**A Survey in Applications of Deep Learning in GIS – Spatiotemporal data mining and forecasting**

**Prepared by**

**Supasin Wuthikulphakdi**

**ID# 608785068**

**Tamkang University**

**June 1, 2021**

**A proposal submitted in partial fulfillment of the requirements for the Master Degree of Science in Computer Science**

**Proposed Supervisor: Isaac Tsai Yihjia, Faculty of Engineering**

**Proposed Reader(s) Faculty of Engineering – Department of Computer Science**

**Date of Submission**

**June 2, 2021**

**TITLE**

A Survey in Applications of Deep Learning in GIS – Spatiotemporal data mining and forecasting

**PROBLEM DESCRIPTION & JUSTIFICATION**

Spatiotemporal data is a dataset collected in the space domain (a.k.a. a map over a location), and the time domain (a.k.a. the time series). The need to S.T.-Data mining rises with the emerging use of intelligent transport system (ITS, i.e. automatic guidance, self-driving cars, traffic prediction, etc.). Several models based on both machine learning have been deployed, but ML models need a human to extract the features. Deep learning models, capable of self-feature extraction, usually outperform regular ML ones. However, DL models have drawbacks, a CNN/GCN can only extract and learn only the space domain, while an RNN/LSTM can only do these on the temporal domain of the S.T. data.

As the result some work referred, overcame the problems by adding such two neural networks into the same framework, allowing them to extract and learn the ST-data of both domains simultaneously.

**BACKGROUND AND HISTORY**

Spatiotemporal data is a dataset collected in the space domain (a.k.a. a map over a location), and the time domain (a.k.a. the time series) My thesis is focused on paper survey of overhead imagery processing & recognition using machine and deep learning algorithms and models, such as CNN (Convolutional Neural Network) for static images, and LSTM-RNN (Long Short-Term Memory RNN) for videos and time series data. With the AI, there are some uses in Spatiotemporal data mining:

* CNN-GCN is used to extract features in the spatial domain.
* LSTM-RNN is used to extract features in the temporal domain.
* In survey engineering, some researchers use CNN observe changes in land use, crop growth, and construction progresses.
* In ITS, ST-Data mining is extremely useful in traffic prediction, guidance systems, route planning, and self-driving cars.
* To make GPS based pathfinding more accurate, Spatiotemporal Data Mining is often utilized in several papers to train the pathfinding ML/DL-architectures.
* All the papers to be referred below, may be added & updated in the future, involve GIS, Computer Vision (CV), and spatiotemporal data mining.

Spatiotemporal data mining is a form of GIS data analysis done in both the spatial and the temporal domains. A CNN learns the spatial graph input, while and LSTM/RNN learns the temporal domain of the input data. With both DNN working together, we can forecast a spatiotemporal data (such as traffic, train ridership, etc.) not only the time series (time domain), but also the heatmap (space domain).

**LITERATURE REVIEW**

According to Doshi et al, 2018 [1], We propose to identify disaster-impacted areas by comparing the change in man-made features extracted from satellite imagery. Using a pre-trained semantic segmentation model from [2] we extract man-made features, the pre- and post-event images on the “before, during and after” imagery of the event-affected area.

According to Amit et al. [3], CNN is a sequence of layers, the convolution layer (who detects features from a data image), the pooling layer (downsamples the input), and the FC layer (who classifies the features detected earlier). with ReLU as the main activation function of the network. Further explained in the section 2 of [3].

Iglovikov et al. 2017 [4], used an FC-CNN named U-NET, along with an embedded multispectral sensor, which detects frequency reflection by the objects, to detect geo-features in satellite images and yielded satisfying results.

According to Bochkovskiy et al. [5], with the help of YOLO V.4, and TensorFlow Keras, CNN’s performance in image recognition is improved. So that it would be our main CNN model in this thesis. We hope that our deep learning model, written in Python, will work as the goals above.

In Terms of video and sequence type photos (such as slideshow), however, the use of LSTM-RNN is needed. According to Fang et al. [6], LSTM is excellent at predicting flood because it could process time series data.

According to Li et al, 2018. [7], spatiotemporal forecasting is a crucial task for a learning system that operates in a dynamic environment. It can be useful in pathfinding, autonomous vehicles, logistics, city planning etc. They used a Diffusion Convolutional Recurrent Neural Network (DCRNN) model to forecast the road traffic within a specific space and timeframe (The dataset was METR-LA, 2014). Diffusion convolution extracts the traffic features, and the RNN processes the traffic volumes in sequence.

Yu et al, 2018. [8] proposed a Spatiotemporal Graph Convolutional Networks (STGCN), to tackle the time series prediction problem in traffic domain. They formulated the problem on graphs and build the model with complete convolutional structures, enabling much faster training speed with fewer parameters. Compared with existing models, STGCN more effectively captured comprehensive spatiotemporal correlations through modeling multi-scale traffic networks and consistently outperforms state-of-the-art baselines on various real-world traffic datasets.

Correa et al, 2017 [9] performed a spatiotemporal data mining of Taxi vs Uber ridership in NYC, 2014+15. According to the heatmap inside the paper, the ridership for both taxi systems depended on several factors – such as personal income, education, jobs, car ownership etc. With 3 spatial models for ridership prediction – linear, spatial error and spatial lag models, the last one outperformed not only the first 2 algorithms, but also yielded a considerable accuracy and performance.

