Solving the TSP with Convolutional Neural Networks

Student: Samuel J. K. Chin, SCM 2022
Advisor: Dr. Matthias Winkenbach

MIT Supply Chain MANAGEMENT

MIT GLOBAL SCALE NETWORK
Research Fest

Motivation / Background

The TSP was reportedly first posed by Hassler Whitney in 1934 at a seminar talk at Princeton University (Flood 1956) and the first mention of the TSP in writing is believed to be from Julia Robinson (Robinson 1949). The TSP has been formally studied since the 1930s and researchers are still constantly trying to push the boundaries of the TSP till today. There are two reasons for this. First, the TSP is classified as an NP-hard problem. This means that it is possible to obtain "good enough" approximations, but finding the true optimal is close to impossible for large problems of 100,000 cities and beyond. While an approximation is good enough in practice, the problem itself is interesting from a theoretical perspective. Second, at first glance, the TSP appears to be only relevant to logistics. Specifically, school bus routing, postal deliveries, Amazon Prime deliveries and many more. However, one will be surprised to find that the TSP is also relevant to genome sequencing, scan chains, drilling problems on circuit boards and many more (Applegate et al. 2006). In fact, the largest TSP ever solved is an 85,900 city optimal tour in a Very Large-Scale Integration (VLSI) application (Applegate et al. 2006). A small improvement in the TSP can lead to a very large impact across many different industries. This is why the TSP has been studied extensively over the past 90 years.

Key Question / Hypothesis

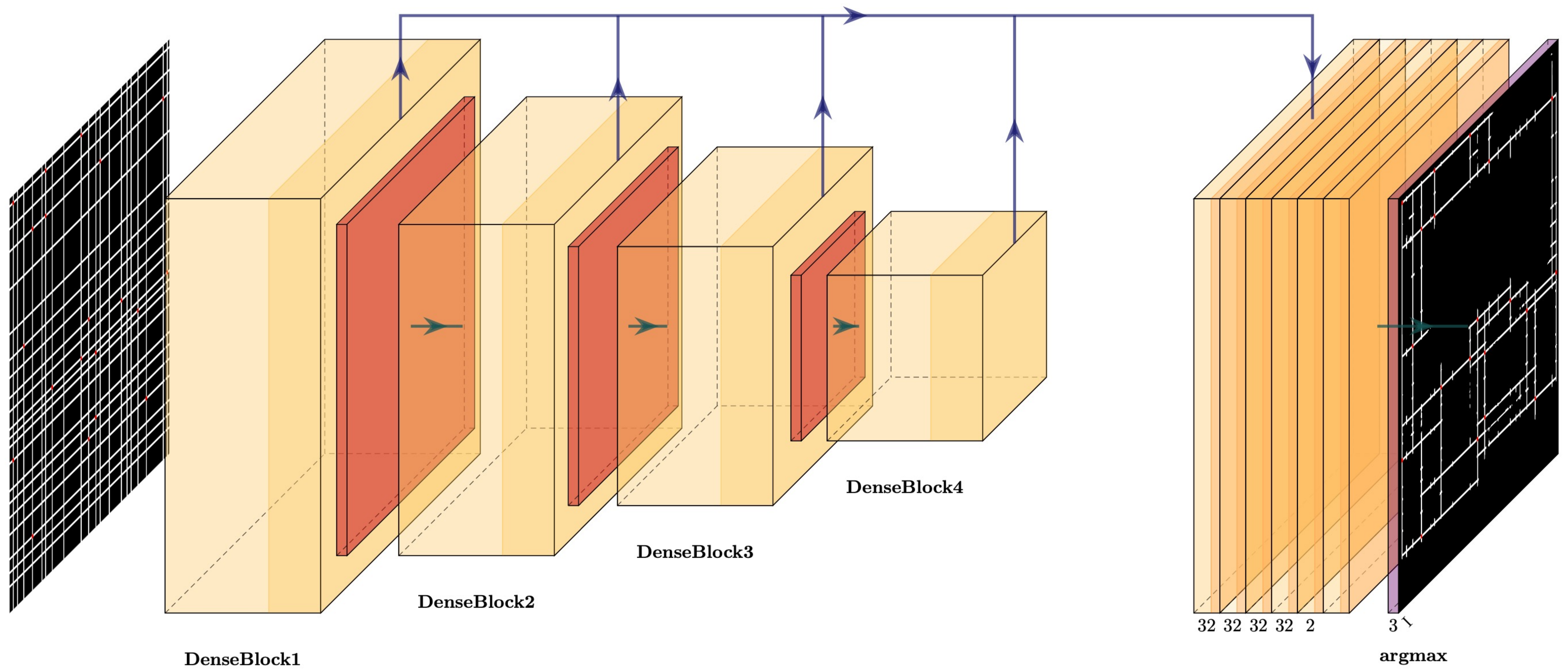
The focus of this thesis is on solving the TSP for transportation applications in a Manhattan Grid in the image domain.

Relevant Literature

Huang, Gao et al. (July 2017). "Densely Connected Convolutional Networks". In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). ISSN: 1063-6919, pp. 2261–2269. doi:10.1109/CVPR.2017.243.

