Road Segmentation for Pedestrian Trajectory Prediction

Ang Xiao

Introduction

The goal of this project is to develop a machine learning model capable of recognizing walkable road areas in drone-captured images.

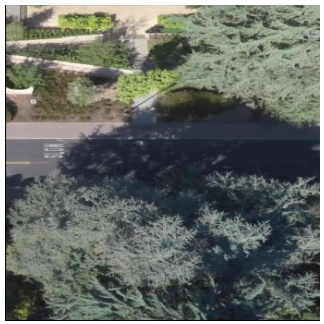
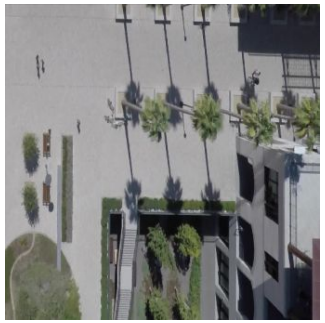
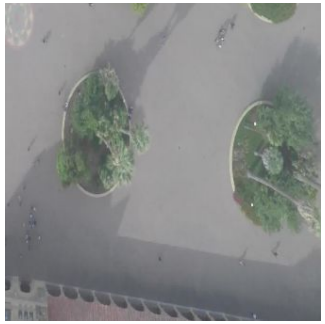
Drone-Based Pedestrian Tracking

Recognizing walkable road areas is helpful to predict the pedestrian trajectory

Introduction

Dataset

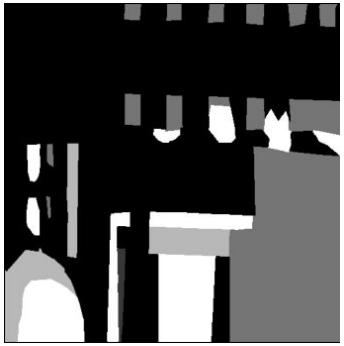
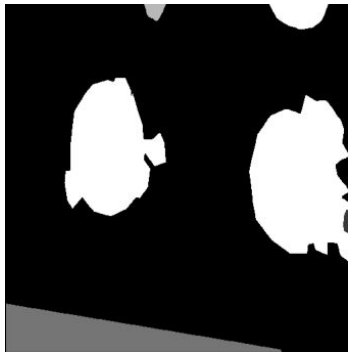
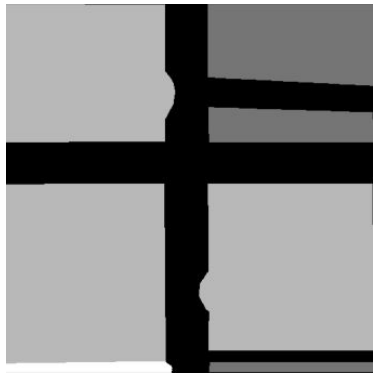
Stanford Drone Dataset[1].
The original dataset consists of videos. I utilize the still image backgrounds as the input.


Dataset

Mask images for supervised learning[2].

These masks help train the machine learning model to recognize walkable areas in drone images.

