Siamese Network for Taxi Driver Trajectory Detection

Mohammad Nur Hossain Khan

1 Introduction

Meta-learning and few-shot learnings are important models for deep learning when the numbers of data are not enough to learn a traditional model. A few-shot learning model can learn the underlying features with a less number of data. In this project, we will use a Siamese network as a few-shot learning model to detect Taxi driver's ID from their trajectory. Unlike previous project, we have only 5 days of data for each driver. Siamese networks have been proven efficient when there is less number of data.

2 Methodology

2.1 Data Processing and Feature Generation

We divide the state into 25 x 48 size grid. We found the maximum longitude is 114.293, min longitude is 113.813, the minimum latitude is 22.502, maximum latitude is 22.751. After making the grid we integrate the trajectory of whole day without passengers and with passengers into the grid as two data for each of the five drivers. We concatenate each grid with a passenger and the grid without a passenger as two channels of one image. After that, we divided the data into positive samples and negative samples. Trajectories from same driver are attributed as positive samples and trajectories from different drivers are attributed as negative samples.

2.2 Network Architecture

The siamese network consists of two similar networks. Two data are passed through similar networks, and pairwise distance is calculated. Pairwise distance has been used to detect similarities between images. Convolutional neural network (CNN) has been used as the backbone of Simaese network. This backbone network consists of three CNN layer with 4, 8, and 8 output channels and three linear layers with 500, 500, and 5 output layers. Number of filters has been used is 3. Batch-normalization and ReLU activation has been used after each convolutional layer.

2.3 Training and validation

We generated 5000 images for training where 2500 were positive samples and 2500 were negative samples. The model was trained with contrastive loss and Adam optimizer with a learning rate of 0.0005. Among 5000 images, 80% were used as train and 20% data were used fro validation. Given validation set were used for testing the model.

3 Evaluation and Result

We found that a learning rate of 0.0005 works best. Also, the 128 batch size and a dropout of 0.4 gives the best performance. Different hyperparameter tested in this project is given later in this section.

3.1 Hyperparameters

- learning rates: Four different learning rates are used 0.001. 0.002, 0.0005, 0.0001
- loss function: Three different loss function was used in this project. Contrastive loss, binary cross entropy and Cross entropy
- batch size was varied from 32 to 128
- number of epoch was varied form 50 to 300
- Dropout was varied from 0.1 to 0.4

Table 1: Confusion Matrix

	Same Driver	Different Driver
Same Driver	72	28
Different Driver	29	21

Table 2: Comparison with different architecture

	Accuracy	F1-score	Precision	Recall
CNN v1 (proposed)	62	57	57	57
CNN v2	67	38	57	40

3.2 Training and validation result

We trained the model for 50 epochs with contrastive loss and adam optimizer. Our best-performing model achieves 62% accuracy on the validation set. The F1-score, precision and recall is 57%, 57%, 57% respectively. The confusion matrix is given at table 1.

3.3 Comparison

We compare the model with a different CNN encoder with different numbers of output layers (32, 64, 128 number of neurons). Although the accuracy of this model is high 67% than our reported model, the F1-score is low (38%). That means it cannot generalize the output better than our reported model. The comparison is given at table 2

4 Conclusion

In this project, we provide a network architecture to predict driver's ID from GPS data using few-shot learning. We got the highest accuracy of 62%. The model performance could be improved by using hand-crafted features as input. Also, data processing and feature generation can also be improved to get accurate results. LSTM model can be also useful which is not experimented here. Siamese networks are proven fruitful in a few shot learning. Other type of meta-learning can also be tested to get improved performance.