

Hidir YESILTEPE

✉ hidir@vt.edu | 📧 hidiryasiltepe | 🌐 yesiltepe-hidir | 📄 yesiltepe-hidir.github.io

SUMMARY

Ph.D. student focusing on foundational image and video models for consistent semantic editing and motion transfer. Currently exploring training-free methods by applying concepts from statistical physics. Aims to develop creative solutions for image and video manipulation.

EDUCATION

Virginia Tech, Blacksburg, VA — *Ph.D. in Computer Science* AUG 2023 - MAY 2028
METU/ODTU, Ankara, Turkey — *B.S. in Computer Engineering* SEP 2018 - JUN 2023

PUBLICATIONS

- **Yesiltepe, H.**, Meral, T. H. S., Dunlop, C., & Yanardag, P. (2024). MotionShop: Zero-Shot Motion Transfer in Video Diffusion Models with Mixture of Score Guidance. **Preprint, 2025**. <https://motionshop-diffusion.github.io/>
- Meral, T. H. S., **Yesiltepe, H.**, Dunlop, C., & Yanardag, P. (2024). MotionFlow: Attention-Driven Motion Transfer in Video Diffusion Models. **Preprint, 2025**. <https://motionflow-diffusion.github.io/>
- Dalva, Y., **Yesiltepe, H.**, & Yanardag, P. (2024). GANTASTIC: GAN-based Transfer of Interpretable Directions for Disentangled Image Editing in Text-to-Image Diffusion Models. arXiv preprint arXiv:2403.19645. **Preprint, 2025**
- **Yesiltepe, H.**, Akdemir, K., & Yanardag, P. (2024). MIST: Mitigating Intersectional Bias with Disentangled Cross-Attention Editing in Text-to-Image Diffusion Models. arXiv preprint arXiv:2403.19738. **Preprint, 2025**. <https://mist-diffusion.github.io/>
- Kara, O., Kurtkaya, B., **Yesiltepe, H.**, Reh, J. M., & Yanardag, P. (2024). Rave: Randomized noise shuffling for fast and consistent video editing with diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6507-6516). **CVPR 2024 - Highlight**. <https://rave-video.github.io/>
- **Yesiltepe, H.**, Dalva, Y., & Yanardag, P. (2024). The Curious Case of End Token: A Zero-Shot Disentangled Image Editing using CLIP. arXiv preprint arXiv:2406.00457. **CVPR 2024 - AI4CC Workshop**
- Zheng, M., Simsar, E., **Yesiltepe, H.**, Tombari, F., Simon, J., & Yanardag, P. (2024). Stylebreeder: Exploring and democratizing artistic styles through text-to-image models. arXiv preprint arXiv:2406.14599. **NeruIPS 2024**. <https://stylebreeder.github.io/>

RESEARCH EXPERIENCE

Amazon — *Applied Scientist Intern* MAY 2024 - AUG 2024

- Research intern in image and video virtual try-on.
- Conducted research on training-free & mask-free image virtual try on methods using diffusion models.
- Explored the capabilities developed method on zero-shot video editing in VTON domain.
- Extended AnimateDiff framework by implementing video-based virtual try-on capabilities, enabling dynamic clothing visualization and garment swapping in motion sequences.

University of Edinburgh — *Visitor Researcher* JULY 2022 - OCT 2022

- Hierarchical Deterministic Regularized Autoencoder: Conducted research on devising generative hierarchical deterministic regularized encoder-decoder pairs in image synthesis, evaluated their performance with respect to their stochastic counterparts Hierarchical Variational Autoencoders, specifically NVAE.

- Supervision: Took supervision by Professor Antonio Vergari.

University College London — *Visitor Researcher*

JULY 2022 - OCT 2022

- Learning Discrete Representations in Hierarchical VAEs: Conducted research on equipping hierarchical VAEs with discrete latent representations. Different than the existing works which utilize Straight-Through Estimators, Gumbel-Max and Gumbel-Softmax Trick we focused on using simple yet effective discrete optimization: Implicit MLE.
- Supervision: Took supervision by Professor Pasquale Minervini.

KTH Royal Institute of Technology — *Visitor Researcher*

JULY 2021 - OCT 2021

- Gesture Generation: Participated in studies concerning synthetic gesture generation for virtual avatars, in particular, decoding underlying emotions into gestures using representation learning.
- Supervision: Took supervision by Professor Hedvig Kjellström and Dr. Taras Kucherenko.

PROFESSIONAL EXPERIENCE

Radius AI — *Research Engineer*

OCT 2022 - FEB 2023

- Designed advanced neural networks for real-time Computer Vision applications, focusing on video stream analysis.
- Developed systems to extract comprehensive behavioral information from video using human pose and gesture estimation.
- Implemented cutting-edge algorithms to interpret and analyze human behavior patterns in complex video data.

HONORS & AWARDS

ICPC, European Regional Finals

FINALIST - 2021

ICPC, European Regional Finals

FINALIST - 2020

inzva Algorithm Competition

RANKED 1ST AMONG 63 TEAMS - 2020

inzva Winter Camp Algorithm Competition

RANKED 1ST AMONG 132 PARTICIPANTS - 2020

EVENTS

Amazon-Virginia Tech Initiative — *Speaker*

OCT 2024

- Gave a talk on internship experience at Amazon as an Applied Scientist Intern.

Oxford Machine Learning Summer School — *Participant*

MAY 2023 - JULY 2023

- Participated Oxford Machine Learning Summer School Finance and Health tracks.