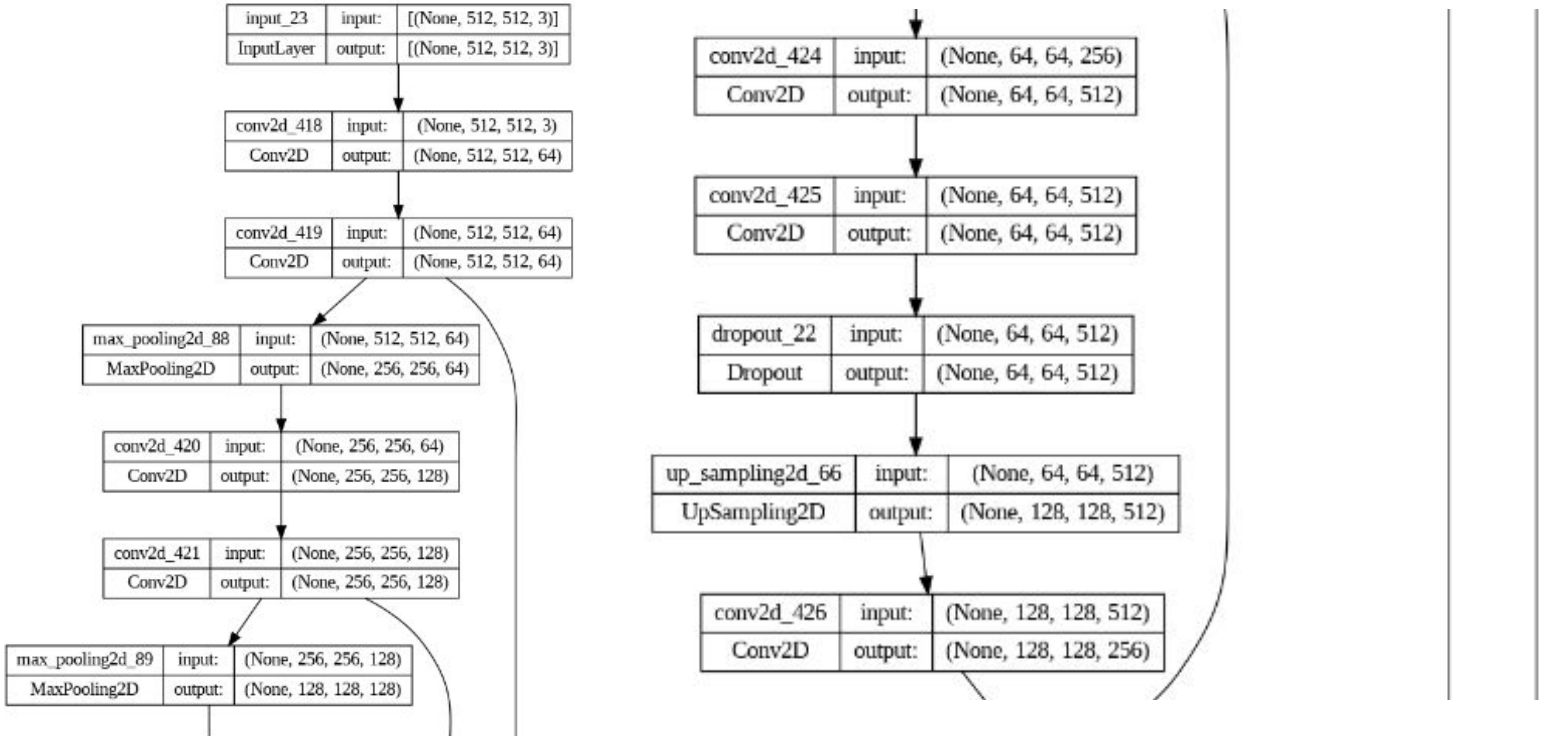
Data Preprocessing

All images are resized to 512x512 pixels and padded with zeros

Dataset Partition: 43 images for training and 17 images for testing.

