Exploring deep learning methods for analyzing land use change

by

Cong Cao

M.Eng., Kyushu University, 2017

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Geography
Faculty of Environment

SIMON FRASER UNIVERSITY
Summer 2019

Approval

Name:	**Cong Cao**
Degree:	**Master of Science**
Title:	**Exploring deep learning methods for analyzing land use change**
Examining Committee:	**Chair:** Kirsten Zickfeld Associate Professor
	Suzana Dragicevic Senior Supervisor Professor
	Songnian Li Supervisor Professor Department of Civil Engineering Ryerson University
	Margaret Schmidt Internal Examiner Associate Professor
Date Defended/Approved:	July 3, 2019

Abstract

The process of land use change (LUC) results from human interactions with the natural environment to meet the needs from societal development. Growing population leads to the depletion of the land resource which entails environmental consequences from local to global scales. Advanced analytical methods can help with the understanding of the complexity of LUC process. They can further benefit sustainable land development. The main objective of this thesis research is to evaluate the deep learning (DL) methods such as convolutional neural networks (CNN) and recurrent neural networks (RNN) for classifying and forecasting LUC. The results demonstrated that the CNN-based LU classification models achieved the model accuracy of above 95%, while the RNN-based models for short-term LUC forecasting had 86% forecasting accuracy. This thesis contributes to advancing the methods for LUC analysis and improving the understanding of LUC process for sustainable land management.

Keywords: land use change; land use classification and forecasting; deep learning; convolution neural networks; recurrent neural networks; geographic information systems

Acknowledgments

I acknowledge the Natural Science and Engineering Research Council (NSERC) of Canada for support of this study under the Discovery Grant Program awarded to Dr. Suzana Dragicevic. And Department of Geography has kindly awarded me the Graduate Fellowship to encourage my research study.

I am grateful for the countless guidance, advice, and feedback from my supportive committee. Dr. Suzana Dragicevic has given me timely motivation, support and professional guidance in the past two years, without whom I cannot finish this master thesis research. In addition, Dr. Songnian Li has provided me with critical opinion during my thesis research which helped me structure and improve my work.

I also appreciate having met wonderful colleagues from the Department of Geography and the Spatial Analysis and Modeling Laboratory, who selflessly share their knowledge and provided me friendly support.

Table of Contents

List of Tables

List of Figures

List of Acronyms

CNN	Convolutional neural networks
DL	Deep learning
GIS	Geographic information system
LC	Land cover
LU	Land use
LUC	Land use change
LULCC	Land use and land cover change
RNN	Recurrent neural networks
RS	Remote sensing
SDGs	Sustainable Development Goals
SVM	Support vector machine

Chapter 1.

Introduction

1.1. Introduction

Land cover (LC) is comprehended to be defined as the physical material covering the earth surface such as water bodies and forest. Land use (LU) refers to the function or use of land for agriculture, recreation, transportation, or residence (Mukherjee et al., 2009). Both LU and LC are dynamic meaning they change over space and time. Land cover and land use changes (LULCC) arise from various human-environment interactions and have diverse impacts on the environment. LULCC is monitored through earth observation techniques (Loveland et al., 1999). Typically, LC types are classified based on properties measured with remote sensing (RS) techniques. LU are classified through field investigation or the interpretation of RS imageries. In recent time, modifications of the landscape for human use have led to significant land use change (LUC) (Foley et al., 2005). However, some of the LU practices have negative impacts on climate, biodiversity, and the ecosystem and have contributed to global warming, mass extinction, and environmental degradation (Fahrenkamp-Uppenbrink, 2013). In addition, the global population growth has created an unprecedented demand and pressure on the land and natural resources.

LUC has attracted the attention of researchers from various disciplines due to its correlation with environmental issues. To meet the demand of a growing population, LUC is inevitable during rapid urbanization and land conversions. Therefore, to maintain optimal benefit to environmental, social, and economic sectors, sustainable land development strategies are the core consideration in policy-making and land management processes. Effective regulation of LU can potentially diminish conflict between human and environment. Correspondingly, LUC studies need global attention and effort to better facilitate urban planning, environmental management, and regional administration (Gustafsson et al., 2018).

Studies of LUC have focused on the quantification and analysis of LUC from historical LU observations and the explanation and projection of land use evolution were

based on expert knowledge. In the literature, many LUC models have been proposed. Modeling methods are designed to simplify, transform and represent complex issues from the real world. (Epstein, 2008). LUC models can help land managers to gain an insight into the temporal and spatial dynamics of the landscape. The modification on the landscape by the human is one of the major causes of LUC. LUC is an indicator of the human footprint (Butt et al., 2015) and it demands to be monitored, analyzed and regulated, so that scientists and decision makers can conduct better land management.

Canada is known for its vast territory, rich and well-protected ecosystems, and natural resources. The favorable immigration policy of Canada has attracted many immigrants in recent decades. According to statistic Canada, from 2001 to 2018, the Canadian population increased from 31.02 million to 37.06 million inhabitants, a growth rate of 19.5%, which has led to increased demand for built-up land and resources (Statistics Canada, 2019). In such circumstances, strategies for sustainable development are emphasized by policies. Sustainable development promotes the balance between meeting human development needs and maintaining the natural system so that the environment can provide resources and energy to human for longer time (Lélé, 1991). Proposed by the United Nations General Assembly, one of the objectives of the Sustainable Development Goals (SDGs) (Rosa, 2017) is to make cities and human settlements inclusive, safe, resilient and sustainable under the circumstance of rapid urbanization. Hence, promoting effective land management policies is one of the key aspects needed to reach the SDGs.

Remote sensing (RS) and geographic information system (GIS) techniques are typically applied to geographic studies (Butt et al., 2015; Chen et al., 2018; Hegazy and Kaloop, 2015). In-situ measurements combined with ground surveys and satellite data are the most precise methodological options for LULC studies. Nowadays, GIS is the mainstream platform for analyzing, processing and sharing geospatial data to solve complex environmental problems (Weng, 2002). RS techniques provide massive earth information data with high resolution and easier access. RS data provide direct, real-time, global earth information, and they are the primary sources for studying LULC (Fritz et al., 2017). Lately, GIS has started to incorporate artificial intelligence (AI) for the tasks of segmenting, classifying and predicting geographical data.

Artificial intelligence (AI) (Russell & Norvig, 1994) can be defined as the intelligence demonstrated by machines or computers that can resolve complex tasks instead of humans. Machine learning (ML) (Bishop, 2006) is encompassing the suite of methods commonly used in AI. The relatively new field, deep learning (DL) (LeCun, Bengio, & Hinton, 2015), has been created as subfield of ML due to the need to augment the capacity and the performance of the ML methods and AI in general. Their relationship can be presented as Figure 1-1. The popularity of DL models has increased due to the availability of larger data, more powerful GPUs, and the advancement of the fields of computer vision (CV) and natural language processing (NLP). DL models automatically and simultaneously complete feature extraction and classification, processes implemented separately with ML methods. LUC is a complicated phenomenon as it originates from complex human-environment interactions. Due to the limitation of expert knowledge and insufficient geospatial datasets, it is necessary to explore the DL models such as convolutional neural network (CNN) (Fukushima, 1980) and recurrent neural network (RNN) (Rumelhart et al., 1986) to classify and forecast LUC, which will help with better understanding and management of LUs.

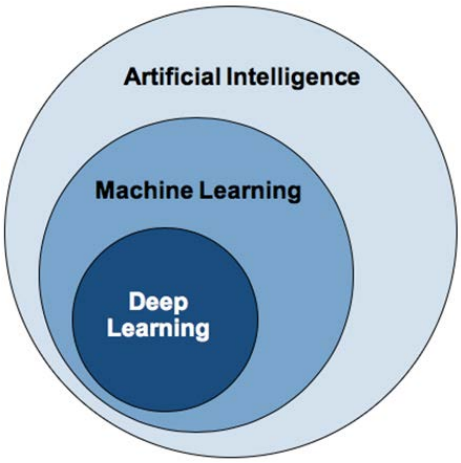


Figure 1-1 The relationship of artificial intelligence, machine learning, and deep learning (Wasicek A., 2018)

1.2. Research Problems

Many approaches have been used for modeling LUC processes in the literature, such as Markov chain (Kumar et al., 2014; Muller and Middleton, 1994), cellular automata (CA) (Batty and Xie, 1994; Charif et al., 2017; Clarke et al., 1997; Clarke and Gaydos, 1998; White and Engelen, 1997; Wu and Webster, 1998), and machine learning (ML) (Hernandez and Shi, 2018; Huang et al., 2009; Nemmour and Chibani, 2006; Otukei and Blaschke, 2010; Samardzic-Petrovic et al., 2016).

To help understand the potential consequences of current land development strategies and to plan for specific urban development needs, this thesis has been focused on study of the LUC using DL models and provided reliable results revealing LUC information. More specifically, the results of this thesis can provide guidance to land managers to avoid unsustainable LU development and choose the optimal location for using certain land types.

In the literature DL based models, have been proposed for LULC classification (Cevikalp et al., 2017; Cheng et al., 2016; Han et al., 2017; Liu et al., 2018; Othman et al., 2016; Qayyum et al., 2017; Romero et al., 2016; Weng et al., 2017; Zhang et al., 2016). Some studies have successfully used trained CNN models for the classification of urban LU with an accuracy of 88% through jointing multilayer perceptron (MLP) and CNN models (Zhang et al., 2019). Furthermore, Zhang et al. (2018) managed the challenges of partitioning LU objects from study sites through proposed object-based CNN (OCNN). Huang, Zhao, & Song (2018) proposed a skeleton-based decomposition method to decide the mapping unit and then performed CNN on each unit. Some studies have used RNNs to classify LC and modeling vegetation (Rußwurm & Körner, 2017, 2018). Recently, RNNs models have been used to forecast weather (Qing & Niu, 2018) and traffic flow (Han et al., 2019), however they were not yet fully exploited for LUC forecasting.

Based on the potential of application of the DL methods in various fields, further exploration of DL models, particularly CNN and RNN, to classify, analyze and forecast the LUC should be investigated. DL models can help improve the efficiency of classification methods for LU studies and based on availability of reliable LU data. Accordingly, two major research questions are raised in this thesis:

1. How do variants of CNN models perform when used on the LU classification tasks and what would be the optimal composition of the training datasets?

2. How effective are the RNN models when implemented to forecast LUC?

1.3. Research Questions and Objectives

The following research objectives are set to answer the outlined research questions:

1. To evaluate several CNN-based models for LUC classification using geospatial datasets and using different training datasets.

2. To evaluate several RNN-based models for forecasting LUC in short term intervals using geospatial datasets.

1.4. Study area and datasets

The proposed methodology from this thesis is implemented using the datasets for several time intervals and case study of land-use change for City of Surrey, British Columbia, Canada know by its fast dynamics of urban growth. The City of Surrey is located in the southern region of (Figure 1-2) with he population that has increased by 10.6% from 2011 to 2016 (Statistics Canada, 2017). The population is expected to continue to grow and even surpass the one in the City of Vancouver. Therefore, the number of residential areas has been developed to accommodate the increased influx of population and to create vibrant neighborhoods where people can live, work, and grow their families (City of Surrey, 2019b). Moreover, The Community of Cloverdale, located in the eastern region of the City of Surrey, has been also chosen to perform the analysis. Starting from the late 1800s, Cloverdale has attracted many young families and successfully transformed itself from a rural agricultural community to prosperous small community.