Amato et al. [10] designed a deep learning-based architecture called “Empirical Orthogonal Functions principal component analysis” EOF-PCA in which the EOF framework decomposes the spatiotemporal input data, in terms of a sum of products of temporally referenced basis functions and of stochastic spatial coefficients which can be spatially modelled and mapped on a regular grid. Then, the input layer spatial covariates are processed by a “Fully Connected Neural Network” (FCNN) to obtain predictive coefficient to be recomposed altogether with the decomposed data stream, obtaining a spatiotemporal signal reconstruction.

Tang et al. [11] designed an LSTM based framework to learn and forecast the rail traffic. The short-term forecast of rail transit is an issue in intelligent transportation system (ITS). Accurate forecast can forewarn travel outburst, helping the passengers with their travel plans. Even though the LSTM is notably effective in temporal data, it cannot correlate the time domain with the space domain. That is why we propose ST-LSTM. Compared with other conventional models, ST-LSTM network can achieve a better performance in experiments.

Lu et al. [12], designed a spatial-temporal deep learning network, termed ST-TrafficNet, for traffic flow forecasting, whose architecture works as follows. 1. The Spatial Aware Multi-Diffusion Convolution Bloc (ADC-Block – who introduces Graph Attention Mechanism (GAM) into the MDC) uncovers unseen spatial dependencies from traffic graph signals automatically 2. From the data stream, the multi-diffusion convolution (MDC) block harvests ST-features of the spatial domain. 3. The ST-TrafficNet, an LSTM based framework, harvests the features of the temporal domain. 4. The output from both ANN are summed up to achieve convolutional results. And 5. The ST-TrafficNet is evaluated on two benchmark datasets and compare it with various baseline methods for traffic forecasting.

**Potential Methods**

From each referred/surveyed paper, we will run the CNN+LSTM code written for each model and dataset. Then, we will run the models in the reference 1. With their own datasets 2. With the Taxi-Uber dataset and compare them in the form of MAE & RMSE metrics – the less of them for each model, the better it performs in the term of learning and forecasting. Finally, we will conclude which model does the best learning/forecast.

The Deep Learning models and ANNs are written in Python with Keras as our preferred library/module.

**Project Phases:**

* May/June <-(CURRENT PHASE)
  + Perform Literature Reviews, Python coding & other essential skills
* July
  + Perform even deeper LR + coding practice + Validation runs.
* August/September
  + Custom Model Design, Training & Testing
* October
  + Final Validation Runs
* November
  + Result & Analysis + Writing a paper.
* December
  + Finish the paper.
* January
  + Defence oral exam

**To be determined/confirmed by the supervisor:**

* Official Title
* Dataset for:
  + Training
  + Testing
  + Validation
    - Kaggle: Taxi vs Uber Data
* Models & Architectures
  + LSTM
  + No. & Types/Roles of hidden layers
  + Activation Functions
  + Etc.
* Quantitative Metrics
  + K1 Score
  + Accuracy
  + RMSE^2.

**Progresses:**

Current Phase: Literature Review and Essential Skill Training

**Works done:**

* Read 4 reference papers
* Run the taxi-uber LSTM time series forecasting model written in Keras
  + Model & coding are to be improved to extrapolate the future trends.

**References: (change to IEEE style)**

[1] Doshi, J., Basu, S. and Pang, G., 2018. From satellite imagery to disaster insights. *arXiv preprint arXiv:1812.07033*.

[2] Doshi, Jigar. 2018. Residual Inception Skip Network for Binary Segmentation. Pages 216–219 of:Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.  
  
[3] Amit, S.N.K.B. and Aoki, Y., 2017, September. Disaster detection from aerial imagery with convolutional neural network. In 2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC) (pp. 239-245). IEEE.

[4] V. Iglovikov, S. Mushinskiy, and V. Osin, “Satellite Imagery Feature Detection using Deep Convolutional Neural Network: A Kaggle Competition,” vol. June, 2017.

[5] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “Yolov4: Optimal speed and accuracy of object detection,” ArXiv, vol. abs/2004.10934, 2020.

[6] Fang, Z., Wang, Y., Peng, L. and Hong, H., 2020. Predicting flood susceptibility using LSTM neural networks. *Journal of Hydrology*, [online] p.125734. Available at: <https://doi.org/10.1016/j.jhydrol.2020.125734> [Accessed 18 April 2021].

[7] Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2017). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926*.

[8] Yu, B., Yin, H., & Zhu, Z. (2017). Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875*.

[9] Correa, D., Xie, K., & Ozbay, K. (2017). Exploring the taxi and Uber demand in New York City: An empirical analysis and spatial modeling. In *96th Annual Meeting of the Transportation Research Board, Washington, DC*.

[10] Amato, F., Guignard, F., Robert, S., & Kanevski, M. (2020). A novel framework for spatio‑temporal prediction of environmental data using deep learning. Scientific Reports, 2020(10), 22243. doi:<https://doi.org/10.1038/s41598-020-79148-7>

[11] Tang, Q., Yang, M., & Yang, Y. (2019). ST-LSTM: A Deep Learning Approach Combined Spatio-Temporal Features for Short-Term Forecast in Rail Transit. Journal of Advanced Transportation, 2019, Article ID 8392592. doi:<https://doi.org/10.1155/2019/8392592>

[12] Lu, H., Huang, D., Song, Y., Jiang, D., Zhou, T., & Qin, J. (2020). ST-TrafficNet: A Spatial-Temporal Deep Learning Network for Traffic Forecasting. Electronics, 2020(9), 1474.