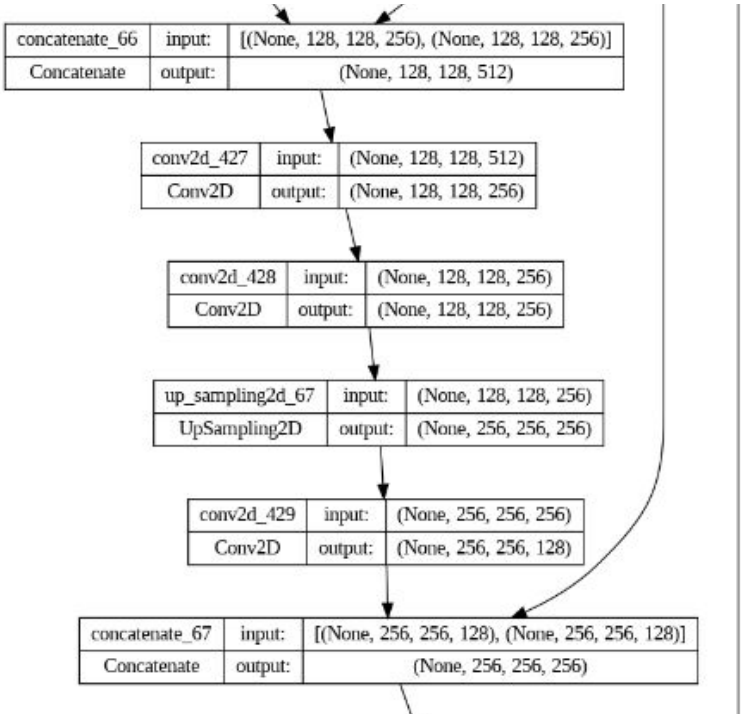
Model Design

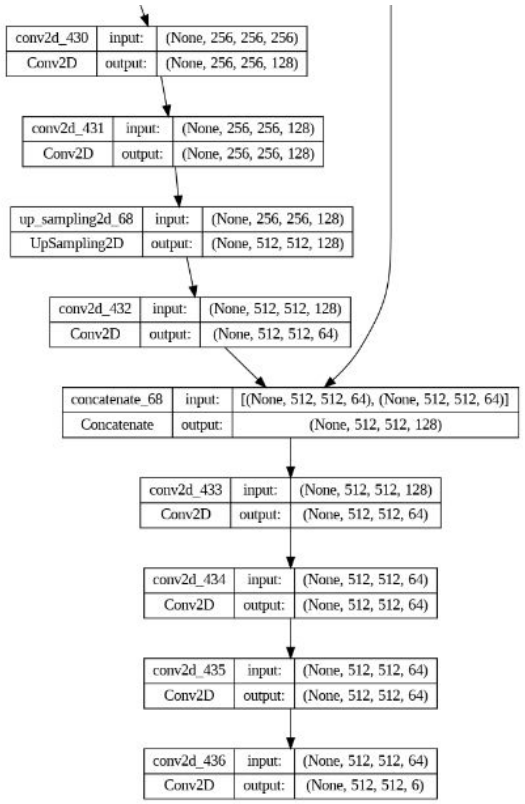
My model is based on the U-Net architecture, renowned for its efficiency in image segmentation tasks.



Model Design

My model is based on the U-Net architecture, renowned for its efficiency in image segmentation tasks.





Model Design

The architecture of my U-Net model includes the following:

Four downsampling stages, each with two convolutional layers and a max pooling layer.

The final stage also includes a dropout layer for regularization.

Model Design

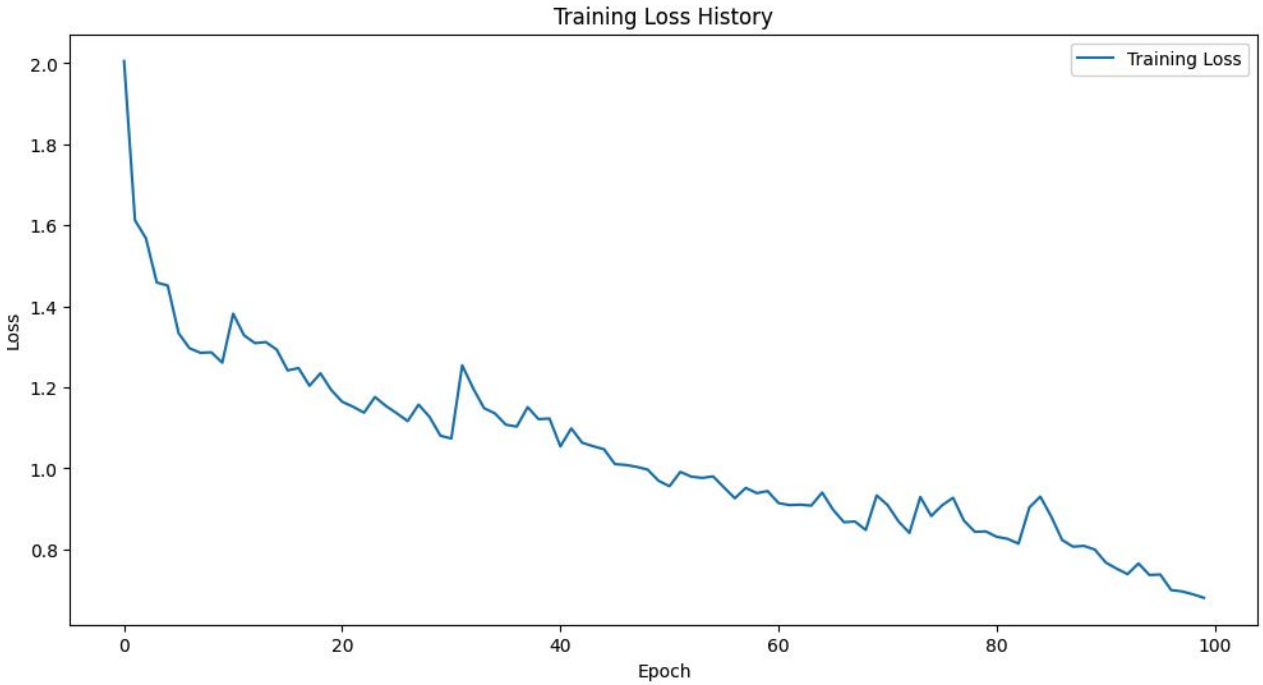
Three upsampling stages, each with an upsampling and two convolutional layers .