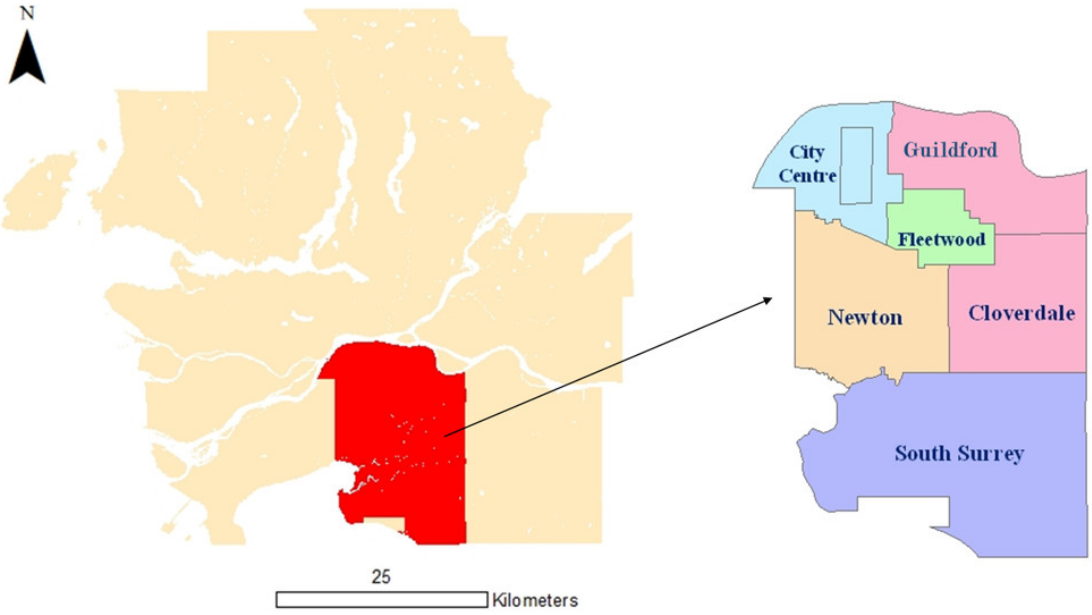


Figure 1-2 Study Area: The City of Surrey, Canada with the community of Cloverdale

In chapter 2, datasets for the community of Cloverdale were chosen to implement the prosed methodology. In this chapter, labeled RS data was used to train CNN-based models and the trained models were applied on the RS data of study area to extract LU information. The training data was obtained from the well-known open source UC-Merced dataset (Yang and Newsam, 2010). It includes high-resolution classified LU images composed of 21 LU types: agricultural, airplane, baseball diamond, beach, buildings, chaparral, dense residential, forest, freeway, golf course, harbor, intersection, medium residential, mobile home park, overpass, parking lot, river, runway, sparse residential, storage tanks, tennis court. The images in UC-Merced dataset were manually extracted from the USGS National Map Urban Area Imagery collection for various urban areas around the country. In this study, the orthophotos of the community of Cloverdale in the past few years were classified by CNNs to obtain LU information of Cloverdale, to provide timely information about local land development. The orthophotos used for classification were obtained from the City of Surrey Open Data Catalog (City of Surrey, 2017) with the resolution of 10cm and covering the size is 2 km × 3.6 km of the study area.

In chapter 3, datasets for the City of Surrey was chosen as larger study area to implement the methodology. To train the RNN models, LU data for the years 1996, 2001, 2006 and 2011 (Figure 1-3) were obtained from Metro Vancouver Open Data Catalog (Metro Vancouver, 2011). The road network shapefiles of the year 1996 and 2001 were obtained from CanMap database (Qing & Niu, 2018) and were used to supplement the missing road network information in LU data of the year 1996 and 2001. The LU data and road network shapefiles were transformed into raster files and the resolution of each raster cell was 10m. Then LUC for the year 2016 was predicted by the trained RNNs.

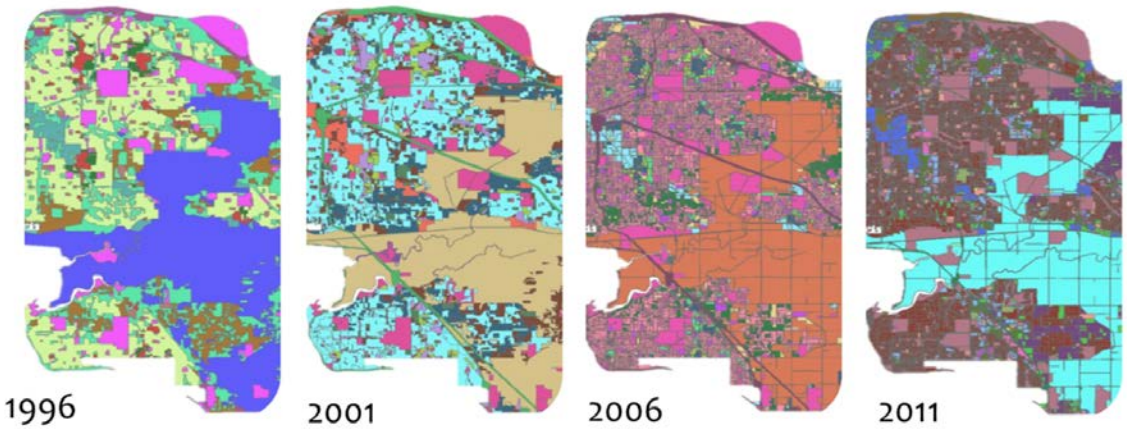


Figure 1-3 Classified LU data for the years 1996, 2001, 2006 and 2011

1.5. Thesis overview

After the introductory chapter, the thesis contains three more chapters. Chapter 2 presents the testing of eight CNN-based models using three pre-trained CNN models, AlexNet, GoogLeNet, and VGGNet and applies the best performing CNN-based model to LUC classification. The best-performed method was GoogLeNet model combined with support vector machine (SVM) as feature classifier (CNN-SVM). Two sources of datasets, UC-Merced dataset and manually sampled LU images from the City of Surrey were used to train the CNN–SVM model to solve a practical issue with the limited data availability. The overall accuracy of re-trained CNN based models was over 95-98%. The transferred CNN-SVM model was then applied to orthophotos of the northeastern Cloverdale as part of the City of Surrey, Canada from 2004 to 2017 to perform LU classification and LU change analysis. The orthophoto images were divided into patches

with size of 70m × 70m, the size was chosen to match the size of training images in UC-Merced dataset. Then the patches were classified by CNN-SVM to obtain the pattern of LUC.

Chapter 3 examines variants of the RNN models to forecast LUC in short-term interval. Historical land use data for the City of Surrey, Canada was used to implement the several variants of the RNN models. The land use data for years 1996, 2001, 2006 and 2011 were used for training the DL models to enable the short-term forecast for the year 2016. Four variants of RNN, LSTM, BiLSTM, GRU, ConvLSTM were applied and evaluated. Two groups of datasets were used for training, where one group only contained data from areas that have changed, and the other group contained all the data from the study area. The variants of RNN models were trained by all raster cells in the study area had 87% model accuracy, higher than the models trained by only changed raster cells which is 62%. Then LSTM trained by all raster cells was used for forecasting the LU change of the City of Surrey for the year 2016. Then the forecasted land use has been compared with 2016 orthophoto image of the City of Surrey has been performed by taking 604 random sample points.

Chapter 4 presents the conclusions of this thesis with the summary of the. It discusses the limitations of the used DL-based models, and it also provides reflections on the avenues for future work, and overall thesis contributions.

1.6. References

Batty, M., & Xie, Y. (1994). From Cells to Cities. *Environment and Planning B: Planning and Design*, *21*(7), S31–S48. https://doi.org/10.1068/b21S031

Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.

Butt, A., Shabbir, R., Ahmad, S. S., & Aziz, N. (2015). Land use change mapping and analysis using Remote Sensing and GIS: A case study of Simly watershed, Islamabad, Pakistan. *The Egyptian Journal of Remote Sensing and Space Science*, *18*(2), 251–259. https://doi.org/10.1016/j.ejrs.2015.07.003

Cevikalp, H., Dordinejad, G. G., & Elmas, M. (2017). Feature Extraction with Convolutional Neural Networks for Aerial Image Retrieval. In *2017 25th Signal Processing and Communications Applications Conference (siu)*. New York: Ieee.

Charif, O., Omrani, H., Abdallah, F., & Pijanowski, B. (2017). A multi-label cellular automata model for land change simulation. *Transactions in Gis*, *21*(6), 1298–1320. https://doi.org/10.1111/tgis.12280

Chen, liping, Sun, Y., & Sajjad, S. (2018). Monitoring and predicting land use and land cover changes using remote sensing and GIS techniques—A case study of a hilly area, Jiangle, China. *PLOS ONE*, *13*(7), e0200493. https://doi.org/10.1371/journal.pone.0200493

Cheng, G., Ma, C., Zhou, P., Yao, X., & Han, J. (2016, July 1). *Scene classification of high resolution remote sensing images using convolutional neural networks*. 767–770. https://doi.org/10.1109/IGARSS.2016.7729193

City of Surrey. (2017). Imagery—City of Surrey Open Data Catalogue. Retrieved September 10, 2018, from http://data.surrey.ca/group/6878e307-9fec-4134-b042-d7e058310255?tags=orthophoto

City of Surrey. (2019a). City of Surrey [Text/xml]. Retrieved May 17, 2019, from http://www.surrey.ca/default.aspx

City of Surrey. (2019b). Planning & Development [Text/xml]. Retrieved May 23, 2019, from http://www.surrey.ca/city-government/645.aspx

Clarke, K C, Hoppen, S., & Gaydos, L. (1997). A Self-Modifying Cellular Automaton Model of Historical Urbanization in the San Francisco Bay Area. *Environment and Planning B: Planning and Design*, *24*(2), 247–261. https://doi.org/10.1068/b240247

Clarke, Keith C., & Gaydos, L. J. (1998). Loose-coupling a cellular automaton model and GIS: Long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science*, *12*(7), 699–714. https://doi.org/10.1080/136588198241617

DMTI Spatial Inc. (2011). CanMap Streetfiles. Retrieved May 24, 2019, from https://www.dmtispatial.com/canmap/

Epstein, J. M. (2008, October 31). Why Model? [Text.Article]. Retrieved May 19, 2019, from http://jasss.soc.surrey.ac.uk/11/4/12.html.bak

Fahrenkamp-Uppenbrink, J. (2013). The Effects of Land-Use Change. *Science*, *339*(6126), 1360–1360. https://doi.org/10.1126/science.339.6126.1360-b

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., … Snyder, P. K. (2005). Global Consequences of Land Use. *Science*, *309*(5734), 570–574. https://doi.org/10.1126/science.1111772

Fritz, S., See, L., Perger, C., McCallum, I., Schill, C., Schepaschenko, D., … Obersteiner, M. (2017). A global dataset of crowdsourced land cover and land use reference data. *Scientific Data*, *4*, 170075. https://doi.org/10.1038/sdata.2017.75

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, *36*(4), 193–202. https://doi.org/10.1007/BF00344251

Gustafsson, S., Hermelin, B., & Smas, L. (2018). Integrating environmental sustainability into strategic spatial planning: The importance of management. *Journal of Environmental Planning and Management*, *0*(0), 1–18. https://doi.org/10.1080/09640568.2018.1495620

Han, X., Zhong, Y., Zhao, B., & Zhang, L. (2017). Scene classification based on a hierarchical convolutional sparse auto-encoder for high spatial resolution imagery. *International Journal of Remote Sensing*, *38*(2), 514–536. https://doi.org/10.1080/01431161.2016.1266059

Han, Y., Wang, C., Ren, Y., Wang, S., Zheng, H., & Chen, G. (2019). Short-Term Prediction of Bus Passenger Flow Based on a Hybrid Optimized LSTM Network. *ISPRS International Journal of Geo-Information*, *8*(9), 366. https://doi.org/10.3390/ijgi8090366

Hegazy, I. R., & Kaloop, M. R. (2015). Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt. *International Journal of Sustainable Built Environment*, *4*(1), 117–124. https://doi.org/10.1016/j.ijsbe.2015.02.005

Hernandez, I. E. R., & Shi, W. (2018). A Random Forests classification method for urban land-use mapping integrating spatial metrics and texture analysis. *International Journal of Remote Sensing*, *39*(4), 1175–1198. https://doi.org/10.1080/01431161.2017.1395968

Huang, B., Xie, C., Tay, R., & Wu, B. (2009). Land-use-change modeling using unbalanced support-vector machines. *Environment and Planning B-Planning & Design*, *36*(3), 398–416. https://doi.org/10.1068/b33047

Huang, B., Zhao, B., & Song, Y. (2018). Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sensing of Environment*, (214), 73–86.

Kumar, S., Radhakrishnan, N., & Mathew, S. (2014). Land use change modelling using a Markov model and remote sensing. *Geomatics, Natural Hazards and Risk*, *5*(2), 145–156. https://doi.org/10.1080/19475705.2013.795502

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lélé, S. M. (1991). Sustainable development: A critical review. *World Development*, *19*(6), 607–621. https://doi.org/10.1016/0305-750X(91)90197-P

Liu, Y., Zhong, Y., Fei, F., Zhu, Q., Qin, Q., Liu, Y., … Qin, Q. (2018). Scene Classification Based on a Deep Random-Scale Stretched Convolutional Neural Network. *Remote Sensing*, *10*(3), 444. https://doi.org/10.3390/rs10030444

Loveland, T. R., Zhu, Z., Ohlen, D. O., Brown, J. F., Reed, B. C., & Yang, L. (1999). An analysis of IGBP global land-cover characterization process. *Photogrammetric Engineering and Remote Sensing*, *65*(9), 10211032.

