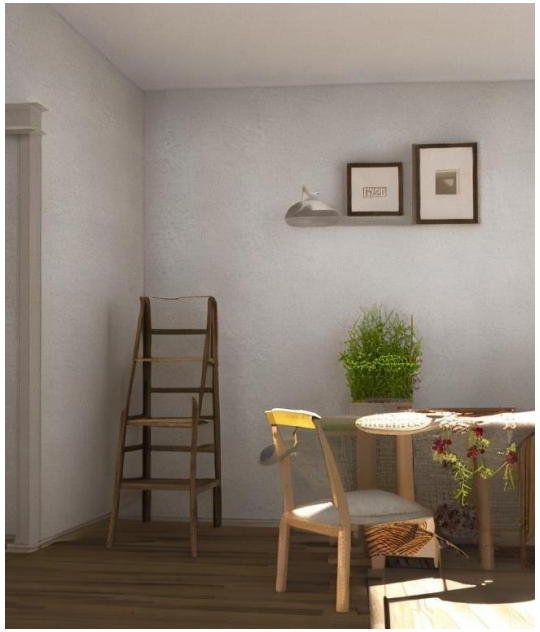
ATHENE
Nationales Forschungszentrum für angewandte Cybersicherheit

Martin Steinebach / Betreuer: RF, SZ
Martin.Steinebach@sit.fraunhofer.de,

Fraunhofer SIT

RECOGNIZING ROOMS/LOCATIONS INSIDE BUILDINGS BASED ON REFERENCE DATA

- There exist several solutions for recognizing objects in images and videos as well as techniques to match visual data. For some use-cases training an ML-based solution is challenging as training data is scarce.
- Training an ML-based classifier on 3D synthesized data to estimate the facial landmarks of human faces has shown to outperform state-of-the-art methods
- The goals of this thesis are:
  - To analyze whether synthesizing training data in 3D space can improve the performance of methods trying to match the interior of a room
  - To evaluate the transferability on real data
- Requirements: Knowledge in 3D modelling and/or game engines
- Earlier works regarding the transfer of synthesized image data:
  - https://openaccess.thecvf.com/content/ICCV2021/papers/Wood_Fake_It_Till_You_Make_It_Face_Analysis_in_the_ICCV_2021_paper.pdf



Stable Diffusion

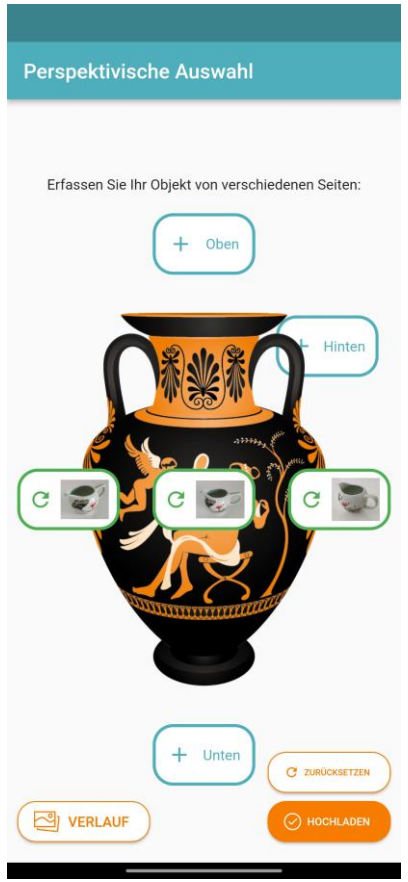
ATHENE
Nationales Forschungszentrum für angewandte Cybersicherheit

Martin Steinebach / Betreuer: RF, SZ
Martin.Steinebach@sit.fraunhofer.de,

Fraunhofer SIT

RECOGNITION OF UNKNOWN CULTURAL ASSETS

- Automatic identification of unknown cultural goods, like archaeological objects, helps prevent illicit trafficking of cultural assets in trade and at customs. The main goal is to determine the provenance of unknown antiquities, including the origin and age of objects by recognizing their appearance.
- Proper deep neural networks are required to train a model to recognize different antiquities. The available datasets for training are limited and numerous objects are not fully tagged, because labeling antiquities poses a high resource demand for archaeological experts.
- The goal of this thesis is to
  - develop or adapt a deep neural network model to identify the origin and age of cultural objects
  - implement the recognition model
  - evaluate the developed model with available datasets
- Related work:
  - https://library.imaging.org/ei/articles/34/8/IMAGE-273



ATHENE
Nationales Forschungszentrum für angewandte Cybersicherheit

Martin Steinebach / Betreuer: HL
Martin.Steinebach@sit.fraunhofer.de,

Fraunhofer SIT

IMAGE MANIPULATION DETECTION

- Digital images are subject to manipulation. Powerful image editing tools make it possible to manipulate images without specialized skills. Therefore, image forensics is needed to verify the integrity and authenticity of digital images.
- Different types of manipulations leave different traces, which can not be easily detected by a single method. Moreover, such traces can be weakened or even eliminated by post-processing.
- The goal of this thesis is to
  - develop or improve a deep learning based or classical forensic algorithm for specific image manipulations
  - implement the forensic algorithm
  - evaluate the implemented algorithm using different datasets
- Related Work:
  - https://www.sciencedirect.com/science/article/pii/S1051200417301938