Each stage also includes a concatenation of the current output and the corresponding output from the downsampling stages.

Two final convolutional layers, with the last applying a softmax activation for class probability output.

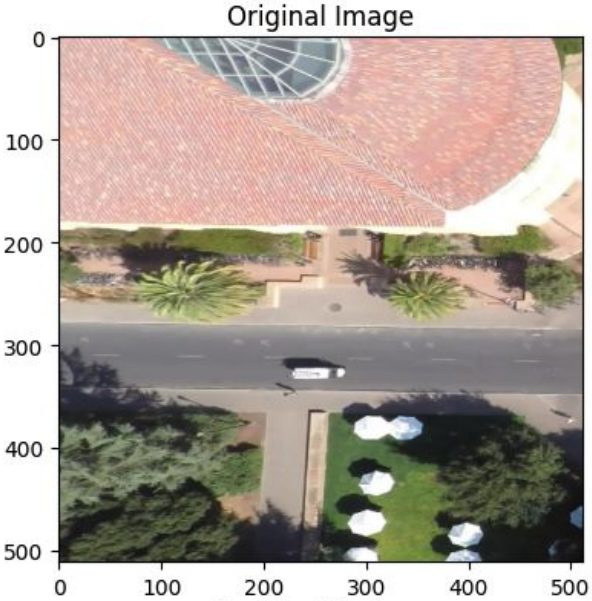
Model Training

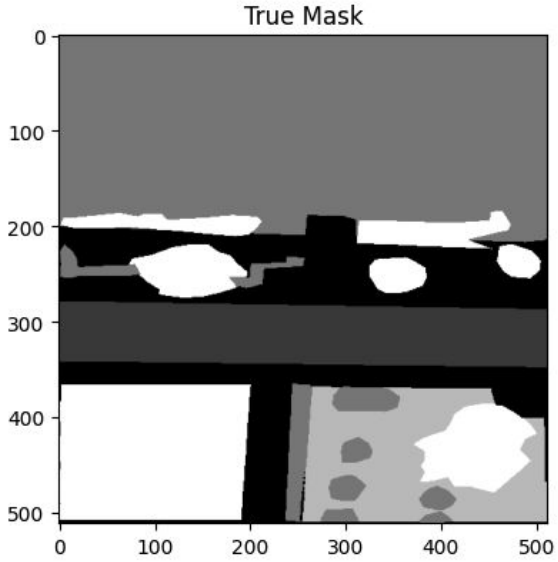
we train over 100 epochs, with a learning rate of 1e-4 and a batch size of 5.