Metro Vancouver. (2011). Open Data Catalogue [Government]. Retrieved December 21, 2018, from Metro Vancouver website: http://www.metrovancouver.org/data

Mukherjee, S., Shashtri, S., Singh, C. K., Srivastava, P. K., & Gupta, M. (2009). Effect of canal on land use/land cover using remote sensing and GIS. *Journal of the Indian Society of Remote Sensing*, *37*(3), 527–537. https://doi.org/10.1007/s12524-009-0042-6

Muller, M. R., & Middleton, J. (1994). A Markov model of land-use change dynamics in the Niagara Region, Ontario, Canada. *Landscape Ecology*, *9*(2), 151–157. https://doi.org/10.1007/BF00124382

Nemmour, H., & Chibani, Y. (2006). Multiple support vector machines for land cover change detection: An application for mapping urban extensions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *61*(2), 125–133. https://doi.org/10.1016/j.isprsjprs.2006.09.004

Othman, E., Bazi, Y., Alajlan, N., Alhichri, H., & Melgani, F. (2016). Using convolutional features and a sparse autoencoder for land-use scene classification. *International Journal of Remote Sensing*, *37*(10), 2149–2167. https://doi.org/10.1080/01431161.2016.1171928

Otukei, J. R., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, *12*, S27–S31. https://doi.org/10.1016/j.jag.2009.11.002

Qayyum, A., Malik, A. S., Saad, N. M., Iqbal, M., Abdullah, M. F., Rasheed, W., … Jafaar, M. Y. B. (2017). Scene classification for aerial images based on CNN using sparse coding technique. *International Journal of Remote Sensing*, *38*(8–10), 2662–2685. https://doi.org/10.1080/01431161.2017.1296206

Qing, X., & Niu, Y. (2018). Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, *148*, 461–468. https://doi.org/10.1016/j.energy.2018.01.177

Romero, A., Gatta, C., & Camps-Valls, G. (2016). Unsupervised Deep Feature Extraction for Remote Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(3), 1349–1362. https://doi.org/10.1109/TGRS.2015.2478379

Rosa, W. (Ed.). (2017). Transforming Our World: The 2030 Agenda for Sustainable Development. In *A New Era in Global Health*. https://doi.org/10.1891/9780826190123.ap02

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533. https://doi.org/10.1038/323533a0

Russell, S., & Norvig, P. (1994). *Artificial Intelligence: A Modern Approach* (Third edition). Retrieved from http://aima.cs.berkeley.edu/

Rußwurm, M., & Körner, M. (2017, July 21). *Temporal Vegetation Modelling using Long Short-Term Memory Networks for Crop Identification from Medium-Resolution Multi-Spectral Satellite Images*. https://doi.org/10.1109/CVPRW.2017.193

Rußwurm, M., & Körner, M. (2018). Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders. *ISPRS International Journal of Geo-Information*, *7*. https://doi.org/10.3390/ijgi7040129

Samardzic-Petrovic, M., Dragicevic, S., Kovacevic, M., & Bajat, B. (2016). Modeling Urban Land Use Changes Using Support Vector Machines. *Transactions in Gis*, *20*(5), 718–734. https://doi.org/10.1111/tgis.12174

Statistics Canada. (2017, February 8). Census Profile, 2016 Census. Retrieved May 24, 2019, from https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E

Statistics Canada. (2019). Population and demography. Retrieved May 24, 2019, from https://www150.statcan.gc.ca/n1/en/subjects/population_and_demography

Wasicek, A. (2018). Artificial Intelligence vs. Machine Learning vs. Deep Learning: What's the Difference? Retrieved August 22, 2019, from Sumo Logic website: https://www.sumologic.com/blog/machine-learning-deep-learning/

Weng, Qian, Mao, Z., Lin, J., & Guo, W. (2017). Land-Use Classification via Extreme Learning Classifier Based on Deep Convolutional Features. *Ieee Geoscience and Remote Sensing Letters*, *14*(5), 704–708. https://doi.org/10.1109/LGRS.2017.2672643

Weng, Qihao. (2002). Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. *Journal of Environmental Management*, *64*(3), 273–284. https://doi.org/10.1006/jema.2001.0509

White, R., & Engelen, G. (1997). Cellular Automata as the Basis of Integrated Dynamic Regional Modelling. *Environment and Planning B: Planning and Design*, *24*(2), 235–246. https://doi.org/10.1068/b240235

Wu, F., & Webster, C. J. (1998). Simulation of Land Development through the Integration of Cellular Automata and Multicriteria Evaluation. *Environment and Planning B: Planning and Design*, *25*(1), 103–126. https://doi.org/10.1068/b250103

Yang, Y., & Newsam, S. (2010). Bag-of-visual-words and Spatial Extensions for Land-use Classification. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 270–279. https://doi.org/10.1145/1869790.1869829

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., & Atkinson, P. M. (2018). An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sensing of Environment*, *216*, 57–70. https://doi.org/10.1016/j.rse.2018.06.034

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., & Atkinson, P. M. (2019). Joint Deep Learning for land cover and land use classification. *Remote Sensing of Environment*, *221*, 173–187. https://doi.org/10.1016/j.rse.2018.11.014

Zhang, F., Du, B., & Zhang, L. (2016). Scene Classification via a Gradient Boosting Random Convolutional Network Framework. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(3), 1793–1802. https://doi.org/10.1109/TGRS.2015.2488681

Chapter 2.

Land use change detection using convolutional neural network-methods[1]

2.1. Abstract

Convolutional neural networks (CNN) have been used increasingly in several land use classification tasks, but there is a need to further investigate its potential since there is limited discussions of the selection of training data and the application of CNNs on real study areas. This study aims to evaluate the performance of CNN models for land classification and to identify land use change (LUC). Eight transferred CNN-based models were trained and their testing accuracy was evaluated on remote sensing data for LU scene classification using three pre-trained CNN models AlexNet, GoogLeNet, and VGGNet. The testing accuracy of eight CNN-based models ranges from 95% to 98%. A best-performed transferred CNN-SVM model was then applied to orthophotos of the northeastern Cloverdale as part of the City of Surrey, Canada from 2004 to 2017 to perform LU classification and LUC analysis. Two sources of datasets were used to train the CNN-SVM model to solve a practical issue with limited data. The obtained results indicate that residential areas were expanding by creating a higher density, while green areas and low-density residential areas were decreasing over the years, which accurately indicates the trend of LUC in the community of Cloverdale study area.

2.2. Introduction

Land use (LU) represents the human use of the natural environment for economic, urban, recreational, conservational, and governmental purposes. Land cover (LC) represents how a region of the earth surface is covered by physical features such as vegetation, water, forest or other (Zin and Lin, 2018). In modern history, the worldwide growth of population has brought significant challenges to society and the environment, such as the increasing demand for housing, food, natural resources, and

[1] A version of this chapter has been published as Cao C., Dragicevic, S., and Li S., (2019). Land use change detection with convolutional neural networks methods. *Environments*. 6(2), 25

basic services (Progress towards the Sustainable Development Goals, 2017). These have resulted in land use and land cover (LULC) changes that have caused adverse effects on the natural environments (Foley et al., 2005). The need for efficient land use planning and management is increasing, not only to eliminate the negative effects of historical LU decisions but also to make future communities healthier and more sustainable (Etingoff, 2017). Therefore, there is a need for more advanced computational methods to analyse geospatial data from the Earth surface to quantify and better understand the complex dynamics of LULC change processes.

Remote sensing (RS) datasets provide significant information documenting the land use and land cover processes (Ienco et al., 2017). RS datasets provide coverages from regional to global scales (Joseph, 2005). Interpretation of RS datasets is a major way to understand the status and changes in both the natural and built environments. In recent decades, RS sensors and techniques have become increasingly sophisticated. They can provide a large volume of datasets with high quality and fine spatial resolution.

Due to easier access to data with higher volume and better quality as well as the development of advanced graphics processing units (GPU), deep learning (DL) has been widely promoted in many recent scientific literature. DL consists of a collection of algorithms as the subset of machine learning (ML), specializing in learning hierarchy of concepts from very big data. Application of DL models have achieved high accuracy especially in speech recognition, computer vision, video analysis, and natural language processing (Arel et al., 2010; W. Liu et al., 2017). Some of the well-known DL models include CNN (Krizhevsky et al., 2012a; LeCun et al., 1989), recurrent neural networks (RNN) (Hochreiter and Schmidhuber, 1997a), deep belief networks (DBN) (Hinton et al., 2006), and stacked auto-encoders (SAE) (Vincent et al., 2010). This study will use CNN models to classify LU.

CNN models have many well-known applications such as face recognition (Lawrence et al., 1997), modeling sentences (Kalchbrenner et al., 2014) and image classification (Krizhevsky et al., 2012a). CNN can identify the health state of rice crops (Lu et al., 2017), or evaluate the degree of building damage (Rashedi Nia, 2017). It can also identify land use from online geo-tagged photos (Xu et al., 2017). As for large-scale images, Isikdogan et al. (2017) used CNN to identify water features from multispectral Landsat imagery and to map water surfaces. These studies identify images with a minor

abnormality and deal with few categories. Identifying LU requires the recognition of several predefined categories thereby increasing the complexity of the classification. Some studies that have configured and evaluated CNN models to classify LUC by training with LU RS dataset (Luus et al., 2015; Scott et al., 2017). Luus et al. (2015) proposed a multiscale LU RS images and designed CNN structures that accept multiscale training images, and their multi-view CNN model achieved total accuracy of 93.48%. Scott et al. (2017) achieved 98.5% total accuracy with retrained CNN structure, ResNet. All these studies used the well-known UC-Merced (Yang and Newsam, 2010) LU image dataset to train the CNN model to identify LU. While most of these studies elaborated the structure of classification models to reach the highest accuracy, their application and performance on real study area have started to be addressed in the scientific literature (Huang, Zhao, & Song, 2018; Zhang et al., 2019).

The typical strategy for identifying land use is aerial scene classification (Cheng et al., 2015; Hu et al., 2015; Liu et al., 2018), which automatically labels an aerial image with predefined semantic categories (Nogueira et al., 2017a). A land use “scene” is a unit that contains several features and demonstrates a unique scenario such as harbor, residence, highway, agricultural land, and park. Several entities exist in the same scene with a complicated layout, making it difficult to identify the scene. Image representation is the key technique in the scene classification task (Hu et al., 2015), i.e., the extraction of core features that represent the original images. Features are divided into low-level and high-level according to the complexity. The low-level features contain minor details of the images such as line, dot, curve edges, gradients, and corners. High-level features are built upon low-level features forming larger shapes or objects. An often used method for scene classification, for example, is bag-of-visual-words (BoVW) (Yang and Newsam, 2010), which utilizes the frequency of low-level features in a scene, but cannot describe the relative location of entities in a scene. Deep learning mode can study low and high-level features and the relative location of features in an image (Hughes et al., 2018; Krizhevsky et al., 2012; Liu et al., 2016; Simonyan and Zisserman, 2014; Yao et al., 2017; Zhu et al., 2017). Land use maps are created by identifying the characteristic scenes from RS imagery (Hu et al., 2015; Nogueira et al., 2017; Scott et al., 2017; Xia et al., 2017).

CNN has become a prevalent model in land use scene classification (Nogueira et al., 2017b; Oquab et al., 2014; Romero et al., 2016). The importance of a dense and

diverse dataset was stressed as they are assumed to contribute to the quality of every task that is performed (Zhou et al., 2018) since the CNNs can learn better from a set of good-quality images. The CNN-based scene classification model was examined on the Places standard dataset (King et al., 2017), which consists of 10 million samples. This study stressed that the key challenge was to find sufficient data to train DL models. In many land use scene classification tasks, CNN, combined with ML classifiers (Razavian et al., 2014), was frequently used as a feature extractor. The CNN-ML based models require less time and training datasets but retain desired classification accuracy. For example, combining CNN based-layers with constrained extreme learning machine (CELM), the CNN-CELM classifier reduced the training time with good generalization capability (Weng et al., 2017). Most studies focus on gaining higher classification accuracy by improving the construction of CNN-based models. For example, Yao et al. used CNN to classify land use from multi-scale samples using the BoVW method to map out land use patterns at the land parcel level.

Besides these efforts, there is a need to examine the performance of CNN on land use classification. Mainly because there are few studies that have used CNN for classifying and identifying the LUC and that achieved good performance (Huang et al., 2018; Yao et al., 2017), as well as considering the customization of training dataset. Consequently, the main objective of this study is to evaluate CNN-based classification models and to detect past LUC based on the classified historical LU data. The selected study area is a part of the Cloverdale study area within the City of Surrey, Canada as it has experienced rapid urbanization in the past decades.

2.3. Overview of CNN methods

2.3.1. The structure of CNN

The CNN models contain a sequence of processing layers where each layer have a group of algorithms (also called filters) that can learn the representation of data from low to high levels (LeCun et al., 2015). Such features provide information which can be merged in later stages to detect higher-level features allowing the CNNs to extract features automatically. Many modern approaches in the computer vision field use CNN to extract local features which rely only on small sub-regions of the image (Bishop, 2016). A basic CNN may include four main types of layers: convolutional layer (Conv),

rectified linear unit (ReLU) layer, pooling layer (Pool), and fully-connected layer (FC). An example structure of CNN is illustrated in Figure 2-1.

