

Serdar Yıldız

🏠 [serdaryildiz](#) 🎓 [Scholar](#) 🐙 [Github](#) 🔗 [Linkedin](#) ✉️ [e-mail](#)

EDUCATION

Yıldız Technical University <i>MSc Computer Engineering</i>	2021 – 2024 GPA: 3.93/4.0
Czech Technical University In Prague <i>Erasmus+ program / Incomplete due to covid-19</i>	Spring 2020
Yıldız Technical University <i>B.S. Computer Engineering (High Honour Student)</i>	2016 – 2021 GPA: 3.54/4.0

EXPERIENCE

The Scientific and Technological Research Council of Turkey <i>Researcher</i> I am engaged in computer vision research, leading a team. My work spans various fields including biometrics, object detection, image captioning, and OCR. I develop models with high generalization performance and throughput.	Jun. 2021 – Present
ID3 Software <i>Intern</i> I worked as an intern on a project focused on monitoring airport ground services, which also contributed to my BSc thesis.	Aug. 2020 – March 2021
Vakıf Emeklilik ve Hayat <i>Intern</i> I worked as a intern, working on a project analyzing customer churn in the private pension system.	Jun. 2019 – Aug. 2019

PROJECTS AND PUBLICATIONS

IMAGE CAPTIONING

I have been addressing the gap in the Turkish image captioning literature by developing new models and creating a unique dataset, aiming to enhance the ability of computers to understand and describe images in Turkish. On the path towards artificial general intelligence, my focus is on understanding the relationships between images and language.

TRCaptionNet++: A high-performance encoder-decoder based deep Turkish image captioning model fine-tuned with a large-scale set of pretrain data <i>Turkish Journal of Electrical Engineering and Computer Sciences</i> (Under Review) – Developed an advanced Turkish image captioning model combining CLIP and BERT technologies, outperforming existing models on the MS COCO and Flickr30K datasets.	2023
Turkish Image Captioning with Vision Transformer Based Encoders and Text Decoders <i>32nd Signal Processing and Communications Applications Conference</i> – Implemented a Turkish image captioning system using vision transformers and text decoders, achieving impressive results on the TasvirEt dataset with high performance metrics.	2024
TRCaptionNet: A novel and accurate deep Turkish image captioning model with vision transformer based image encoders and deep linguistic text decoders <i>Turkish Journal of Electrical Engineering and Computer Sciences</i> – Developed an advanced Turkish image captioning model combining CLIP and BERT technologies, outperforming existing models on the MS COCO and Flickr30K datasets.	2023
Automatic Turkish Image Captioning: The Impact of Deep Machine Translation <i>8th International Conference on Computer Science and Engineering</i> – Conducted a study on enhancing Turkish image captioning with deep machine translation, created a novel dataset from the MS COCO database using the NLLB model, and evaluated a custom Turkish captioning model.	2023

SCENE TEXT RECOGNITION

I identified a lack of real-world datasets in scene text recognition research and created the largest dataset in the literature for this field, introducing a lightweight model.

Turkish Scene Text Recognition: Introducing Extensive Real and Synthetic Datasets and a Novel Recognition Model 2024

Engineering Science and Technology, an International Journal – Introduced the Turkish Scene Text Recognition and Synthetic Turkish Scene Text Recognition datasets, significantly enhancing STR research with a novel Masked Vision Transformer model.

SEMI-SUPERVISED LEARNING

I have worked on semi-supervised learning to enhance the generalization performance of the models I developed.

Iterative Ensemble Pseudo-Labeling for Convolutional Neural Networks 2024

Sigma Journal of Engineering and Natural Sciences – Proposed a semi-supervised learning method combining ensemble approaches and pseudo-labeling, significantly enhancing CNN model accuracy by 15.9% on the STL-10 dataset.

PERSON RE-IDENTIFICATION

Existing literature on person re-identification models lacks the generalization capability for real-world application. In pursuit of this, I am working on developing models with superior generalization performance.

ENTIRE-ID: An Extensive and Diverse Dataset for Person Re-Identification 2024

The 18th IEEE International Conference on Automatic Face and Gesture Recognition – Introduced the ENTIRE-ID dataset, a vast and diverse collection designed to enhance person re-identification research by addressing domain variability and model generalization challenges.