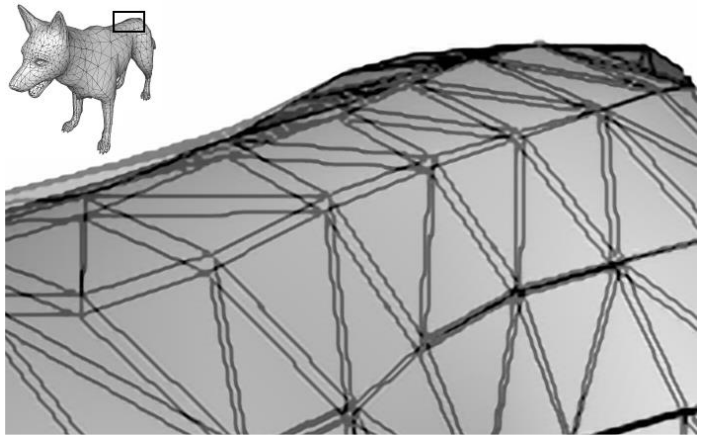
ATHENE
Nationales Forschungszentrum für angewandte Cybersicherheit

Martin Steinebach / Betreuer: HL
Martin.Steinebach@sit.fraunhofer.de,

Fraunhofer SIT

3D-MODEL WATERMARK

- Decentralized production becomes more important with 3D printers
- 3D models are sent to the printers and printed/manufactured on-site
- Sharing the print data, or scanning a 3D model and duplicating it, is possible and that is where this process fails to progress
- The goal of this thesis
  - Is to design a digital watermarking technique for 3D models with the following properties
    - Written for G Code
    - Watermark extractable after 3D printing and scanning
    - Watermark is imperceptible to a person
  - Implement the designed watermark and evaluate it
  - A good starting point is e.g. https://arxiv.org/abs/2109.07202



ATHENE
Nationales Forschungszentrum für angewandte Cybersicherheit

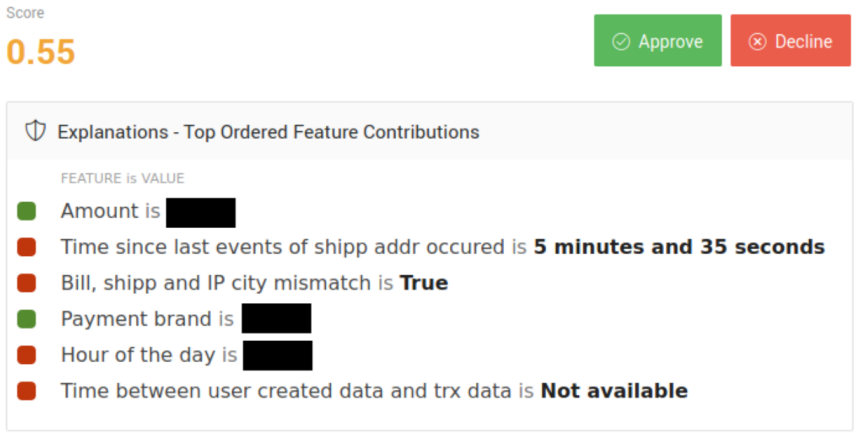
Martin Steinebach / Betreuer*in: WB
Martin.Steinebach@sit.fraunhofer.de,

Fraunhofer SIT

EXPLAINABLE AI IN FINANCE

- Machine learning methods are already used in numerous applications today. In many cases, however, it is unclear how such a model arrived at its decision or which factors favored a particular decision. Particularly in sensitive areas such as finance and healthcare, the use of machine learning methods requires the algorithms employed to be traceable.

- The scope of this work is negotiable, possible approaches are e.g.:
  - Identify current state-of-the-art approaches and describe how decision-making is made comprehensible or explainable in **black-**, **white-** and **grey-box** models applied to tabular data.
  - Development of a system whose decisions are understandable from start to finish - based on OSINT, network analysis, use of inherently understandable algorithms, etc.



Source: How can I choose an explainer? An Application-grounded Evaluation of Post-hoc Explanations, 2021, https://arxiv.org/abs/2101.08758

- A good starting point for this topic is, for example, https://ieeexplore.ieee.org/document/9446887 and https://arxiv.org/abs/2101.08758

ATHENE
Nationales Forschungszentrum für angewandte Cybersicherheit

Martin Steinebach / Betreuerin: VB
Martin.Steinebach@sit.fraunhofer.de,

Fraunhofer SIT

SCALABLE EFFICIENT GRAPH ANALYSIS

- Network graph analysis is a powerful tool, but the calculation time often scales poorly with large amounts of data.

- The aim of this work is to research and develop approaches to apply graph analysis methods and algorithms to large amounts of financial transaction data, preferably in real time.