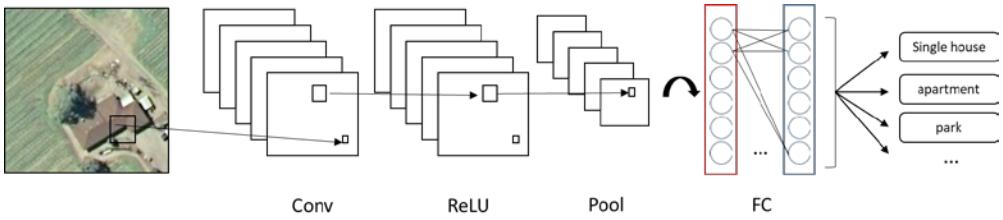


Figure 2-1 An example structure of CNN.

An image is organized as a matrix of pixel values. Given a random input image, the CNN recognizes the class that belongs to and the values of probability that the input belongs to each class. There are four basic types of layers making up a basic CNN as presented in Figure 2-2 and their roles are:

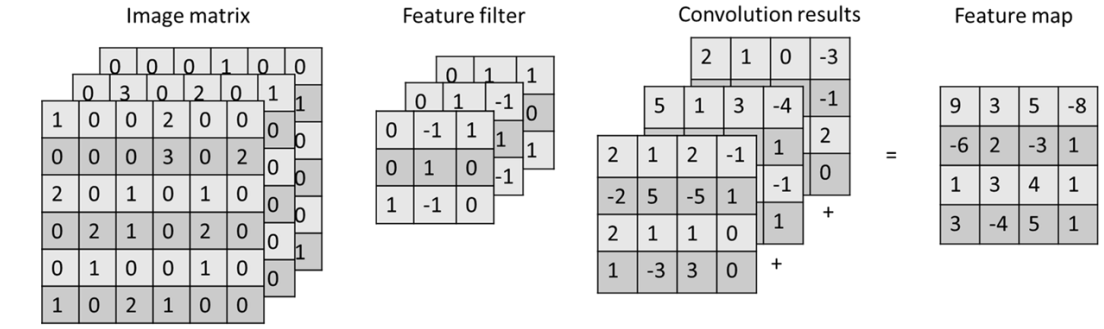
a. Convolutional layer performs the linear operation by doing element-wise multiplications and summations on every sub-input, using a group of weight matrixes called feature filters (Figure 2-2a). Feature filters learn their weights from training data, then each filter can detect the existence of a specific feature. The outputs are called feature maps. A feature map records not only the operation values but also the relative spatial location of these values. On the feature map, a higher output value indicates the potential existence of the corresponding feature at its relative location.

b. Rectified linear units (ReLU) layer performs a nonlinear function on the input (Figure 2-2b) to rescale the input values and output values that range from -1 to 1 (*tanh*) or from 0 to 1 (*sigmoid*). Introducing nonlinearity to the system can improve computational efficiency without losing much accuracy. When performing a threshold operation: f*(x) = max (0, x),* where *f* is a nonlinear function applied to all values *x* in the matrixes, the negative values of input matrixes are rectified as 0, and keeping the size of input volume unchanged.

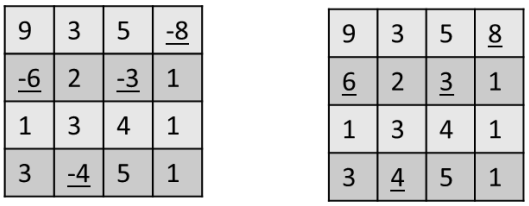
c. Pooling layer is a down-sampling layer (Figure 2-2c), aiming to decrease the size of the feature maps and to reduce the computational cost. A max pooling layer keeps only the maximal number of every sub-array (2x2) as the element of output array. Output omits unimportant features while keeps important features.

d. The fully connected layer multiplies the input by a weight matrix and then adds a bias vector to the output, an *N*-dimensional vector (Figure 2-2d). The output gives the probability of the input image belonging to each class, where *N* is the number of the category in a classification task.

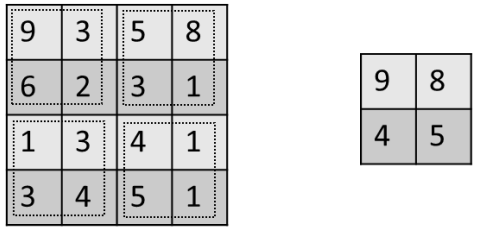
a. Convolutional layer



b. ReLU layer



c. Max Pooling layer



d. Fully connected layer

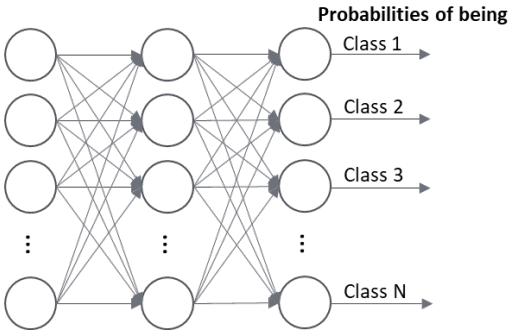


Figure 2-2 CNN layer functions: (a) Convolutional layer; (b) Rectified linear unit layer; (c) Max pooling layer; (d) Fully connected layer.

2.3.2. Training and testing of CNN

AlexNet, GoogLeNet, and VGGNet are typical CNN models that have been used in this study as the pre-trained models that have been trained by using millions of images. Moreover, the algorithms of CNN model can rely on millions of parameters. For instance, AlexNet (Krizhevsky et al., 2012c) model consists of 25 layers and has 62.3 million parameters. GoogLeNet was developed by Szegedy et al. (2014) and it has greatly reduced the number of parameters used in the network to 4 million, compared with 60 million typically used in AlexNet. VGGNet (Simonyan and Zisserman, 2014) has more layers than AlexNet, it has 140 million parameters so it uses more computer memory and time to be fully trained. Initially, parameters are randomly generated and they are learnable through the training and testing process. The parameters of AlexNet, GoogLeNet, and VGGNet are adjusted once through learning from millions of images.

Training of CNN can be performed by using supervised learning algorithms. Well-trained CNN models must learn from sufficiently labeled datasets so that the parameters could be greatly adjusted. Each training process of CNN consists of forwarding pass, loss function, backpropagation, and parameter update, and can be presented in the following six steps:

1. Model initialization: all the parameters are randomly generated;
2. Forward propagation: the input passes through the network layers and calculates the estimated output from the model;
3. Loss function calculation: loss function is used to evaluate the prediction ability of the method. The most common loss function uses the mean square error *(E)* between the model's expected output $y_{expected}$ and estimated output $y_{estimated}$ and can be formulated as:

$$E = \frac{1}{n}\sum_{i=1}^{n}(y_{estimated} - y_{expected})^2 \quad (2\text{-}1)$$

The value of *E* decreases during the training process until reaching a constant that is used as a threshold for the next step. The objective of the training process is to minimize the loss function by feeding the network many inputs and the desired output;

4. Backpropagation: it is a process that backward propagates the derivative of errors;

5. Parameters update: In general, the weights (parameters) update by the delta rules (Russell, 2018) defined as:

$$New\ weight = Old\ weight - Derivative\ Rate * learning\ rate \quad (2\text{-}2)$$

If the deviation is positive the parameter should be decreased; if the deviation is negative the parameter should be increased; if the deviation equals to 0, then the parameter is optimal. The learning rate is a constant and should be set based on experience; if too small, it will take a long time to get optimal weights; if too large, the weights will deviate from optimal;

6. Iterating the process until convergence. Based on the previous step, the weights get updated very slowly, thus requesting many iterations to get the desired weights and minimize the loss function. In reality, the CNN model was trained by a batch of images simultaneously at one training epoch in order to increase efficiency.

The labeled dataset from the same probability distribution as the training dataset (Ripley, 1996) is split in the ratio of 70:30, and 70% were training set and 30% were testing set, the idea of choosing this ratio is that more training data is better for model to fine tune the parameters whilst test data were used to estimate the error. The 70:30 is a one of the commonly used ratio to split the data (Gholami, Chau, Fadaie, Torkaman, & Ghaffari, 2015), since this research study focuses more on the selection of training datasets and the application of CNN on real study area. The other ratios such as 50:50, 60:40 or 80:20 were not discussed, but the influence of different split ratio is factor that can impact the output results and need additional consideration in future work.

The model is trained by the training set, after which the testing set is used for model evaluation. In a classification task, the following indicators (Equations 2-3~2-5) are used to assess the performance of the classifier model：

$$Total\ accuracy = \frac{TurePositive+TrueNegative}{TurePositive+TrueNegative+FalsePostive+FalseNegative} \quad (2\text{-}3)$$

$$Precision = \frac{TruePositive}{TurePositive+FalsePositive} \quad (2\text{-}4)$$

$$Recall = \frac{TruePositive}{TurePositive+FalseNegative} \quad (2\text{-}5)$$

where true positive (TN) = correctly identified, false positive (FP) = incorrectly identified, true negative (TN) = correctly rejected, false negative (FN) = incorrectly rejected pixels.

2.4. The land use classification method

2.4.1. Datasets for transfer learning

The UC Merced Land Use Dataset contains 21-class high-resolution land use images of the size of 256×256 pixels with 30cm resolution. Examples of images are shown in Figure 2-4. Each category in the UC Merced dataset has 100 images. These images were manually extracted from the USGS National Map Urban Area Imagery collection for various urban areas around the USA (Yang and Newsam, 2010). But different countries and cities have different land use patterns entailed by different urban morphology. For example, most cities in China have very dense high-rise buildings, similar to downtown areas such as New York City, USA. By contrast, most residential areas in the Canada and USA are characterised by urban sprawl with low urban density where majority of housing consist of low-rise buildings. Therefore, if only the LU images from USA are used to train to CNN models, the CNN models can recognize the LU features from the USA but China. The LU features from the USA and Canadian cities also have some differences despite both belonging to North America. Currently, there is a lack of a comprehensive dataset that includes LU features sampled from different cities and countries. In the study area, i.e. Community of Cloverdale, the agriculture, building and road network LU types are not very similar with the LU images from the UC Merced dataset based on observation, so manually sampled images were used for these categories instead.

In order to address these varieties of possible land use features and their classifications, this research used a mixture of the UC Merced dataset and the manually sampled images from orthophotos of the City of Surrey in Canada. The size of the manually sampled images is 700×700 pixels and the resolution is 10cm. The size of images from the UC Merced dataset is 256×256 and the resolution is 30.5cm. Therefore, the images from both sources cover almost the identical ground area. The manually sampled images were based on manual classification of images from LU types of agriculture, building and road networks, and each LU types contained 400 images.

The six main land-use classes used for classification and analysis were: (1) Green areas (GA); (2) Industrial and commercial areas (In&Co); (3) High-density residential (HR); (4) Low-density residential (LR); (5) Parking lot (PL); and (6) Road network (RN). For the purpose of model training and testing with 400 images for each LU class, the images were split randomly by the ratio of 70% (280 images) and 30% (120 images) with a total of 2520 and 1080 images for training and testing for each land use, respectively.

2.4.2. Transfer learning strategies

Transfer learning is a procedure that reuses the exiting ML or DL models that have been developed for other tasks and then re-applied to a new task without re-development. Pre-trained CNN models mentioned in section 2.3.2 were already developed. Therefore, using pre-trained CNN models in the LU classification task is transfer learning method. These pre-trained networks have been trained by millions of images. In reality, it is difficult to have datasets with sufficient size to train a complete network. In order to complete the training of an entire network, the training process can even take 2-3 weeks on multiple GPUs. Due to time restrictions or computational constraints in many pieces of research, transfer learning becomes the first choice to start a new task by adopting pre-trained CNN models.

There are three main options to apply transfer learning (Goodfellow, Bengio, & Courville, 2016; Yosinski, Clune, Bengio, & Lipson, 2014). As shown in Figure 2-4, each option has different computational effort and accuracy. These options are:

a. Fine-tuning the whole network (Figure 2-4a): Loading the pre-trained network, replacing the fully-connected layer of the network, retraining the network using new dataset, and fine-tuning the weights of all layers through backpropagation.

b. Fine-tuning the higher layers of the network (Figure 2-4b): Loading the pre-trained network, and freezing the parameters in the earlier layers of the network, so when retraining with new dataset, only the weights of last several layers will be updated.

c. Convolutional base as a feature extractor and link a support vector machine (SVM) to form a CNN-SVM based model (Figure 2-4c): Loading the pre-trained network and feeding the data through the convolutional base; the output as the feature representations are then classified by SVM. This solution also requires less computational effort.

