Giada Zingarini

Detection of fully and locally AI-generated images

Tutor: Luisa Verdoliva

Cycle: XXXVIII

Year: Second

My background

- **MSc degree** in Biomedical Engineering, curriculum in Biorobotic and Bionic – Università degli Studi di Napoli Federico II
- **Research group**: GRIP (Image Processing Research Group)
- **PhD start date**: 01/11/2022
- **Scholarship type**: UNINA - DII, DISCOVER project, funded by DARPA under the SEMAFOR program

Research field of interest

- **Multimedia Forensics**:
  - Analysis of multimedia data for forensic applications
- **Image forgery localization and detection**:
  - Development of methods for detecting synthetic images and localizing the manipulated parts

Which image is synthetic ?

Research field of interest

- **Multimedia Forensics**:
  - Analysis of multimedia data for forensic applications
- **Image forgery localization and detection**:
  - Development of methods for detecting synthetic images and localizing the manipulated parts

Which image is synthetic ?

Research field of interest

- **Multimedia Forensics**:
  - Analysis of multimedia data for forensic applications
- **Image forgery localization and detection**:
  - Development of methods for detect synthetic images and localize the manipulated parts

Why?

The food is AI generated

Summary of study activities

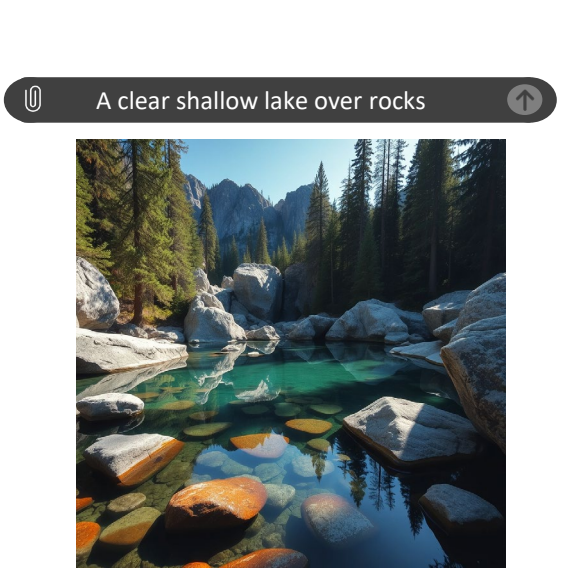
	Courses	Seminars	Research	Tutorship
Total	9	5.4	41	0.52
Expected	10 - 20	5 - 10	30 - 45	0 – 1.6

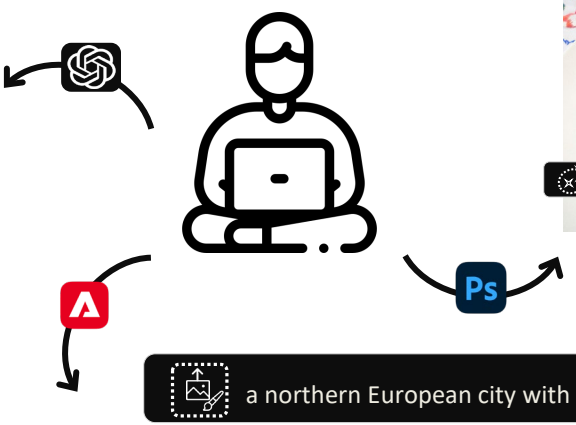
- Study of the state-of-the-art methods for synthetic image generation and for detection of fake images

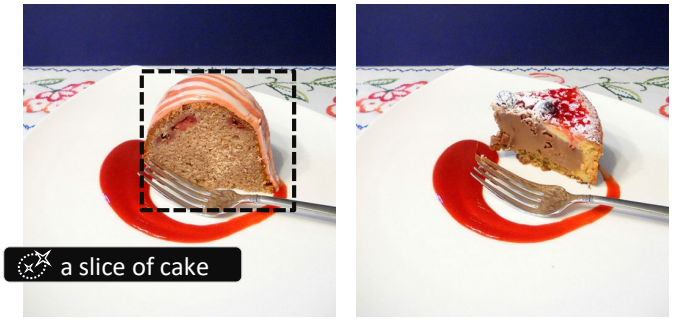
- **PhD School**:
  - "2024 IEEE SPS - Summer School "Understanding and modeling the world around us" at University Federico II of Naples "
- **PhD courses:**
  - "Strategic Orientation for STEM Research & Writing" - University Federico II of Naples (Dr. Chie Shin Fraser.)
- **Conferences**:
  - IEEE International Conference on Acoustics, Speech and Signal Processing 2024, Seoul, Republic of Korea.