GEO-LOCALIZATION

To address a gap in the emerging field of geo-localization for computer vision, I created the YILDIZ-VGL dataset, and I am currently investigating existing VPR models in the literature.

YILDIZ Visual Geo-localization: A Novel Dataset with Dense Coverage Under Diverse Environmental Conditions for Visual Place Recognition 2024

Image and Vision Computing(Under Review) – Introduced the YILDIZ-VGL dataset for visual place recognition, offering rich visual content with precise GPS data to enhance model development and evaluation under diverse conditions.

NUCLEI INSTANCE SEGMENTATION (MSc THESIS)

For my MSc thesis focused on nuclei instance segmentation, I've contributed to developing methods for improving accuracy in medical image analysis. The thesis includes exploring deep learning models and addressing challenges in segmenting histology images with imbalanced data.

A Comparative Analysis of Loss Functions in Segmentation of Medical Images with Highly Imbalanced Class Distribution 2024

International Conference on Innovations in Intelligent Systems and Applications – Conducted a comparative analysis of loss functions for medical image segmentation with unbalanced class distribution.

Semantic and Instance Segmentation of Multi-organ Cell Nuclei Using Deep Learning Based Methods 2024

32nd Signal Processing and Communications Applications Conference – Presented a study on cell nuclei segmentation in multi-organ histology images using various deep learning models

Segmentation of Cell Nuclei in Histology Images with Vision Transformer Based U-Net Models 2024

32nd Signal Processing and Communications Applications Conference – Analyzed the performance of vision transformer-based UNet models for cell nuclei segmentation in colon histology images.

Nuclei Instance Segmentation in Colon Histology Images with YOLOv7 2023
International Conference on Advanced Engineering, Technology and Applications – Implemented the first nuclei instance segmentation study using YOLOv7 on colon histology images.

Size-based Adaptive Instance Pruning for Refined Segmentation of Cell Nuclei in Histology Images 2023
31st Signal Processing and Communications Applications Conference – Presented an adaptive size-based instance pruning approach for more accurate cell nuclei segmentation in colon histology images.

Nuclei Segmentation in Colon Histology Images by Using the Deep CNNs: A U-Net Based Multi-class Segmentation Analysis 2022
Medical Technologies National Conference – Conducted a study on nuclei segmentation in colon histology images using the U-Net model, achieving promising results on the CoNIC Challenge dataset.

MITOTIC CELL DETECTION (TUSEB PROJECT)

I currently work as a scholar researcher on the TUSEB-managed melanoma cancer decision support system project, where I contribute to object detection.

CNMI-YOLO: Domain Adaptive and Robust Mitosis Identification in Digital Pathology 2024
Laboratory Investigation – proposes a YOLOv7 and ConvNeXt based novel architecture for mitosis detection

Performance analysis of the YOLO series for object detection: Detection of mitosis cells in histopathology images 2023
Medical Technologies National Conference – Explored the potential of YOLO architectures for real-time mitosis cell detection in pathology, demonstrating YOLOv8's superior performance on the MIDOG 2022 dataset across various cancer types.

TURNAROUND CONTROL SYSTEM (B.S. THESIS)

During my ID3 Software internship and as part of my bachelor's degree thesis in 2021, I focused on developing a turnaround control system to automatically detect and monitor airport ground services, utilizing object detection, single object tracking, and optical flow techniques.

A turnaround control system to automatically detect and monitor the time stamps of ground service actions in airports: A deep learning and computer vision based approach 2022
Engineering Applications of Artificial Intelligence – Developed a deep learning and computer vision system for detecting and monitoring airport ground services which aids in efficient management and cost reduction related to service delays.

CUSTOMER CHURN ANALYSIS (B.S. PROJECT)

During my internship at Vakıf Emeklilik ve Hayat, I worked on a project to analyze customer churn in the private pension system as part of my third-year bachelor's degree project. This was my initial experience with machine learning.

Customer Churn Analysis 2020
28nd Signal Processing and Communications Applications Conference – Developed a system to predict customer churn in Turkish private pension system with an average accuracy of 99.01%, using attribute selection and machine learning models.