Figure 2-3 Example images used for training the models; the UC-Merced dataset (the first two rows), and the third row presents the manually extracted samples.

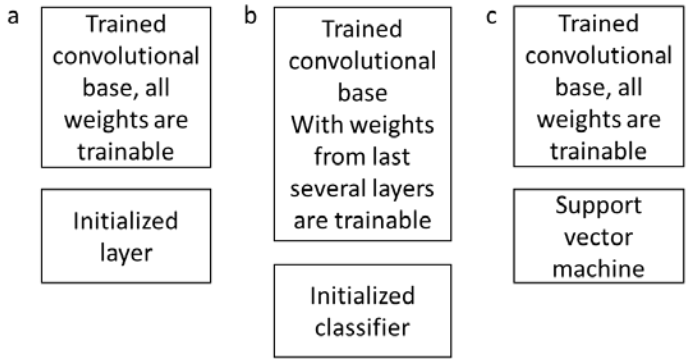


Figure 2-4 Three different transfer learning scenarios: (a) Fine-tuning the whole network; (b) Fine-tuning higher layers of the network; (c) Convolutional base as feature extractor.

2.5. Methodology

2.5.1. Study area and primary data

The study area is a section of the community of Cloverdale, at the north-eastern part of the City of Surrey, Canada (Figure 2-5). The community of Cloverdale has been transformed over the years from a rural agricultural community to a prosperous small city attractive for young families, and it is known by its fast developing residential areas (City of Surrey, 2018a), therefore, it was chosen as the suitable area to detect LU change through DL models. The City of Surrey database provide orthophotos from 2004 to 2017, from which the LU classification was implemented. Since each orthophoto was large and took average 4 hours to classify, only the selected images of good quality were processed for simplicity reasons. Particularly, the orthophotos for years 2004, 2006, 2011, 2013, 2015 and 2017 were considered in order to document the LU change for the past 12 years. The digital orthophotos cover a section of the community of Cloverdale (City of Surrey, 2017) with size 2.0km × 3.6km and with 10cm spatial resolution. The 2011 land use data from the Metro Vancouver Open Data Catalogue (Metro Vancouver, 2011) and from DMTI Spatial Inc were also used as references.

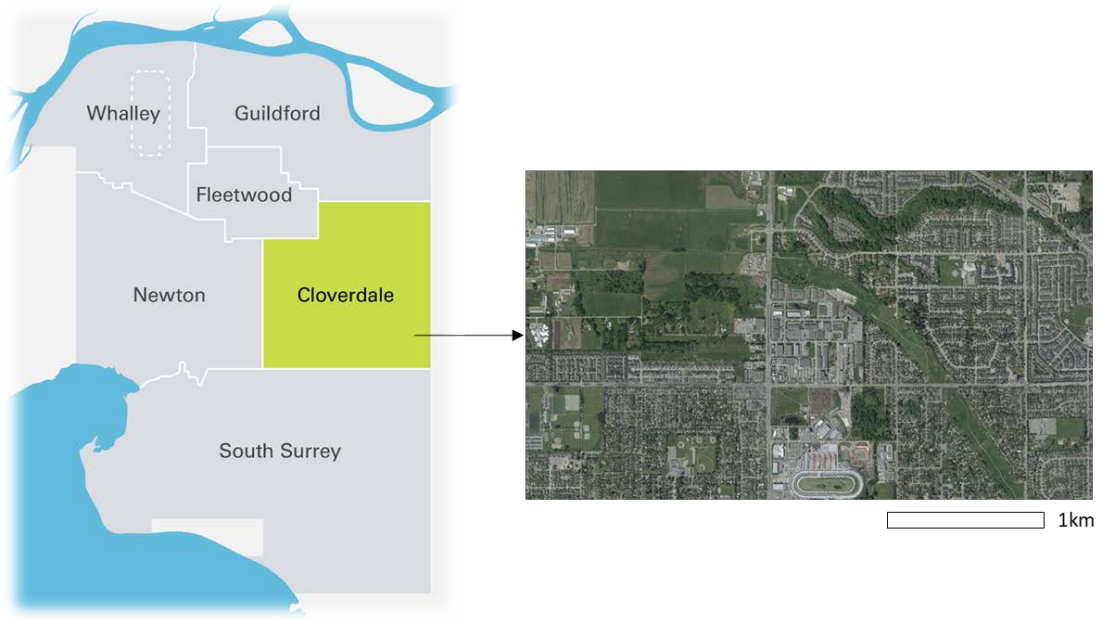


Figure 2-5 Study area: north-east section of Cloverdale, community of the City of Surrey, Canada (City of Surrey, 2017; COSMOS, n.d.)

2.5.2. Land use change analysis

Figure 2-6 shows the process used to perform the land use classification to identify the changes.

1. Using the enriched UC-Merced dataset to fine-tune ‘GoogLeNet-‘pool5-drop’-SVM’ models.

2. The image of the study area is covering 20 000×36 000 pixels at a spatial resolution of 10cm. The window with a pixel size of 700×700 shifts over and crops the image in a specific order (from top to bottom and from left to right), with a stride of 100 pixels, therefore, there are {(20000-700)/100+1}×{(36000-700)/100+1} sub-patches, i.e., 194×354. The moving window covers all the image after 194×354 iterations with the same number of sequentially generated patches. These patches are sent to the classifier and land use labels were returned. The numbers from 1 to 6 represent the six land-use labels. To extract the road network and green area land use classes, the second process was repeated, while the size of moving window was set to 300×300 pixel, the matrix was only updated when the return label is 1 or 6. This process is done due to green area and road network can exit in a small ground area independently. They can also become background of another feature if the moving window size is larger. For example, in a 700×700 pixel image, a small house can exist as agriculture land simultaneously, but the model will still think the image is sparse residential. Therefore, the smaller window size, i.e. 300×300 pixel, was used again to separate out individual features from potentially misclassified 700×700 pixel images.

3. A 194×354 void matrix was created, the cells in the matrix with index i correspond to the ith image patches. The ith cell values were updated as the label number of the ith patch. Each cell represents a land use category of a 10m spatial resolution for ground truth.

4. The same classification model was applied to orthophotos for different years 2004, 2006, 2011, 2013, 2015, and 2017. Six matrixes represent the digital LU maps from different year, where each cell of the matrix store the index of classified LU type, so that each LU types can be quantified and analyzed.

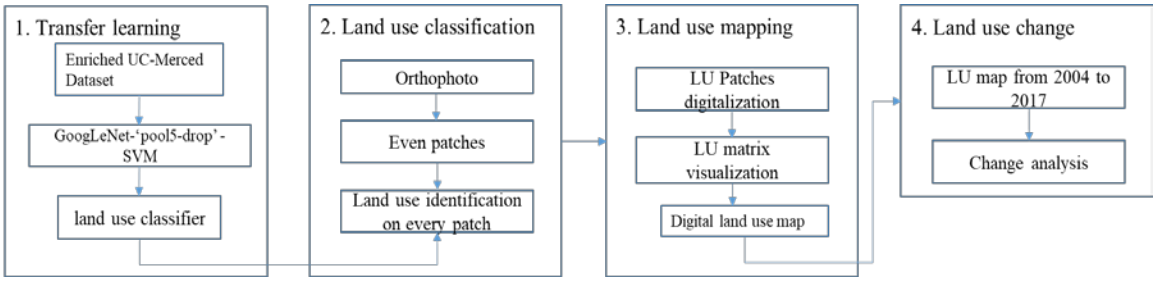


Figure 2-6 Flowchart of the process for land use change analysis based on the retrained CNN.

2.6. Result and discussion

The obtained values of total accuracy are given in Table 2-1 based on the eight transferred CNN-based models. The total accuracy was calculated based on equation (2-3). Models (a) and (b) were respectively structured by fine-tuning the whole AlexNet and GoogLeNet. Models (c) and (d) used 'fc7' or 'fc6' layers from AlexNet as feature extractors combined with SVM. Models (e) and (f) were respectively based on the 'loss classifier' or 'pool5-drop' from GoogLeNet as feature extractors combined with SVM. Models (g) and (h) used respectively 'fc7' or 'fc6' from VGGNet as feature extractors and were combined with SVM. Models (c) to model (h) were set to test the effect of using different layers from CNN as feature extractors, the results show the difference in testing accuracy which is not significant. Models (a) and (b) had slightly lower accuracy after training, compared with other models, which indicated that models that using CNN combined with SVM can be more efficient than fine-tuning the CNNs.

Table 2-1 shows that the obtained values for model total accuracy from each transferred model ranges from 95.1% to 98.1%. The transferred CNN models perform well on the land use classification, with more than 95.0% accuracy. The accuracy difference is not large, however the GoogLeNet -'pool5-drop'- SVM classifier was chosen for further classification in the next step given its best performance on our training dataset. As an example, Figure 2-7 depicts the confusion matrix for the highest performing GoogLeNet -'pool5-drop'- SVM. All the transferred models were trained on a single GPU 'Quadro P2000'. The average computing time of models (a) and (b) were 45 min with single GPU, and models (c) to (h) took less than 5 minutes to generate classifier, which proved the time efficiency of using the transferred models, especially the CNN-SVM models.

Table 2-1 **The obtained values for the model total accuracy of the eight transferred CNN-based models where training and testing datasets are kept the same for each model.**

New models	Total accuracy
a. AlexNet, full-trained	95.80%
b. GoogLeNet, full-trained	95.06%
c. AlexNet-'fc7'-SVM	95.68%
d. AlexNet-'fc6'-SVM	97.11%
e. GoogLeNet-'loss classifier'-SVM	97.47%
f. GoogLeNet-'pool5-drop'-SVM	98.14%
g. VGGNet-'fc6'-SVM	95.93%
h. VGGNet-'fc7'-SVM	96.99%

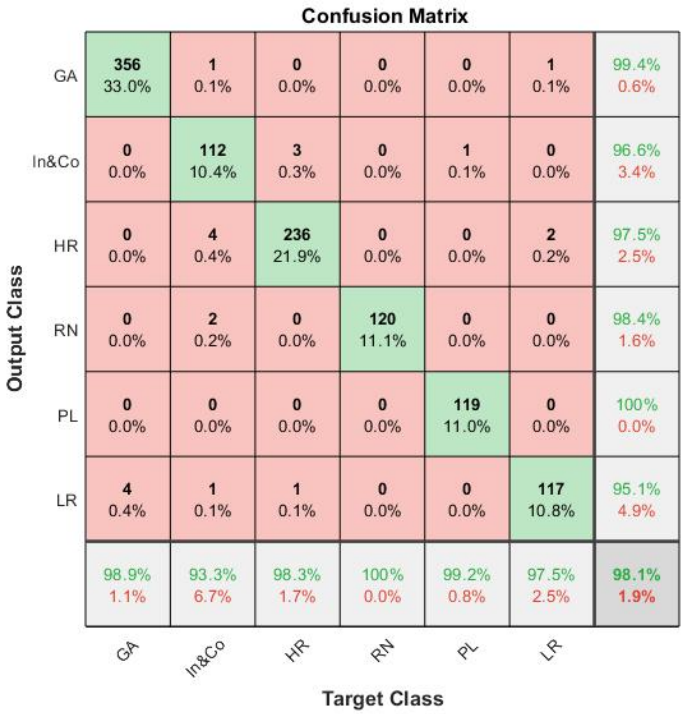


Figure 2-7 **Confusion matrix for the highest performing GoogLeNet -'pool5-drop'- SVM. GA: Green area; In&Co: Industrial and Commercial areas; HR: High-density Residential areas; LR: Low-density Residential areas; PL: Parking lot; RN: Road network.**

Figure 2-8 depicts the comparison of the obtained land use classification map with the digital orthophoto of the study area in 2017. The boundary between different land use classes is clearly defined in the resulting land use map. Except for the north-western part of the study area, it can be seen that most of the remaining areas are classified as denser residential land-use. The middle of the study area was classified as industrial or commercial buildings land-use. The outline of major roads is distinct. Figure

2-8c presents the orthophoto image and the classified image of the track field at the north of the study area, indicating that the track field was fully occupied by cars. At the same location on the classified image, there is a black elliptical area at the south of the study area which was labeled as a parking lot. Even though this elliptical area was misclassified, the model still demonstrated the sensitivity to capture the features of cars and the shape of the track field.