Research activity: Overview

- Problem
  - AI generative tools are now easily accessible by any user with internet connection







Research activity: Overview

- Problem
  - AI tools can be maliciously used to spread **disinformation**

‘Verified’ Twitter accounts share fake image of ‘explosion’ near Pentagon, causing confusion

By Donie O'Sullivan and Jon Passantino, CNN
3 minute read · Updated 11:35 AM EDT, Tue May 23, 2023

NEWS | 22 November 2023

ChatGPT generates fake data set to support scientific hypothesis

Researchers say that the model behind the chatbot fabricated a convincing bogus database, but a forensic examination shows it doesn't pass for authentic.

By Miryam Naddaf

HOME › ARCHIEF › SCAMMERS ARE STEALING HOMES FROM UNDER THEIR OWNERS' NOSES. AI IS MAKING IT SCARILY EA

Scammers are stealing homes from under their owners' noses. AI is making it scarily easy.

Jordan Pandy, Katie Balevic
22 okt 2024

AUGUST 15, 2024 SHARE DOWNLOAD PDF

PEER REVIEWED

How spammers and scammers leverage AI-generated images on Facebook for audience growth

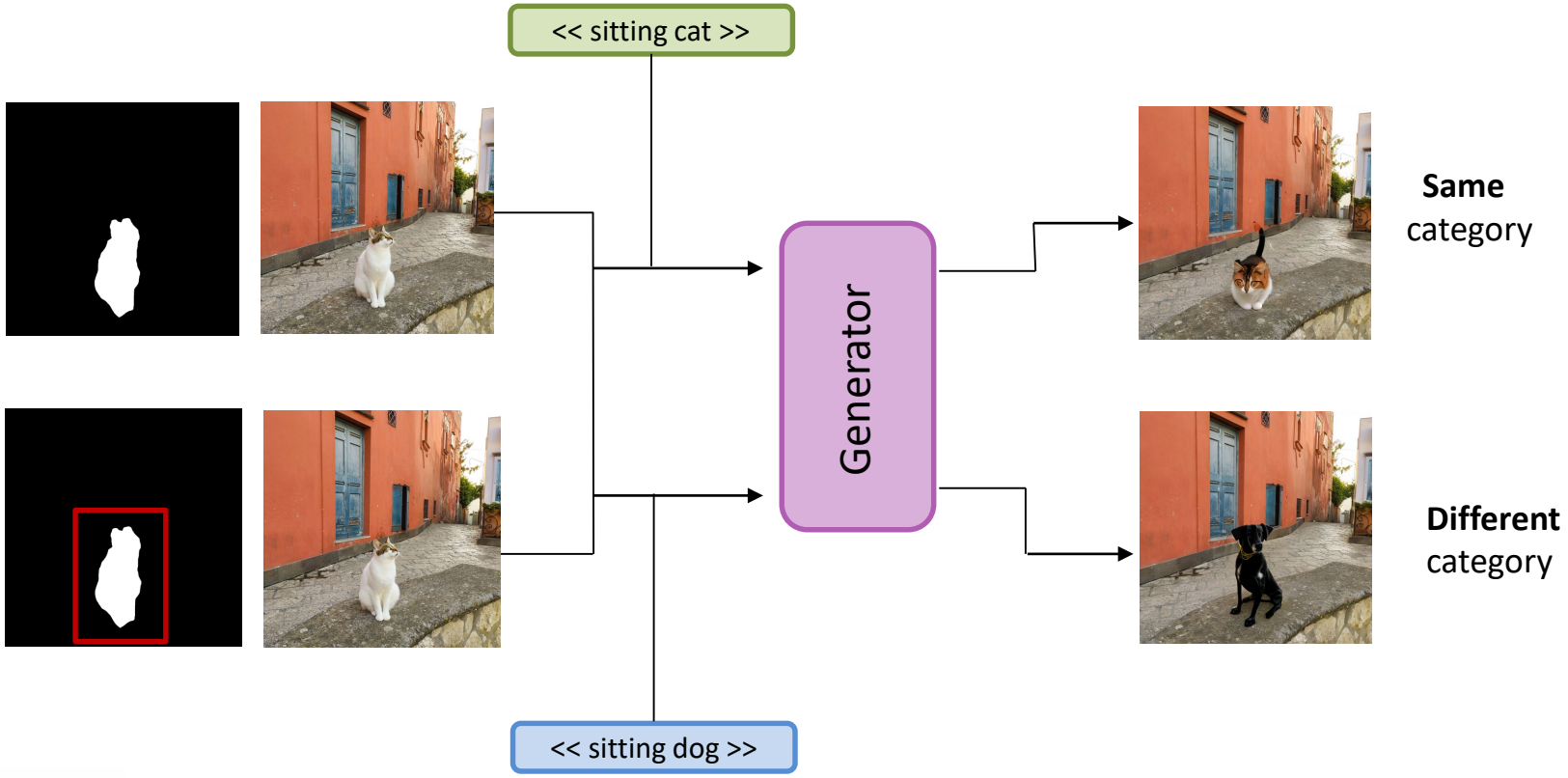
Research activity: Overview

- Problem

- Objective
  - **Create** synthetic datasets with different generators to train robust detection methods
  - Improve the performance of methods for the **detection** of synthetic data shared over the web

