

# Zeyun Zhong

📍 Karlsruhe, Germany | ✉ zeyun.zhong@kit.edu

🌐 [linkedin.com/in/zeyun-zhong](https://www.linkedin.com/in/zeyun-zhong) | 🐙 [github.com/zeyun-zhong](https://github.com/zeyun-zhong) | 🎓 Google Scholar

PhD candidate working on streaming video understanding, action anticipation, and real-time vision-language models.  
Seeking 2026 internship / full-time roles in video understanding and multimodal learning.

## Education

<b>Karlsruhe Institute of Technology (KIT)</b> Ph.D. Candidate in Streaming Video Understanding Advisor: Prof. Dr. Juergen Beyerer, Prof. Dr. Juergen Gall	Karlsruhe, Baden-Württemberg Nov. 2021 – present
<b>Leibniz University Hannover</b> M.Sc. in Mechatronics and Robotics (with distinction)   Final Grade: 1.2 (Excellent, German Scale)	Hannover, Lower Saxony Apr. 2018 – Mar. 2021
<b>University of Applied Sciences and Arts Hannover</b> B.Eng. in Process Engineering and Energy Technology   Final Grade: 1.7 (Good, German Scale)	Hannover, Lower Saxony Sep. 2014 – Feb. 2018

## Experience

<b>Research Assistant</b> IES, Karlsruhe Institute of Technology (KIT) & Fraunhofer IOSB	Nov. 2021 – present Karlsruhe, Baden-Württemberg
<b>• Research &amp; Algorithmic Contributions:</b> <ul style="list-style-type: none"><li>– Proposed E<sup>2</sup>-TTT to efficiently learn long-context beyond the fixed-window constraints of transformer models, sustaining &gt;90% passkey-retrieval accuracy at 8× the training context.</li><li>– Developed FlowNar, an efficient streaming video narration framework that dynamically manages the visual context, supporting 10× longer videos at 3× higher throughput than prior online LMMs.</li><li>– Designed several novel action anticipation architectures achieving superior performance on large-scale benchmarks, including Epic-Kitchens and Ego4D.</li><li>– Authored a comprehensive survey of 90+ deep learning action anticipation methods, introducing a taxonomy by contribution and consolidating datasets, metrics, and future directions.</li></ul>	
<b>• Grant Acquisition &amp; Industry Impact:</b> <ul style="list-style-type: none"><li>– KIMoS Project: Led grant acquisition (€180K) for real-time assembly assistance, securing three industrial partners, and represented Fraunhofer IOSB to exhibit the system at Hannover Messe 2026.</li><li>– JuBot Project: Built and deployed a multimodal action recognition system that identifies multiple people's actions in real time on the KIT ARMAR humanoid robot.</li></ul>	
<b>Master Thesis &amp; Research Assistant</b> IPI, Leibniz University Hannover	Oct. 2020 – Sep. 2021 Hannover, Lower Saxony
<b>• Designed probabilistic models for uncertainty estimation in dense stereo matching (Grade: 1.0) [Paper][Code].</b>	
<b>Research Intern</b> Volkswagen AG   Research on adversarial attacks and robustness for semantic segmentation	Nov. 2019 – Apr. 2020 Wolfsburg, Lower Saxony
<b>Development Intern &amp; Bachelor Thesis</b> Robert Bosch GmbH   Simulation of gasoline engine exhaust systems	May 2017 – Dec. 2017 Schwieberdingen, Baden-Württemberg

## Selected Publications

- Z. Zhong, et al. "Rethinking Expressivity and Efficiency in Test-Time Training." Under Review.
- Z. Zhong, et al. "FlowNar: Scalable Streaming Narration for Long-Form Videos." ICML 2026. [Page][Paper][Code]
- Z. Zhong, et al. "Scalable Video Action Anticipation with Cross Linear Attentive Memory." WACV 2026. [Paper][Code]
- Z. Zhong, et al. "A survey on deep learning techniques for action anticipation." TPAMI Under Review. [Paper]
- Z. Zhong\*, D. Schneider\*, et al. "Anticipative feature fusion transformer for multi-modal action ..." WACV 2023. [Paper][Code]

## Honors & Awards

- 2nd Place out of 15 teams, Ego4D Long-term Action Anticipation Challenge @ CVPR 2024 [Page][Report]
- 6th Place out of 30 teams, KIT Neuland Innovation Contest [Page]

## Skills

**Languages:** Chinese (Native), English (Fluent), German (Fluent)  
**Programming:** Python, Matlab, C++  
**Libraries & Tools:** PyTorch, HuggingFace, DeepSpeed, TensorFlow, OpenCV, ROS