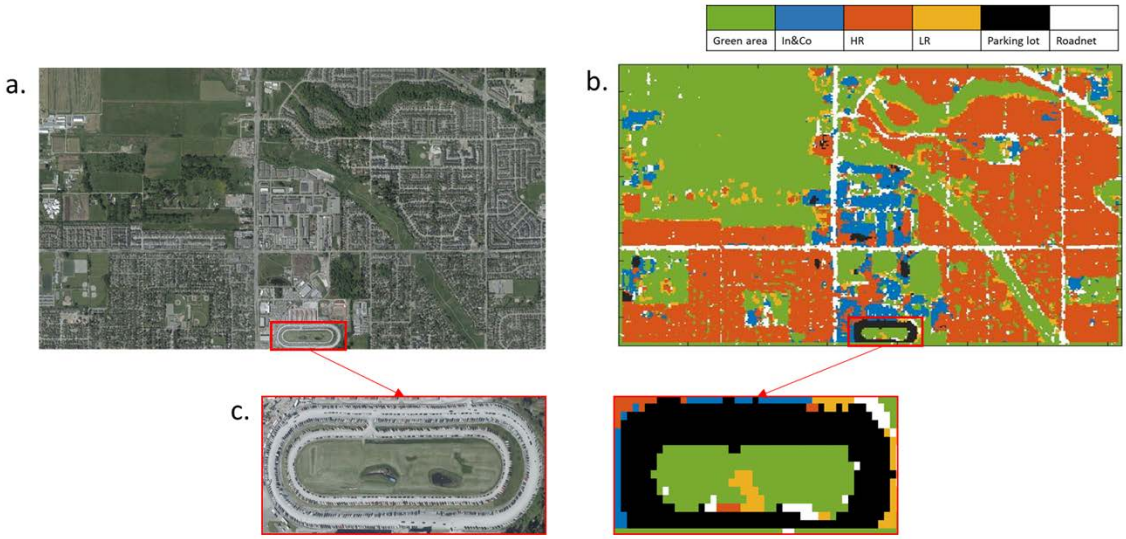


Figure 2-8 **(a) The orthophoto image of the study area with (b) The digital land use map based on classification results for the year 2017: (1) Green area (GA), green color; (2) Industrial and Commercial areas (In&Co), blue color; (3) High-density Residential areas (HR), red color; (4) Low-density Residential areas (LR), yellow color; (5) Parking lot (PL), black color; (6) Road network (RN), white color. (c) Detailed image of the zone with the field track filled with cars.**

Based on the transferred "GoogLeNet-'pool5-drop'-SVM" model with higher accuracy trained by the enriched UC-Merced dataset, the land use classification was performed for years 2004, 2006, 2011, 2013, 2015 and 2017 as shown in Figure 2-9. The composition result of land use classes is given in Table 2-2, which shows that from 2004 to 2017 the green area decreased during urban development. Dense residential areas were increasing over the years and low-density residential areas were decreasing, as can be expected from the growing population (City of Surrey, 2018) in this area. From 2006 to 2011, it can be seen that a new residential community appear at the north of the study area. The percentage of road network was decreasing, because the density of building and residential areas increased, which have occupied the road area. In 2017,

the old track field at the north of the study area was used for parking. The parking lots were usually located at nearby shopping malls and industrial buildings. These features can imply the location and number of Industrial and Commercial (In&Co) land use class, which also increased, while the change was not significant. The overall results indicate that the community of Cloverdale tends to develop itself as a predominantly residential area and not as an industrial or agricultural area.

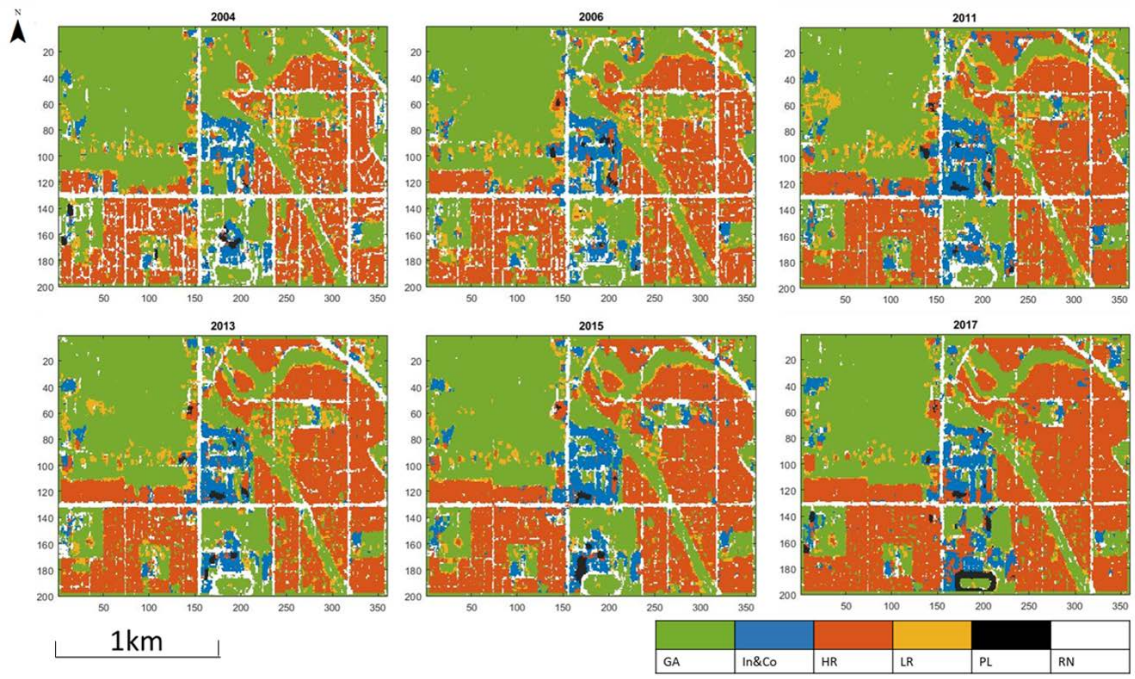


Figure 2-9 **Land use classifications obtained for years 2004, 2006, 2011, 2013, 2015 and 2017. GA: Green area; In&Co: Industrial and Commercial areas; HR: High-density Residential areas; LR: Low-density Residential areas; PL: Parking lot; RN: Road network.**

Table 2-2 **Land use constitution of the study area from 2004, 2006, 2011, 2013, 2015 and 2017. GA: Green area; In&Co: Industrial and Commercial areas; HR: High-density Residential areas; LR: Low-density Residential areas; PL: Parking lot; RN: Road network.**

	GA	In&Co	HR	LR	PL	RN
2004	44.13%	5.45%	26.61%	5.52%	0.37%	15.25%
2006	40.77%	5.69%	30.39%	7.51%	0.37%	12.35%
2011	36.79%	8.44%	35.46%	6.37%	0.45%	9.19%
2013	39.30%	5.85%	34.77%	6.43%	0.44%	10.00%
2015	38.91%	7.55%	35.93%	4.91%	0.52%	8.87%
2017	37.61%	6.60%	40.89%	3.83%	1.01%	6.62%

Figure 2-10 presents the comparison of the obtained land use classification with the reference land use data for year 2011 (DMTI Spatial Inc., 2011; Metro Vancouver, 2011). 2011 land use map is the newest map the municipality has provided to the public. Typically, public LU data for multiple years are difficult to obtain. They also contain different classifications standards that make their use challenging in computationally intensive data processing necessary for studies of LUC process evolved over time. The 2011 land use data (Figure 2-10b), from the Metro Vancouver Open Data Catalogue, contains more land use types such as townhouse, high-rise building, and institutional buildings, while land use data for more recent years are not yet available. The land use data from DMTI Spatial Inc. for the same year (Figure 2-10c) has lower thematic resolution and contains only the residential, open area, parks, and government institutional land use classes. Due to the lack of high-quality land use data and the orthophoto images, the proposed classification approach provides an economical solution for the LUC analysis.

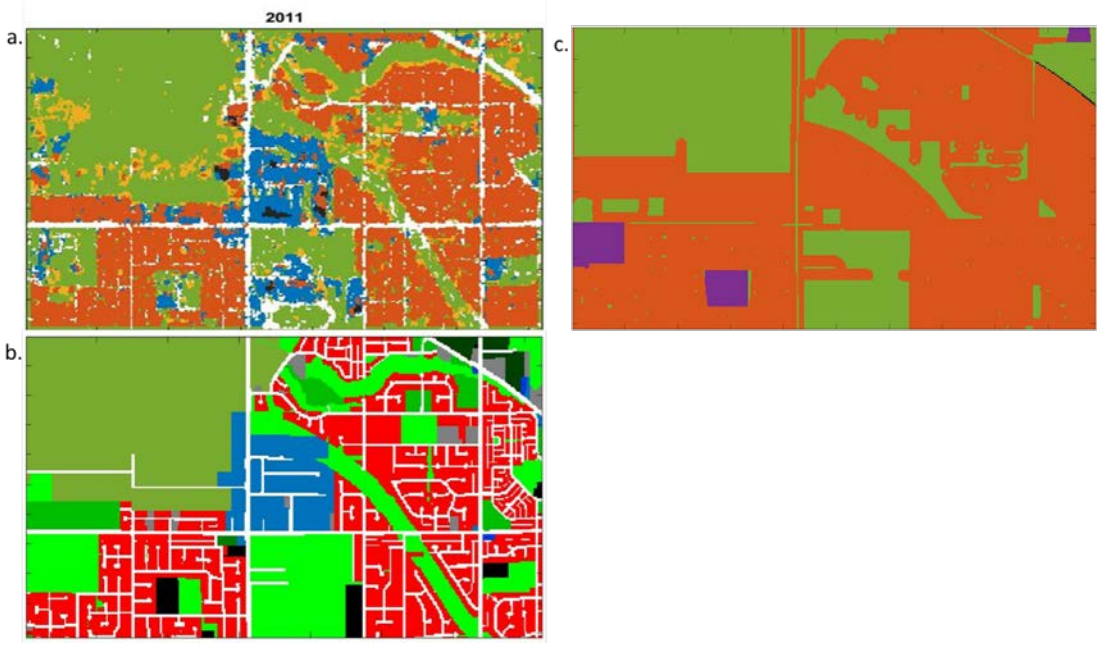


Figure 2-10 Comparison of the LU classifications for the year 2011: (a) resulting from the proposed CNN approach for the LU classification; (b) from Metro Vancouver Open Data Catalogue and (c) from DMTI Spatial Inc

2.7. Conclusions

This research has focused on the (1) evaluation of eight transferred CNN-based models on land use classification tasks, and (2) application of the best performing transferred CNN-based model as a classifier to classify and map the land use from orthophotos of the study area. The results showed that successful land use classification performance from all transferred models with at least 95.0% accuracy was obtained, compared with other studies such as 98.67% achieved by Othman et al. (2016) using CNN combined with neural networks (NN). The accuracy from the CNN models in this study is comparably high. The training dataset combined the off-the-shelf UC-Merced dataset with the manually classified images from the study area in order to enrich the data and teach the model the local land use features. Using the digital maps, the land use change of the study area over the years was interpreted.

The study has indicated the effectiveness of CNN-SVM models for land-use classification, especially with the “GoogLeNet-‘pool5-drop’-SVM” model. The training data that was obtained from a real study area, including the manual image interpretation, can improve the overall classification results with significant time efficiency. Incorporating manual image interpretation makes the experiment less automatic, and it increases time for project completion. The proposed methodology requires selection and augmentation in every specific study area to reach better LU classification accuracy, which needs a little priori knowledge of the study area. Augmenting the amount of training datasets to cover wide variety of world regions will be helpful to improve classification and achieve higher accuracy using CNN models. Overall, the proposed land use classification method can be transferred to analyse real-time digital orthophotos datasets from other urban-rural fringe areas, providing a fast and low-cost solution to aid land use change analysis and to assist municipal land management.

2.8. References

Arel, I., Rose, D. C., & Karnowski, T. P. (2010). Deep Machine Learning—A New Frontier in Artificial Intelligence Research. IEEE Computational Intelligence Magazine, 5(4), 13–18. https://doi.org/10.1109/MCI.2010.938364

Bishop, C. M. (2016). Pattern Recognition and Machine Learning. Springer New York.