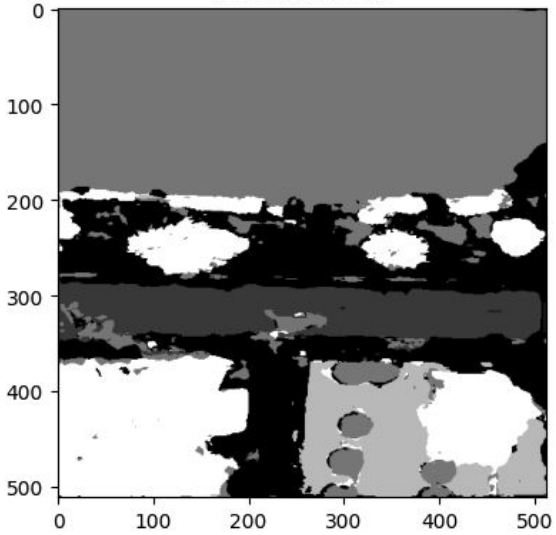


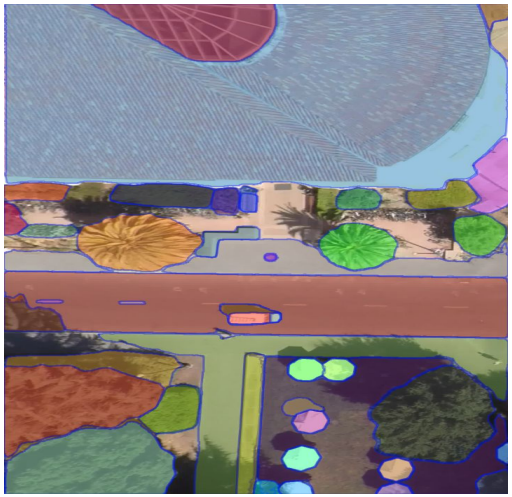
Test loss: 1.055034875869751

Results



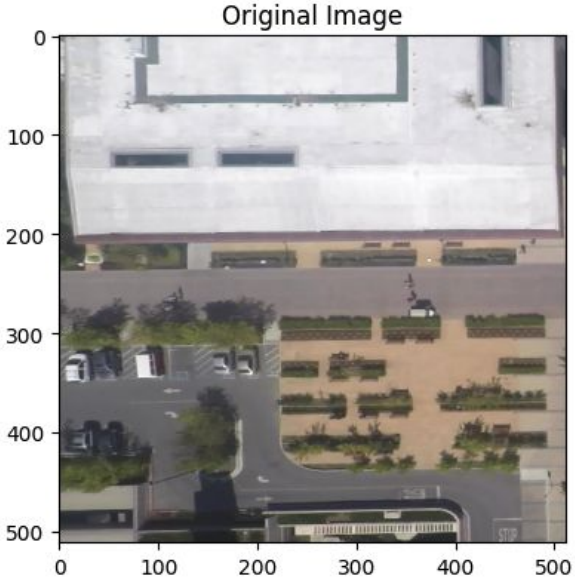


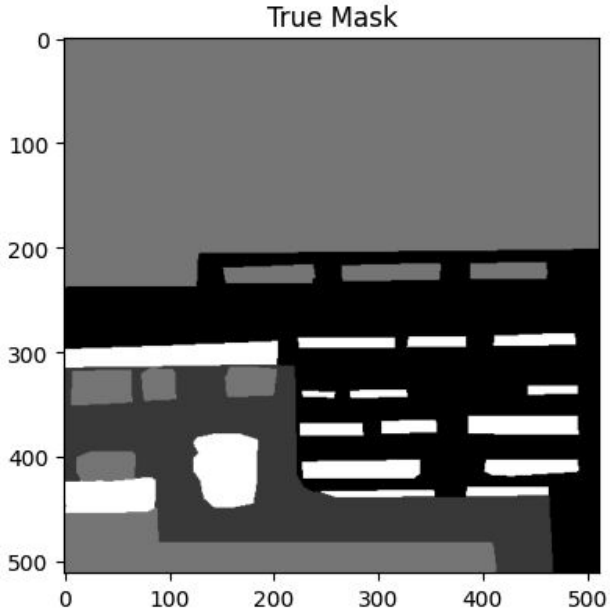


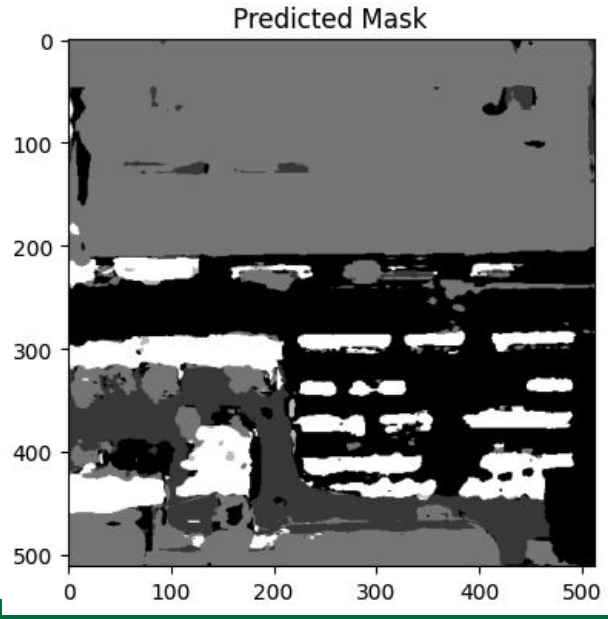


By Segment Anything

Results

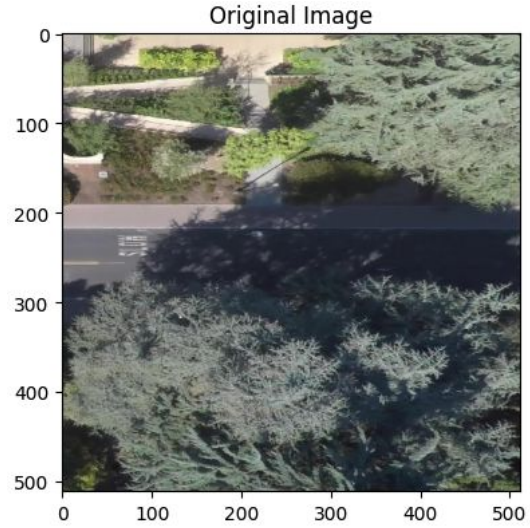


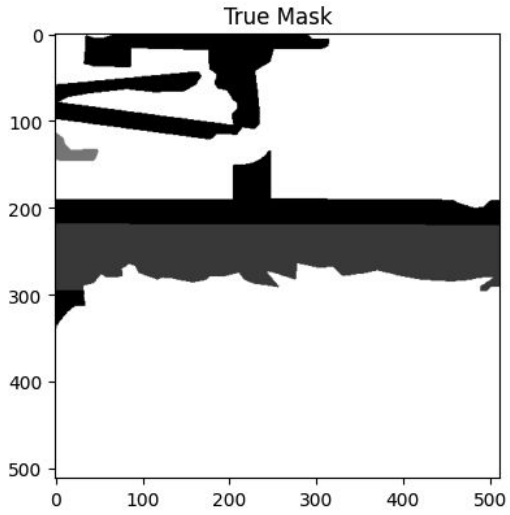


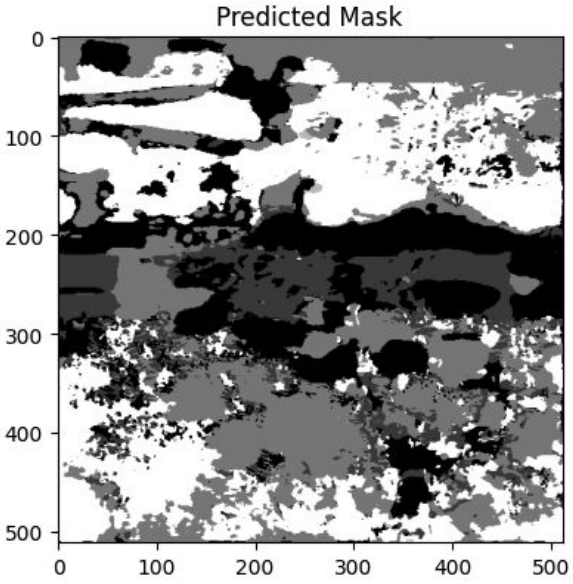


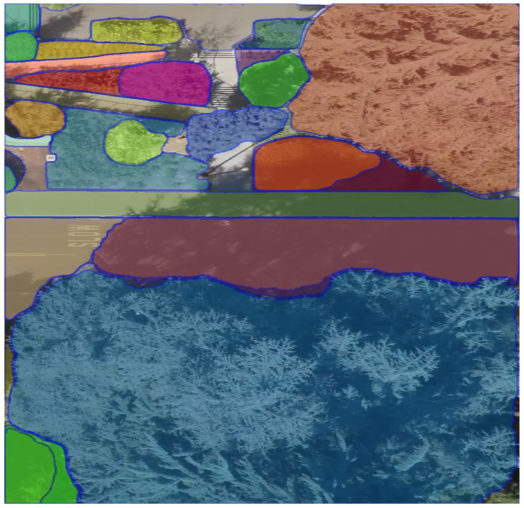
By Segment Anything

Results









By Segment Anything

Challenges and Future Work

Key challenges:

Hard to tell detail textures, such as trees and shadows.
Distinguishing between roofs and the ground in drone images.
Recognizing obscured walkable areas under trees or buildings.

Challenges and Future Work

Future Work:

Use transfer learning: Pretrain the model on a large dataset like ImageNet and fine-tune it on the specific task with the available smaller dataset.

References

[1] A. Robicquet, A. Sadeghian, A. Alahi, S. Savarese, "Learning Social Etiquette: Human Trajectory Prediction In Crowded Scenes," in European Conference on Computer Vision (ECCV), 2016.

[2] K. Mangalam, Y. An, H. Girase, J. Malik, "From Goals, Waypoints & Paths To Long Term Human Trajectory Forecasting," arXiv preprint arXiv:2012.01526, 2020.

[3] Kirillov, Alexander, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao et al. "Segment anything." arXiv preprint arXiv:2304.02643 (2023).

Thank You!