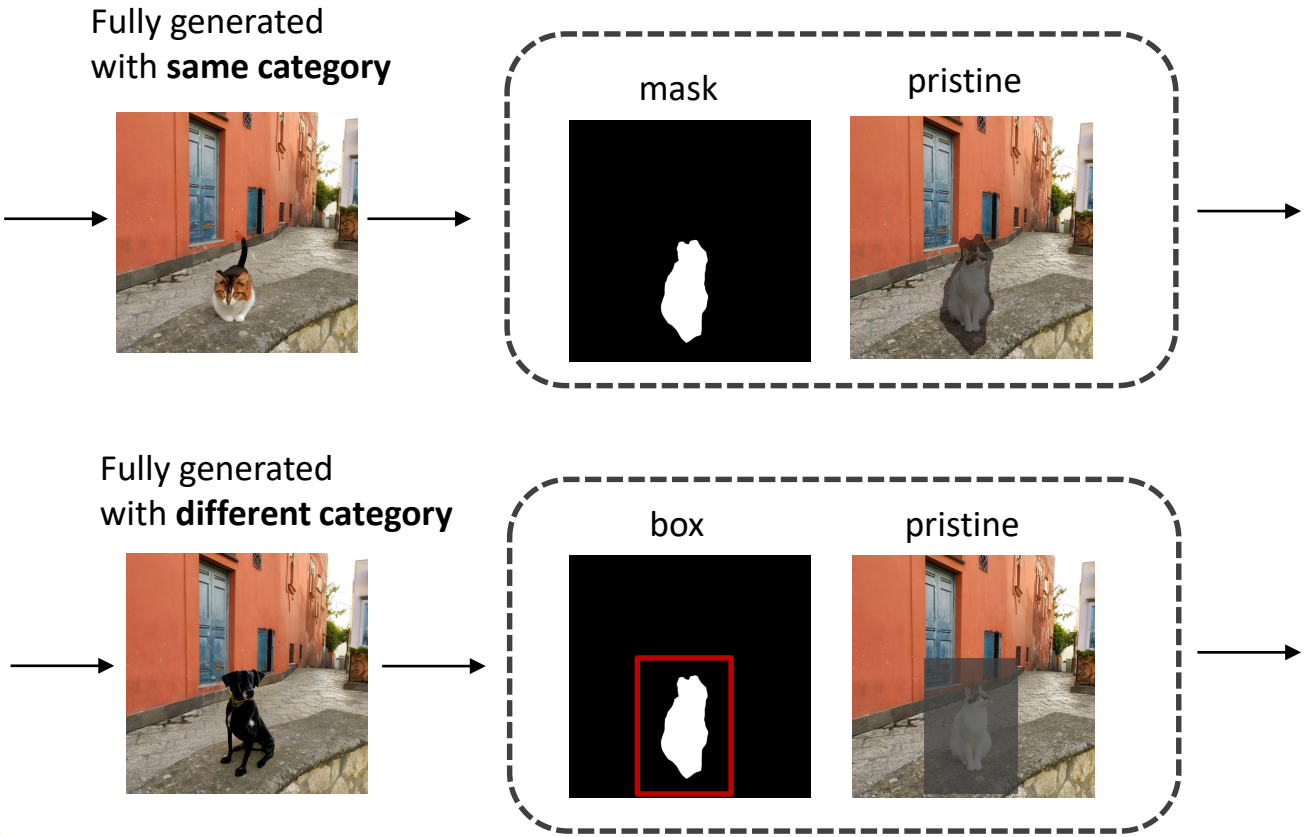
Research activity: Overview

- **Methodology**
  - We generated manipulated images by employing several different synthetic generators