Cheng, G., Han, J., Guo, L., Liu, Z., Bu, S., & Ren, J. (2015). Effective and efficient midlevel visual elements-oriented land-use classification using VHR remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 53(8), 4238–4249. https://doi.org/10.1109/TGRS.2015.2393857

City of Surrey. (2017). Imagery—City of Surrey Open Data Catalogue. Retrieved September 10, 2018, from http://data.surrey.ca/group/6878e307-9fec-4134-b042-d7e058310255?tags=orthophoto

City of Surrey. (2018a). Cloverdale. Retrieved May 27, 2019, from https://www.surrey.ca/community/6799.aspx

City of Surrey. (2018b). Population Estimates & Projections [Text/xml]. Retrieved May 24, 2019, from http://www.surrey.ca/business-economic-development/1418.aspx

Cloverdale. (n.d.). Retrieved January 19, 2019, from City of Surrey website: http://www.surrey.ca/community/6799.aspx

COSMOS. (n.d.). City of Surrey Mapping Online System (COSMOS). Retrieved February 5, 2019, from http://cosmos.surrey.ca/external/

DMTI Spatial Inc. (2011). CanMap Streetfiles. Retrieved May 24, 2019, from https://www.dmtispatial.com/canmap/

Etingoff, K. (2017). Urban Land Use: Community-Based Planning. CRC Press.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., … Snyder, P. K. (2005). Global Consequences of Land Use. Science, 309(5734), 570–574. https://doi.org/10.1126/science.1111772

Gholami, V., Chau, K. W., Fadaie, F., Torkaman, J., & Ghaffari, A. (2015). Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers. Journal of Hydrology, 529, 1060–1069. https://doi.org/10.1016/j.jhydrol.2015.09.028

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A Fast Learning Algorithm for Deep Belief Nets. Neural Computation, 18(7), 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527

HLPF. (2017). Progress towards the Sustainable Development Goals (No. N1713409). Retrieved from http://www.un.org/ga/search/view_doc.asp?symbol=E/2017/66&Lang=E

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hu, F., Xia, G. S., Hu, J., & Zhang, L. P. (2015). Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery. Remote Sensing, 7(11), 14680–14707. https://doi.org/10.3390/rs71114680

Huang, B., Zhao, B., & Song, Y. (2018). Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. Remote Sensing of Environment, (214), 73–86.

Hughes, L. H., Schmitt, M., Mou, L., Wang, Y., & Zhu, X. X. (2018). Identifying Corresponding Patches in SAR and Optical Images with a Pseudo-Siamese CNN. IEEE Geoscience and Remote Sensing Letters, 15(5), 784–788. https://doi.org/10.1109/LGRS.2018.2799232

Ienco, D., Gaetano, R., Dupaquier, C., & Maurel, P. (2017). Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks. IEEE Geoscience and Remote Sensing Letters, 14(10), 1685–1689. https://doi.org/10.1109/LGRS.2017.2728698

Isikdogan, F., Bovik, A. C., & Passalacqua, P. (2017). Surface Water Mapping by Deep Learning. Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(11), 4909–4918. https://doi.org/10.1109/JSTARS.2017.2735443

Joseph, G. (2005). Fundamentals of Remote Sensing. Universities Press.

Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A Convolutional Neural Network for Modelling Sentences. ArXiv:1404.2188 [Cs]. Retrieved from http://arxiv.org/abs/1404.2188

King, J., Kishore, V., & Ranalli, F. (2017). Scene classification with Convolutional Neural Networks (Course Project Reports No. 102; p. 7). Stanford, CA, USA: Stanford University.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012a). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 25 (pp. 1097–1105). Retrieved from http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012b). ImageNet Classification with Deep Convolutional Neural Networks. Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, 1097–1105. Retrieved from http://dl.acm.org/citation.cfm?id=2999134.2999257

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012c). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 25 (pp. 1097–1105). Retrieved from http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

Lawrence, S., Giles, C. L., Tsoi, A. C., & Back, A. D. (1997). Face recognition: A convolutional neural-network approach. IEEE Transactions on Neural Networks, 8(1), 98–113. https://doi.org/10.1109/72.554195

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. Neural Computation, 1(4), 541–551. https://doi.org/10.1162/neco.1989.1.4.541

LeCun, Yann, Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. Neurocomputing, 234, 11–26. https://doi.org/10.1016/j.neucom.2016.12.038

Liu, Y., Zhong, Y., Fei, F., & Zhang, L. (2016). Scene semantic classification based on random-scale stretched convolutional neural network for high-spatial resolution remote sensing imagery. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 763–766. https://doi.org/10.1109/IGARSS.2016.7729192

Liu, Yanfei, Zhong, Y., Fei, F., Zhu, Q., Qin, Q., Liu, Y., … Qin, Q. (2018). Scene Classification Based on a Deep Random-Scale Stretched Convolutional Neural Network. Remote Sensing, 10(3), 444. https://doi.org/10.3390/rs10030444

Lu, Y., Yi, S., Zeng, N., Liu, Y., & Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. Neurocomputing, 267(Supplement C), 378–384. https://doi.org/10.1016/j.neucom.2017.06.023

Luus, F. P. S., Salmon, B. P., Bergh, F. van den, & Maharaj, B. T. J. (2015). Multiview Deep Learning for Land-Use Classification. IEEE Geoscience and Remote Sensing Letters, 12(12), 2448–2452. https://doi.org/10.1109/LGRS.2015.2483680

Metro Vancouver. (2011). Open Data Catalogue [Government]. Retrieved December 21, 2018, from Metro Vancouver website: http://www.metrovancouver.org/data

Nogueira, K., Penatti, O. A. B., & Santos, J. A. dos. (2017a). Towards Better Exploiting Convolutional Neural Networks for Remote Sensing Scene Classification. Pattern Recognition, 61, 539–556. https://doi.org/10.1016/j.patcog.2016.07.001

Nogueira, K., Penatti, O. A. B., & Santos, J. A. dos. (2017b). Towards Better Exploiting Convolutional Neural Networks for Remote Sensing Scene Classification. Retrieved December 19, 2018, from Undefined website: /paper/Towards-Better-Exploiting-Convolutional-Neural-for-Nogueira-Penatti/3719960f974173f23b88a207a42d67d7a393a89a

Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2014). Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks. 1717–1724. Retrieved from https://www.cv-foundation.org/openaccess/content_cvpr_2014/html/Oquab_Learning_and_Transferring_2014_CVPR_paper.html

Othman, E., Bazi, Y., Alajlan, N., Alhichri, H., & Melgani, F. (2016). Using convolutional features and a sparse autoencoder for land-use scene classification. International Journal of Remote Sensing, 37(10), 2149–2167. https://doi.org/10.1080/01431161.2016.1171928

Rashedi Nia, K. (2017). Automatic Building Damage Assessment Using Deep Learning and Ground-Level Image Data (Thesis, Applied Sciences: School of Computing Science). Retrieved from http://summit.sfu.ca/item/16986

Razavian, A. S., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN Features off-the-shelf: An Astounding Baseline for Recognition. ArXiv:1403.6382 [Cs]. Retrieved from http://arxiv.org/abs/1403.6382

Ripley, B. D. (1996). Pattern Recognition and Neural Networks. Cambridge University Press.

Romero, A., Gatta, C., & Camps-Valls, G. (2016). Unsupervised Deep Feature Extraction for Remote Sensing Image Classification. IEEE Transactions on Geoscience and Remote Sensing, 54(3), 1349–1362. https://doi.org/10.1109/TGRS.2015.2478379

Russell, I. (n.d.). The Delta Rule. Retrieved December 21, 2018, from http://uhaweb.hartford.edu/compsci/neural-networks-delta-rule.html

Scott, G. J., England, M. R., Starms, W. A., Marcum, R. A., & Davis, C. H. (2017). Training Deep Convolutional Neural Networks for Land -Cover Classification of High-Resolution Imagery. IEEE Geoscience and Remote Sensing Letters, 14(4), 549–553. https://doi.org/10.1109/LGRS.2017.2657778

Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. ArXiv:1409.1556 [Cs]. Retrieved from http://arxiv.org/abs/1409.1556

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol, P. A. (2010). Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. Journal of Machine Learning Research, 11(Dec), 3371–3408.

Weng, Q., Mao, Z., Lin, J., & Guo, W. (2017). Land-Use Classification via Extreme Learning Classifier Based on Deep Convolutional Features. Ieee Geoscience and Remote Sensing Letters, 14(5), 704–708. https://doi.org/10.1109/LGRS.2017.2672643

Xia, G.-S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., & Zhang, L. (2017). AID: A Benchmark Dataset for Performance Evaluation of Aerial Scene Classification. IEEE Transactions on Geoscience and Remote Sensing, 55(7), 3965–3981. https://doi.org/10.1109/TGRS.2017.2685945

Xu, G., Zhu, X., Fu, D., Dong, J., & Xiao, X. (2017). Automatic land cover classification of geo-tagged field photos by deep learning. Environmental Modelling & Software, 91, 127–134. https://doi.org/10.1016/j.envsoft.2017.02.004

Yang, Y., & Newsam, S. (2010). Bag-of-visual-words and Spatial Extensions for Land-use Classification. Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, 270–279. https://doi.org/10.1145/1869790.1869829

Yao, Y., Li, X., Liu, X., Liu, P., Liang, Z., Zhang, J., & Mai, K. (2017). Sensing Spatial Distribution of Urban Land Use by Integrating Points-of-interest and Google Word2Vec Model. Int. J. Geogr. Inf. Sci., 31(4), 825–848. https://doi.org/10.1080/13658816.2016.1244608

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 27 (pp. 3320–3328). Retrieved from http://papers.nips.cc/paper/5347-how-transferable-are-features-in-deep-neural-networks.pdf

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., & Atkinson, P. M. (2019). Joint Deep Learning for land cover and land use classification. Remote Sensing of Environment, 221, 173–187. https://doi.org/10.1016/j.rse.2018.11.014

Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., & Torralba, A. (2018). Places: A 10 Million Image Database for Scene Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(6), 1452–1464. https://doi.org/10.1109/TPAMI.2017.2723009

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A review. IEEE Geoscience and Remote Sensing Magazine, 5(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307

Zin, T. T., & Lin, J. C.-W. (2018). Big Data Analysis and Deep Learning Applications: Proceedings of the First International Conference on Big Data Analysis and Deep Learning. Springer.

Chapter 3.

Short-term forecasting of land use change using Recurrent Neural Network models[2]

3.1. Abstract

Land use change (LUC) is a dynamic process that significantly impacts the environment and various approaches have been proposed to analyze and model the LUC process for sustainable land use management and decision making. Recurrent neural network (RNN) models are part of deep learning (DL) approaches that has the capability to capture spatial and temporal information from sequential data. The main objective of this study was to examine variants of the RNN models by applying and comparing them when forecasting LUC in short time periods. Historical land use data for the City of Surrey, British Columbia, Canada were used to implement the several variants of the RNN models. The land use data for years 1996, 2001, 2006 and 2011 were used for training the DL models to enable the short-term forecast for the year 2016. For the 2011 to 2016 period, only 4.5% of the land use in the study area changed. The results indicate that an overall accuracy of 86.9% was achieved while actual changes in each LU type were forecasted with a relatively lower accuracy. This research study demonstrates that RNN models provide a suite of valuable tools for short-term LUC forecast that can inform and complement the traditional long-term planning process.

3.2. Introduction

Land use change (LUC) arises from human-environmental interactions (Kleemann, Baysal, Bulley, & Fürst, 2017) and so far only about 39% of the earth's land has never been exploited or used for the benefits of humans (De Palma et al., 2018). Land use change (LUC) with urban intensification result in pressures to the natural environment that can produce irreversible damages if not adequately addressed. Better knowledge and understanding of the LUC process can help policymakers to make

[2] A version of this chapter coauthored with Dragicevic, S., and Li S has been submitted to *Sustainability* journal for peer review.

informed decisions for sustainable land management. Sustainable land management practices promote activities that increase the benefit of utilization and development of land resources for individual, social, and economic purposes. LUC analysis and modeling methods can assist with the projection of future LU patterns, thus helping and guiding the management of land towards sustainable urban development.

The LUC phenomenon is typically studied through earth observations (EO), remote sensing (RS) and field measurements (Green, Kempka, & Lackey, 1994) all of which provide the opportunity for monitoring and quantifying change of LU patterns at local, regional and global levels. LUC is a complex phenomenon occurring locally and with implications for global geographic scales. Decades ago, RS sensors provided data with lower resolution and the availability of this data to the public was very limited. Besides EO and RS techniques that require advanced satellite equipment and expert knowledge for data interpretation, researchers have been using LUC modeling approaches for decades. LUC models provide representations and strategies that can help analyze, understand and assist the planning and management of land and natural resources. Many LUC models are based on inductive approaches which start with studying the observations and then developing explanations (Anderson, Hardy, Roach, & Witmer, 1976; National Research Council, 2014). LUC models are often based on a suite of explanatory variables that potentially drive the change process. But the main factors of LUC are directly related to human interactions and decision-making processes that are often difficult to accurately model and predict.