Bengio, Yoshua, Andrea Lodi, and Antoine Prouvost (Mar. 12, 2020). "Machine Learning for Combinatorial Optimization: a Methodological Tour d'Horizon". In: arXiv:1811.06128 [cs, stat]. arXiv: 1811.06128. url: http://arxiv.org/abs/1811.06128

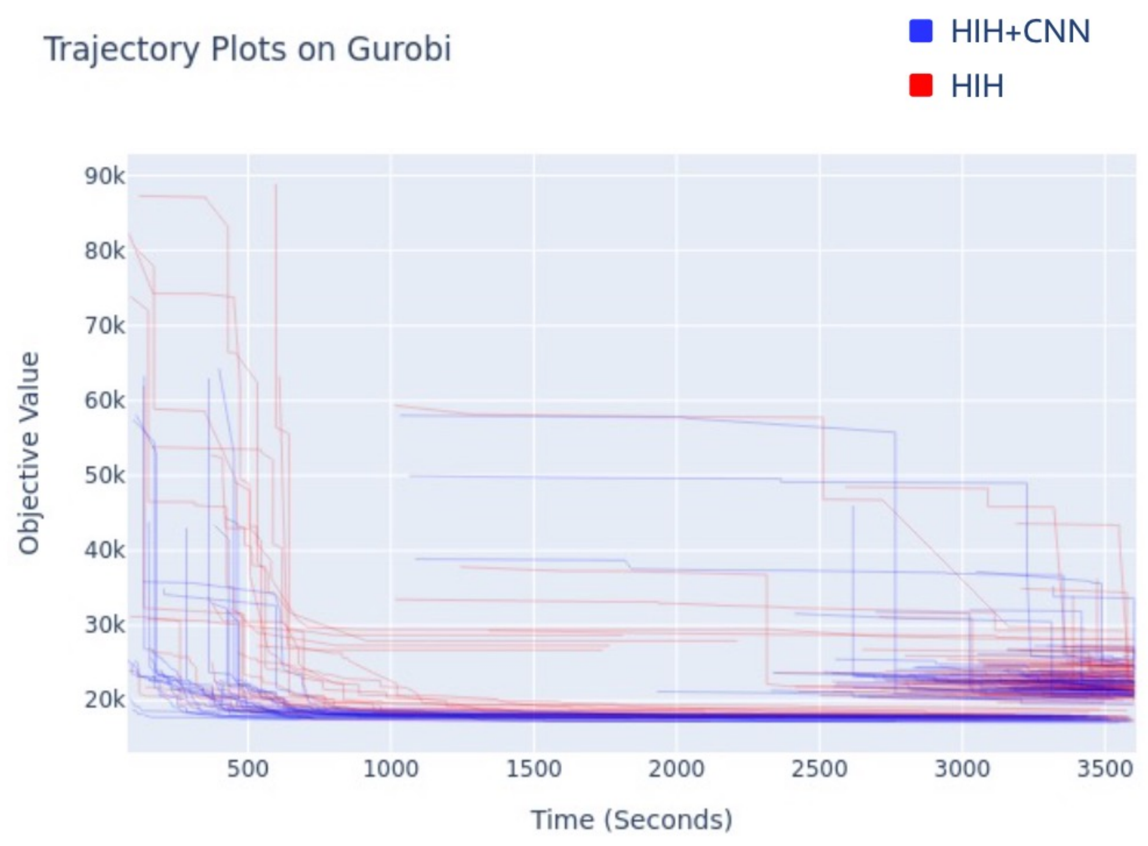
Applegate, David L. et al. (2006). "Computation". In: The Traveling Salesman Problem. A Computational Study. Princeton University Press, pp. 489–530. isbn: 978-0-691-12993-8. url: https://www.jstor.org/stable/j.ctt7s8xg.19



Methodology

We adapted the original DenseNet-161 architecture from Huang et al. 2017 and extract features from every DenseBlock. The features are then put through a BatchNorm, ReLU, 2D Convolution followed by a bilinear upsample. The transformed features are then concatenated in a final layer and then convolved one final time to produce values that can be put through an argmax to yield the semantic class of the specific pixel. The pruned semantic road network map is obtained and we apply our human inspired heuristic (HIH) to convert the semantic road network map into a truncated distance matrix, which we then pass to a solution algorithm to yield the final TSP solution. The models are trained with a sample size of 10,000 samples, pretrained weights from DenseNet-161 and 200 epochs. Cross Entropy Loss is used and the loss function is weighted in a ratio of [1, 1, 50], with the highest weightings on the paths.

Results



1. 1,000 examples from the 400 node instances were run using HIH+CNN and HIH on Gurobi with a time limit of 1 hour.
2. Approximately 100 instances each from HIH+CNN and HIH yielded tour lengths below 30K at the end of an hour. HIH+CNN shows faster convergence and lower average tour lengths.

Contributions

1. Our results show faster convergence and lower average tour length, indicating that the CNN was an effective pre-processing step.
2. We take a first step in the direction of potentially solving routing problems directly with CNNs.

Samuel J. K. Chin