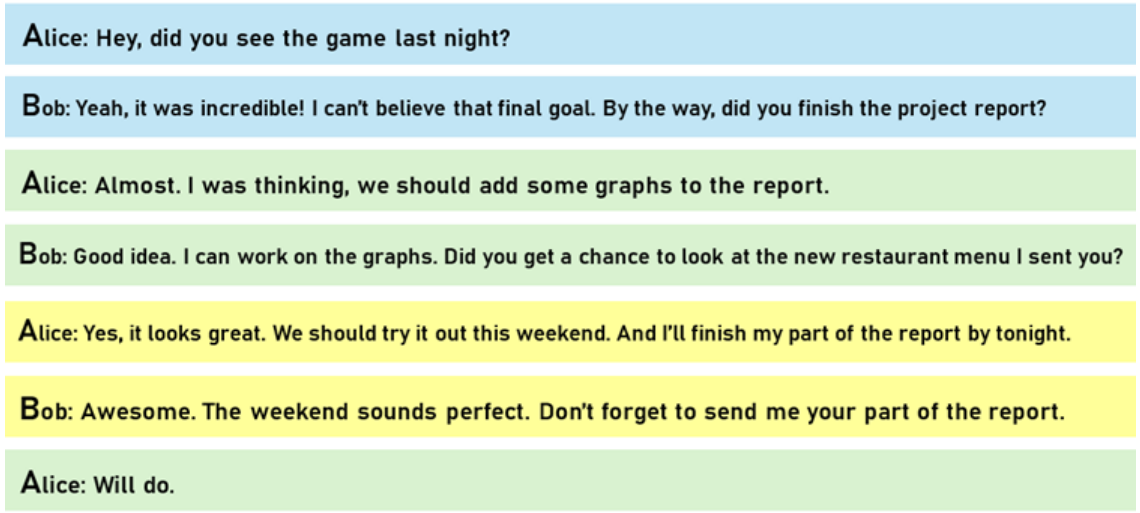
Martin Steinebach / Betreuerin: VB
Martin.Steinebach@sit.fraunhofer.de,

TOPIC MODELLING IN CHATS

- The analysis of topics within chat conversations facilitates a comprehensive understanding of the main discussion points and trends observed in communication. Topic modelling techniques can be used to improve customer service or detect criminal activity in chats.
- The characteristics of the language used in chats and the dynamic nature of the development of topics in chat conversations present a great challenge to topic identification in chats.
  The ability to accurately identify the main topic of a chat conversation depends on several factors, including the context awareness of the algorithm.
- Goal of the thesis:
  - Investigate the state of the art in topic modelling techniques
  - Discuss the robustness of their implementation in chat analysis
  - Develop a concept for topic analysis in chats
  - Implement and evaluate the proposed method
  - Relevant dataset: NPS Chat Corpus

THREAD DETECTION IN CHAT STREAMS

- Unlike written discourse such as reviews, comments or news articles, chat streams typically lack a linear discussion of a single topic. Instead, they comprise a multitude of partially threaded and interrelated topics presented in the form of shorter, incomplete messages that do not always follow a clear narrative thread.
- To gain a deeper comprehension of contextual relatedness among chat messages, the development of algorithms capable of tracking the dynamic nature of chat streams is essential.
- Goal of the thesis:
  - Research the topic and existing techniques
  - Develop a concept for identifying threads in chat streams
  - Implement and evaluate the proposed method
  - Relevant dataset: NPS Chat Corpus



Source: ChatGPT

Martin Steinebach / Betreuerin: SF
Martin.Steinebach@sit.fraunhofer.de,

AUTHOR PROFILING IN CHAT LOGS

- Author profiling is the task of identifying demographic characteristics (e.g. gender, age, geographic origin, level of education, native language) or psychometric traits of the author of a text. It has a wide range of applications, including commercial, sociological or cybercrime contexts.
- The incomplete casual language used by interlocutors in chats presents a challenge for author profiling. However, it also offers opportunities for predicting author identity, such as geographic origin or educational level, due to the specific words used in chat conversations.
- Goal of the thesis:
  - Investigate and discuss the state of the art
  - Develop a concept for author profiling in chats
  - Implement and evaluate the proposed method
  - Relevant datasets: PAN Datasets, PJ Dataset

Martin Steinebach / Betreuerin: SF
Martin.Steinebach@sit.fraunhofer.de,

SEXUAL PREDATOR IDENTIFICATION IN CHATS

- Online sexual predation represents a significant social issue. The development of effective methods for identifying predators can have a profound impact on enhancing online safety, particularly for vulnerable underage groups.
- The objective of this thesis is to develop a system which enables
  - (1) the differentiation of sexual predators from victims engaged in online chat conversations, and
  - (2) the identification of the precise lines of the conversation which indicate a propensity for sexual exploitation of other chat participants.
- Available dataset: PAN 12

Source: WikiHow

Martin Steinebach / Betreuerin: SF
Martin.Steinebach@sit.fraunhofer.de,