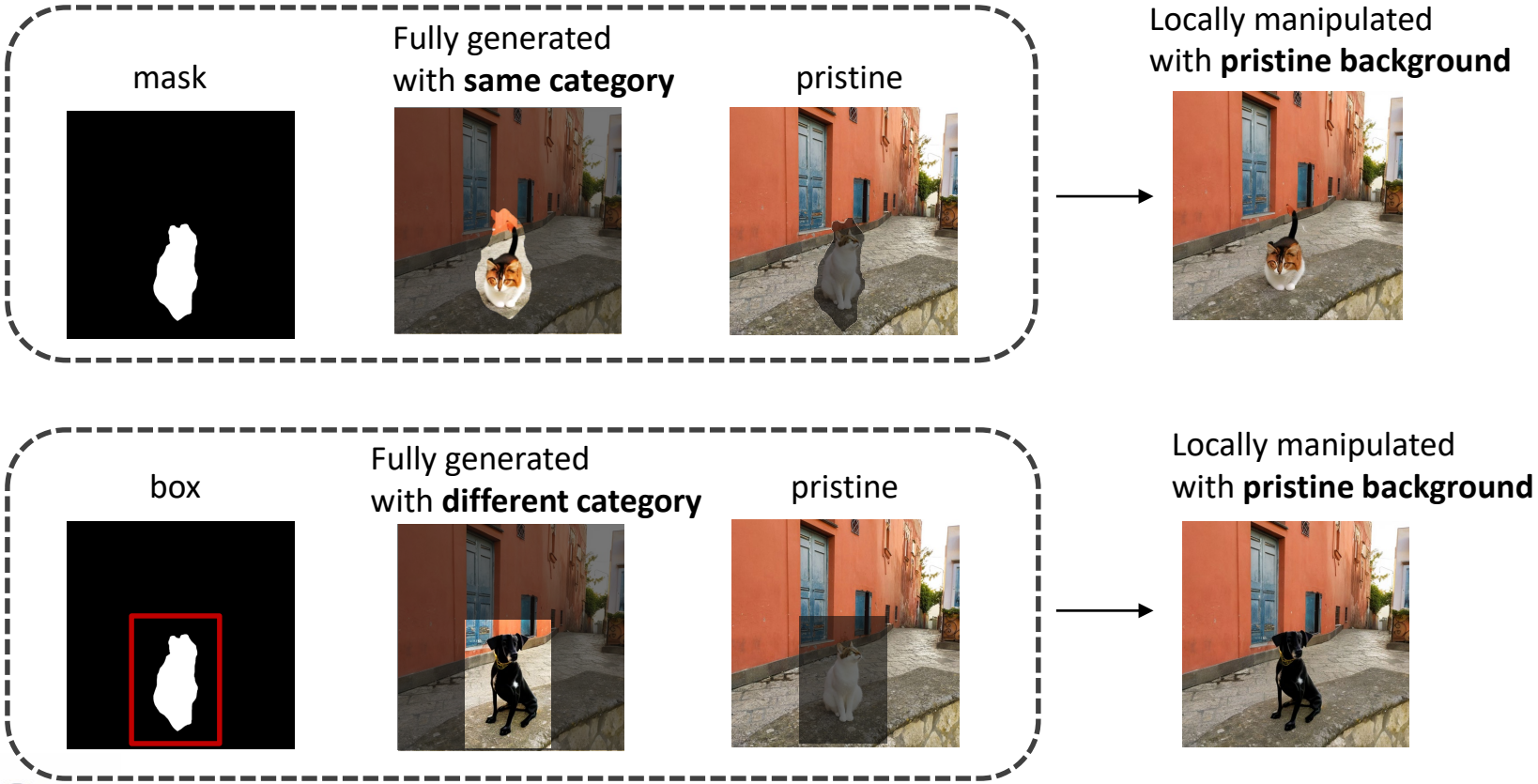
Research activity: Overview

- **Methodology**
  - Each generated image has an alternative version with pristine background



Research activity: Overview

- **Methodology**
  - Each generated image has an alternative version with pristine background



Research activity: Overview

- **Methodology**
  - We use the dataset to train a baseline that leverages pre-trained vision-language models
  - We investigate the use of fully and locally generated data during training to improve the detection of AI-generated images

AUC	Adobe Ps	Autosplice	Laion Ip2p	Synthbuster	AVG
Corvi et al.	62.59	47.87	**99.99**	90.50	75,24
Ojha2023	80.13	85.48	88.46	79.70	83.39
PSCC-Net	80.12	**99.31**	68.39	55.20	75.75
HiFi-Net	65.06	57.39	83.78	43.56	62.44
Ours	**98.68**	95.67	96.94	**92.07**	**95.84**

Good results on both fully and locally generated data

Products

[P1]	**Conference Paper** R. Corvi, D. Cozzolino, **G. Zingarini**, G. Poggi, K. Nagano, and L. Verdoliva, “On the detection of synthetic images generated by diffusion models”, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2023, Rhodes, Greece.
[P2]	**Conference Paper** **G. Zingarini**, D. Cozzolino, R. Corvi, G. Poggi, L. Verdoliva, "M3Dsynth: A dataset of medical 3D images with AI-generated local manipulations", in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2024", Seoul, Republic of Korea.

Next year

- Develop a method that can also localize the manipulated regions
- Work on the attribution task, i.e. identify the type of synthetic generator

Thank you for the attention!