Francesco Vultaggio [1,2]
Predrag Djindjic [1,3]
Phillipp Fanta-Jende [1]
Markus Gerke [2]
Sebastian Tschiatschek [3]

{name.surname}@ait.ac.at, m.gerke@tu-bs.de, sebastian.tschiatschek@univie.ac.at

Funded by the European Union

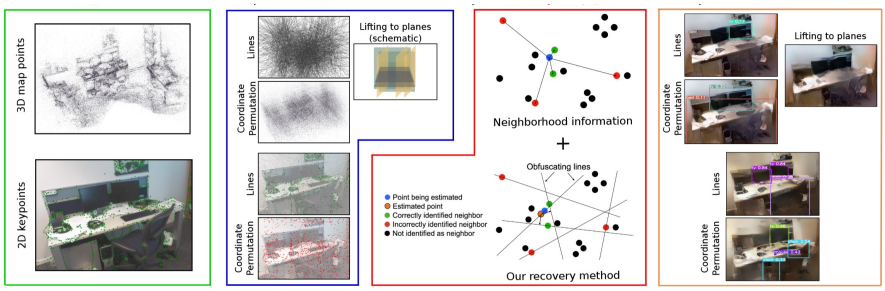
MISANTHROPE: A PRIVACY PRESERVING KEYPOINT DETECTOR

Problem definition

(a) SfM point cloud (top view) (b) Projected 3D points (c) Synthesized Image (d) Original Image

Pittaluga, F., Koppal, S. J., Kang, S. B., & Sinha, S. N. (2019). Revealing scenes by inverting structure from motion reconstructions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 145-154).

Modern *inversion* networks can reconstruct images from sparse set of keypoints and descriptors. Inversion attacks pose a threat to the privacy of any system in which image features leave the user device, such as in distributed robotics, visual localization and many others.

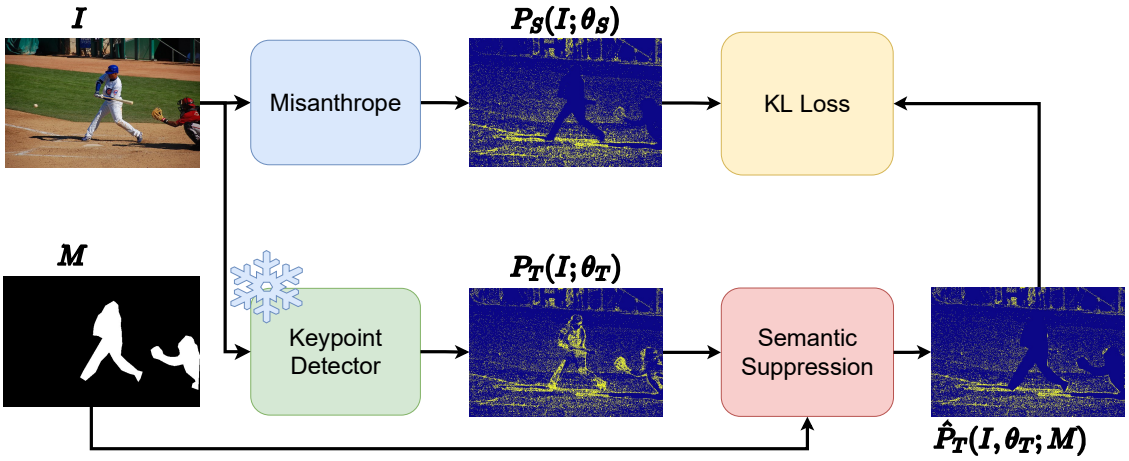


Chelani, K., Benbihi, A., Kahl, F., Sattler, T., & Kukelova, Z. (2025, March). Obfuscation Based Privacy Preserving Representations are Recoverable Using Neighborhood Information. In *2025 International Conference on 3D Vision (3DV)* (pp. 189-199). IEEE.

