

# Tom Fischer

☎ (+49) 157 8965 5134    ✉ [fischertom@outlook.com](mailto:fischertom@outlook.com)    🏠 [Homepage](#)

Ph.D. Candidate in 2D / 3D Computer Vision and Machine Learning, specialized in object-level representation learning for detection, 9D object pose estimation, and continual learning, eager to contribute to and support your team.

## Education

---

Ph.D. Candidate in Computer Science    (expected graduation: late 2026)    *NÜREMBERG, GERMANY*  
*UNIVERSITY OF TECHNOLOGY NÜREMBERG*    04.2025 – present

Supervised by Prof. [Eddy Ilg](#) in the CVMP Lab

Ph.D. Candidate in Computer Science    *SAARBRÜCKEN, GERMANY*  
*SAARLAND UNIVERSITY*    10.2023 – 04.2025

Supervised by Prof. [Eddy Ilg](#) in the CVMP Lab

M.Sc. in Computer Science    GPA: 1.5/4.0 (best: 1.0, worst: 4.0)    *SAARBRÜCKEN, GERMANY*  
*SAARLAND UNIVERSITY*    10.2020 – 05.2023

Thesis title: “Optical Flow with Explicit Diffusion”

B.Sc. in Cybersecurity    *SAARBRÜCKEN, GERMANY*  
*SAARLAND UNIVERSITY*    10.2016 – 10.2020

## Publications

---

Unified Category-Level Object Detection and Pose Estimation from RGB Images using 3D Prototypes,  
[Tom Fischer](#)<sup>\*</sup>, Xiaojie Zhang<sup>\*</sup>, Eddy Ilg. In *Proceedings of the 20<sup>th</sup> International Conference on Computer Vision (ICCV)*

iNeMo: Incremental Neural Mesh Models for Robust Class-Incremental Learning,  
[Tom Fischer](#), Yaoyao Liu, Artur Jesslen, Noor Ahmed, Prakhar Kaushik, Angtian Wang, Alan Yuille, Adam Kortylewski, Eddy Ilg. In *Proceedings of the 18<sup>th</sup> European Conference on Computer Vision (ECCV)*

Neuroexplicit Diffusion Models for Inpainting of Optical Flow Fields,  
[Tom Fischer](#), Pascal Peter, Joachim Weickert, Eddy Ilg. In *Proceedings of the 41<sup>st</sup> International Conference on Machine Learning (ICML)*

<sup>\*</sup> Equal contribution.

## Professional Experience

---

Research Intern    *SAARBRÜCKEN, GERMANY*  
*SAARLAND UNIVERSITY*    05.2023 – 10.2023

Researching Neuroexplicit Diffusion Models under Prof. [Eddy Ilg](#) and Prof. [Joachim Weickert](#)

Working Student    *KARLSRUHE, GERMANY*  
*ATRU VIA AG*    01.2022 – 04.2023

## Programing Languages and Deep Learning Frameworks

---

Python, PyTorch, PyTorch3D, CUDA

## Reviewing

---

ICCV (1x), CVPR (2x), AAAI (1x), ECCV (1x), TPAMI (3x)

## Tutoring

---

Computer Vision, Master’s Project, 3D Computer Vision, Cryptography

## Author Contribution Statements & Experience in Brief

---

### ANNOTATED PUBLICATIONS

1. Unified Category-Level Object Detection and Pose Estimation from RGB Images using 3D Prototypes [ICCV/25].  
**Problem:** Detect objects and estimate **9D pose** (position, orientation, size) from single RGB image.  
**Method:** Learn set of **3D prototypes** for 2D/3D correspondence estimation from feature matching with novel transformer-based architecture and size refinement.  
**My Role:** Led method & implementation; designed training protocol & 9D inference pipeline; designed experiments; wrote major sections and coordinated submission.  
**Impact:** **SOTA** on REAL275 (+**22.9%** avg across scale-agnostic metrics); more robust to image corruptions (**-14%** vs. **-19–37%** for baselines).
2. iNeMo: Incremental Neural Mesh Models for Robust Class-Incremental Learning [ECCV/24].  
**Problem:** Learn **new object categories over time** without catastrophic forgetting, while still estimating **3D pose**.  
**Method:** Uses **expandable** set of Neural Mesh Models plus regularization/distillation to preserve past knowledge; **cameraview-aware** exemplar selection and **uncertainty-aware** classification.  
**My Role:** Led design, implementation, losses, and experiments; wrote major sections and coordinated submission.  
**Impact:** Beats baselines by **2–6%** in-domain and **6–50%** OOD; **first** incremental pose-estimation framework; strong under heavy occlusions (**60–80%** occluded) outperforming baselines by up to **70%**.
3. Neuroexplicit Diffusion Models for Inpainting of Optical Flow Fields [ICML/24].  
**Problem:** Restores missing motion cues between frames to improve downstream vision tasks.  
**Method:** Combines a network that predicts **how motion should diffuse** with a stable, explicit update step across multiple scales based on the theory of diffusion.  
**My Role:** Conceived and implemented the full pipeline; designed all experiments and ablations; wrote major sections and coordinated submission.  
**Impact:** New **SOTA** for filling in missing **optical flow** (pixel-wise motion): avg EPE gains vs. model-based +**11–27%**, and learning-based +**42–47%**; for KITTI, on-par with strongest baseline, while showing fewer outliers at low density.

### WORKING EXPERIENCE

1. Research Intern at Saarland University (05.2023-10.2023)  
**Scope:** Independent research on **diffusion-based optical-flow inpainting** distilled from my master's thesis. This work matured into our **ICML 2024** publication.  
**Outcome:** Evolved the thesis prototype into a publishable method; authored the manuscript and produced a fully reproducible codebase used for subsequent experiments.
2. Working Student at the Meaningful Conversational Intelligence Team at Atruvia AG (01.2022-04.2023)  
**Scope:** Contributed to **Text-to-Speech (TTS)** R&D and **Speech-to-Text (STT)** production systems used in the banking context.  
**Outcome:**  
**TTS:** Delivered an end-to-end prototype (model architecture, inference path, serving design) that informed later product decisions.  
**Contribution:** Selected and justified the model architecture; implemented the inference/serving pathway with a focus on latency and resource use; prepared integration points for internal platforms.  
**STT:** Hardened and shipped the production service via Kubernetes and CI/CD; the system was put into production early in 2023.  
**Contribution:** Managed day-to-day deployments on **Kubernetes** using the company **CI/CD** pipeline; improved rollout reliability and resolved deployment issues alongside platform/product teams.