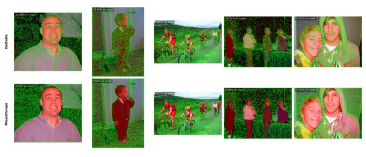
Prior art focused on obfuscating or quantizing all the information contained in the scene or image but obfuscation methods have been counteracted by newer inversion techniques and descriptor quantization comes at the cost of accuracy

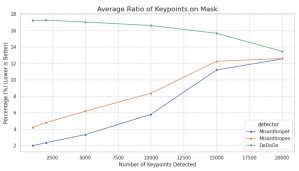
Method

We propose to mitigate the threat posed by inversion attacks by removing private information, in this work defined to be people, through a process of *guided self distillation*. A teacher instance of a detector model is used to generate keypoint probability maps which are then made private through the use of semantic masks and finally used as labels for the student model. We call the resulting model **Misanthrope** and show it to be more robust to inversion attacks and a better feature extractor in scenes where people act as distractors.



Privacy Experiments





Keypoint detection on people: Misanthrope can effectively reduce the percentage of keypoints detected on people. Even in failure cases it avoids detecting keypoints on faces, showcasing misanthrope learned a robust concept of personhood

Method	LKT		HKT	
	PSNR	mIoU	PSNR	mIoU
DeDoDe	17.71	0.744	19.32	0.818
Misanthrope	**16.69**	**0.397**	**18.09**	**0.636**
GT Mask	17.00	0.578	18.12	0.684

Person detection on inverted images: Misanthrope can counteract so called *inversion attacks*. We showcase how by avoiding sampling any information on people these become invisible to subsequent person detection attempts

Matching Experiments

Method	British Museum	Florence Cathedral	Lincoln Memorial	London Bridge	Milan Cathedral	Mount Rushmore	Piazza San Marco	Sagrada Familia	St Paul's Cathedral	Avg. Rank
Aliked	72.69	71.24	82.87	50.79	82.24	37.47	57.73	75.86	78.83	2.78
Dedode	82.36	77.87	**84.14**	52.59	87.11	33.47	63.21	81.31	88.87	2.00
Disk	68.53	71.07	78.46	48.22	81.37	**43.03**	56.38	74.83	74.53	3.56
Xfeat	61.52	61.46	71.27	40.09	63.39	23.39	47.29	56.04	63.62	5.22
Misanthrope	**83.25**	**78.08**	67.35	**53.00**	**87.41**	28.18	**64.13**	**81.62**	**89.01**	**1.89**

IMC2021 Phototurism: Misanthrope all other models we tested in the phototourism test set from the IMC2021 in seven out of nine scenes. It does underperform only in the Lincoln memorial and Mount Rushmore scenes where it does not detect keypoints on the statue mistaking them for people

Method	MRE [°]	MTE [°]	mAA @1°	mAA @5°	mAA @10°
Aliked	0.43	1.39	35.02	85.95	93.14
Dedode	0.43	1.38	35.31	83.12	88.98
Disk	0.45	1.56	31.65	83.12	91.91
Superpoint	0.43	1.53	31.34	86.20	**93.94**
Xfeat	0.46	1.82	27.85	79.97	88.99
Misanthrope	**0.42**	**1.21**	**39.94**	**87.21**	91.45

WILD-SLAMGS: We test misanthrope in an indoor dataset in which people act as distractors by appearing in front of the camera. By avoiding detecting keypoints on people misanthrope can outperform other feature extraction methods

Method	MMA			MHA		
	1px [%]	3px [%]	5px [%]	1px [%]	3px [%]	5px [%]
Aliked	**44.03**	74.88	82.71	52.93	**84.14**	**90.00**
Disk	38.09	**76.82**	**84.47**	50.69	80.34	88.28
Dedode	33.49	73.69	83.24	**56.90**	83.97	**90.00**
Superpoint	25.83	61.02	72.14	50.69	81.55	89.14
Xfeat	19.89	56.56	71.13	47.07	81.72	89.83
Misanthrope	33.17	73.05	82.64	54.31	81.72	88.79

Hpatches: We test misanthrope on a standard image matching dataset where no people are present and observe how it retains most of the performances of the original DeDoDe model from which it was self-distilled.