Deep learning (DL) are a subset of machine learning (ML) approaches and can be considered as deep machine learning. ML and DL both work by reducing the dimensionality and extracting features of large datasets. Compared with traditional ML methods, DL models can simultaneously extract and classify features with faster computation. Recent access to larger volumes of data from open sources coupled with superior computational abilities have made DL models as a great potential to be valuable tools capable of exploring and analyzing LUC phenomena. DL models include both the convolutional neural network (CNN) (Fukushima, 1980) and the recurrent neural network (RNN) (Rumelhart, Hinton, & Williams, 1986). CNNs have been used for image classification (Schmidhuber, 2017) and LUC classification and mapping (Huang, Zhao, & Song, 2018) while RNNs have been used for natural language processing (NLP) tasks (Graves, Mohamed, & Hinton, 2013; Mahoney, 2017). Deep learning (DL) has been

identified as intelligent models for advancing the field of LU modeling (Yao, Li, et al., 2017; Yao, Liang, Li, Zhang, & He, 2017).

The RNNs have the capability to capture information within sequential datasets such as in spatial and temporal sequences (J. Liu, Shahroudy, Xu, Kot, & Wang, 2018). Due to the spatio-temporal nature of LUC processes, the main objective of this research study is to examine the capabilities of RNN-based models to model LUC from an integrated space-time perspective and to perform a short-term forecast of LU. Historical land use data for the City of Surrey, British Columbia, Canada were used to implement the RNN models and generate the forecasted LUC.

3.2.1. Land use change models

The usual approaches for monitoring urban growth and LUC detection are based on geographic information systems (GIS) and RS (Hegazy & Kaloop, 2015) techniques and available geospatial datasets, which may require intensive pre-processing and interpretation. Efforts have been made to model LUC with the projection of possible future scenarios for spatial patterns of change to provide solutions and assist to land management. In the published research literature various LUC modeling methods have been reported such as Markov chains (Kumar, Radhakrishnan, & Mathew, 2014; Muller & Middleton, 1994); cellular automata (Batty & Xie, 1994; K C Clarke, Hoppen, & Gaydos, 1997; Keith C. Clarke & Gaydos, 1998; White & Engelen, 1997; Wu & Webster, 1998); neural networks (X. Li & Yeh, 2002; Lin, Lu, Espey, & Allen, 2005; Pijanowski, Brown, Shellito, & Manik, 2002); logistic regression (Cheng & Masser, 2003); multi-agent systems (Brown, Page, Riolo, Zellner, & Rand, 2005; Sanders, Pumain, Mathian, Guérin-Pace, & Bura, 1997); and machine learning (Huang, Xie, Tay, & Wu, 2009; Nemmour & Chibani, 2006; Otukei & Blaschke, 2010; Samardzic-Petrovic, Dragicevic, Kovacevic, & Bajat, 2016).

Markov chains is a stochastic model that can capture time dependency among sequential data, and usually used to describe a sequence of possible events or states whose probabilities only depend on previous events or states. Markov chains cannot preserve the information from the event that is not within the neighborhood of a current event. There are some Markov chain-based models for representing the dynamics of LU systems (Muller & Middleton, 1994), which project the future LU by applying transition

probability matrix on the primary matrix recording LU information. Markov chains model assume that the transition probability between each pair of states is stationary over time; hence, these models can forecast LUC from short to long-term. However, Markov chains models cannot consider socioeconomic and human related factors that can potentially lead to changes in LU patterns.

Cellular automata (CA) is a discrete modeling approach that has been used for representing LUC given their capability to capture both spatial and temporal dynamics of the phenomenon and considering the changes at a very local scale (Batty, Xie, & Sun, 1999; Chaudhuri & Clarke, 2013; White, Engelen, & Uljee, 1997). The CA consists of a regular grid of cells, where each cell has one of many finite states, and the states of a cell change in the next time iteration according to the function of transition rules based on the state of the cell and in its spatial neighborhood. The structure of CA models has a close affinity with raster-based GIS and RS datasets. However, Stevens and Dragićević (2007) proposed a LUC model using irregular CA cells although this requires longer computation time. The integration of Markov and CA allows the simulation of spatial and temporal LUC processes (Guan et al., 2011). Even though the fixed rules enable the various possibilities of transition, the forecast of LUC is more precise when the system is stable over years and under the assumptions that the land always changes with the same transition rules. The assumption of an ideal and stable environment is not realistic as the LU changes are governed by human decisions changing over time that are difficult to predict. However, these types of models are more sensitive to spatial than temporal factors.

The machine learning (ML) methods depend on strong statistical learning theory where the size and quality of the training datasets significantly influence the performance of the ML methods. ML-based LUC models can extract and learn from earlier LU observations the driving forces of LUC and their impact. Otukei and Blaschke (2010) evaluated several ML methods, such as artificial neural networks (ANN), support vector machine (SVM), maximum likelihood, and decision trees to investigate LUC detection, and their study demonstrated better classification performance of SVM and ANN algorithms. Samardzic-Petrovic et al. (2017) compared the performance of the common ML methods for LUC short-term forecasting. Urban LUC was also modeled using decision trees (Samardžić-Petrović, Dragićević, Bajat, & Kovačević, 2015) and SVM (Samardzic-Petrovic et al., 2016). However, traditional ML methods have limited

performance while the data is high-dimensional and the number of observations are large (Arel, Rose, & Karnowski, 2010).

Post-classification comparison is the strategy of some ML-based LUC detection methods. The change analysis of the multi-temporal images generally employs two basic methods: raster-to-raster comparison and post-classification comparison (Mukherjee, Shashtri, Singh, Srivastava, & Gupta, 2009). As other ML methods assume the independence of data (Lipton, 2015), however, spatial data are often known for their dependency and spatial autocorrelations. Therefore, using DL models compared to ML, brings the advantage of automating the extraction of representations (abstractions) from a larger amount of data. The success of DL models has started to attract attention for studies of LUC classification using RS datasets with CNNs.

Recently, some studies have shown the effectiveness of RNN for analyzing LUC. Byeon et al. (2015) conducted LU scene classification with LSTM instead of CNN without pre-processing, whose result was comparable to that of CNNs. This study indicates that the LSTM model can learn the spatial neighboring context information for every raster and capture the global dataset dependency through the recurrent connections. Using sequence-to-sequence processing of LSTM models, Rußwurm & Körner (2017a) classified land cover through learning from multi-temporal land cover RS datasets. The CNN, RNN and LSTM models were used to model vegetation from temporal RS data, LSTM outperformed CNN as temporal information was used in training (Rußwurm & Körner, 2017b). Due to the realistic dynamics of fast-developing urban areas, it is necessary to propose more efficient models to capture the change and study the LUC from a spatio-temporal perspective. Bengio (2009) used CNN and RNN for learning long-term dependencies and functions from complex phenomenon. The RNN models can consider larger numbers of data layers and thus be more effective than traditional ML methods such as SVM or decision tree (DT) to name a few. So far, RNN has recently been used for land cover classification (Ienco, Gaetano, Dupaquier, & Maurel, 2017; Rußwurm & Körner, 2018; Sharma, Liu, & Yang, 2018), land cover change detection (Lyu, Lu, & Mou, 2016; Mou & Zhu, 2018). Du et al. (2018) has used RNNs for spatio-temporal modeling of LUC and the overall accuracy results were close to 50%. Therefore, there is a need to further explore the potential of RNNs to forecast LUC. In this research study, previous geospatial LU datasets with 5-year intervals for the

City of Surrey, British Columbia, Canada were used to implement the concepts for RNN-based short-term forecasting of LUC.

3.2.2. RNN and its variants

Artificial Neural Networks (ANN) are the collections of connected neurons (also called layers) inspired by the human brain (McCulloch & Pitts, 1943). CNN and RNN are types of ANN. Unlike CNN, which contains different types of layers that perform different functions (e.g. convolutional, pooling and nonlinear layers), the basic RNN consists only of recurrent layers. Each layer in an RNN shares the same group of functions and parameters, while the parameters are updated in each layer. As shown in Figure 3-1, the connections between recurrent layers are cyclical when presented compactly (Graves, 2012), for the convenience of visualization, it can be unfolded like a chain-like structure.

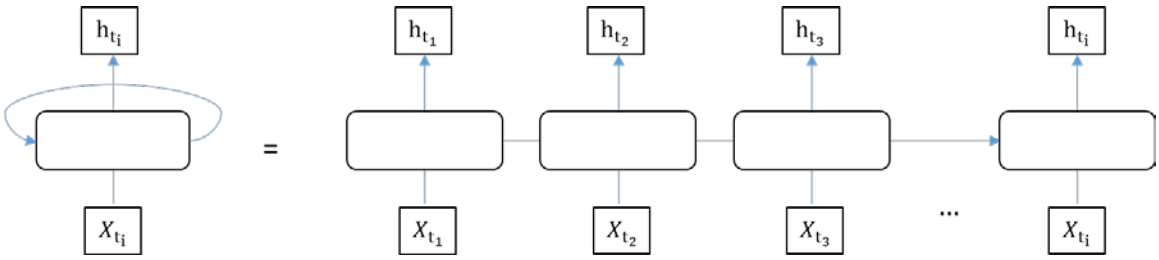


Figure 3-1 A simple RNN structure allowing information to loop in the layer, it can be unfolded as a neural network indicated at the right side. X_{t_i} is a temporal sequence input. h_{t_i} is the hidden state.

CNNs are often used for image classification tasks. The inputs and outputs of CNNs have fixed sizes and are processed independently among each other. By contrast, RNNs can deal with sequential inputs and have sequential outputs, while the inputs are considered dependent so that RNNs can capture their dependence. In reality, many datasets such as text, speech, audio, video, weather and stock price are sequential and internally dependent. Some applications areas of RNNs and its variants include for example music composition (Eck & Schmidhuber, 2002); handwriting recognition (Graves & Schmidhuber, 2009), speech synthesis (Zen & Sak, 2015), and video captioning (Yang et al., 2018), to name a few. The LSTMs are often used for time-series problems such as predicting stock market price movement (Nelson, Pereira, & Oliveira, 2017), weather (Qing & Niu, 2018), and traffic flow (Fernandes et al., 2019), passenger flow (Han et al., 2019).

The traditional RNN has the limitation of short-term memory caused by a vanishing gradient problem (Y. Bengio, Simard, & Frasconi, 1994). Gradients are used to update the weights of networks, which shrink through time during the backpropagation process and may become too small to contribute new significant weights based on equation (3-1):

$$\text{New weights} = \text{old weights} - (\text{learning rate} \times \text{gradients}) \quad (3\text{-}3)$$

The problem of short-term memory of traditional RNN can be solved by Long-Short Term Memory (LSTM) (Hochreiter & Schmidhuber, 1997). The LSTM is one of the variants of RNN, which has unique internal mechanisms called gates that can regulate the flow of information compared with traditional RNN. The function of gates is to decide if the data in a sequence is important or not, then to keep or discard the information from that data. Through gates, the essential information can be preserved even if the sequential data is long. Gated Recurrent Unit (GRU), another variant of RNN, was first introduced by Cho (2014). GRUs have fewer gates and relatively shorter memory but have faster training processes than LSTMs. GRU has comparable performance to LSTM for music and speech modeling tasks (Chung, Gulcehre, Cho, & Bengio, 2014). Bidirectional LSTM (BiLSTM) (Schuster & Paliwal, 1997) processes two sequential inputs with the opposite direction so that current BiLSTM layer has two hidden states which accept past information and future information. Besides, the algorithms within BiLSTM layers are the same as that of unidirectional LSTM. BiLSTM adapted with more complicated situations such as speech recognition (Graves, Jaitly, & Mohamed, 2013), phoneme recognition (Graves, Fernández, & Schmidhuber, 2005), where the current inputs are influenced by previous and future inputs.

The inputs to RNNs like LSTM and GRU are one-dimensional (1D) vectors, while a Convolutional LSTM (ConvLSTM) network receives inputs as 3D vectors which can encode both spatial and temporal information. ConvLSTM is an effective model for nowcasting short-term precipitation within a study area (Shi et al., 2015). There also are ConvLSTM based models for short-term forecast of traffic accidents (Yuan, Zhou, & Yang, 2018), video anomaly detection (Luo, Liu, & Gao, 2017) and short-term forecast of traffic flow (Y. Liu, Zheng, Feng, & Chen, 2017), and these phenomena have spatial patterns that can be captured by convolutional layers. When using LSTM, it is assumed

that each cell is independent. Figure 3-2a presents the structure of a ConvLSTM as a group of cells that are located in the same neighborhood. ConvLSTM can simultaneously incorporate the spatial neighborhood for each raster cells and the temporal LU information. The GRU model is very similar to LSTM, however, the differences of GRU is in the update gate and reset gate where the update gate learns and decides how much of the past information to pass to the future, reset gate decides how much of the past information to forget (Figure 3-2c). BiLSTM (Figure 3-2d) can be considered to have the inputs with the original order and reversed order respectively to feed into the LSTM. RNNs and LSTMs can receive complex sequential inputs or form as a hybrid model with other layers or networks (Fan, Wang, Soong, & Xie, 2015; S. Li, Li, Cook, Zhu, & Gao, 2018; Lv et al., n.d.).

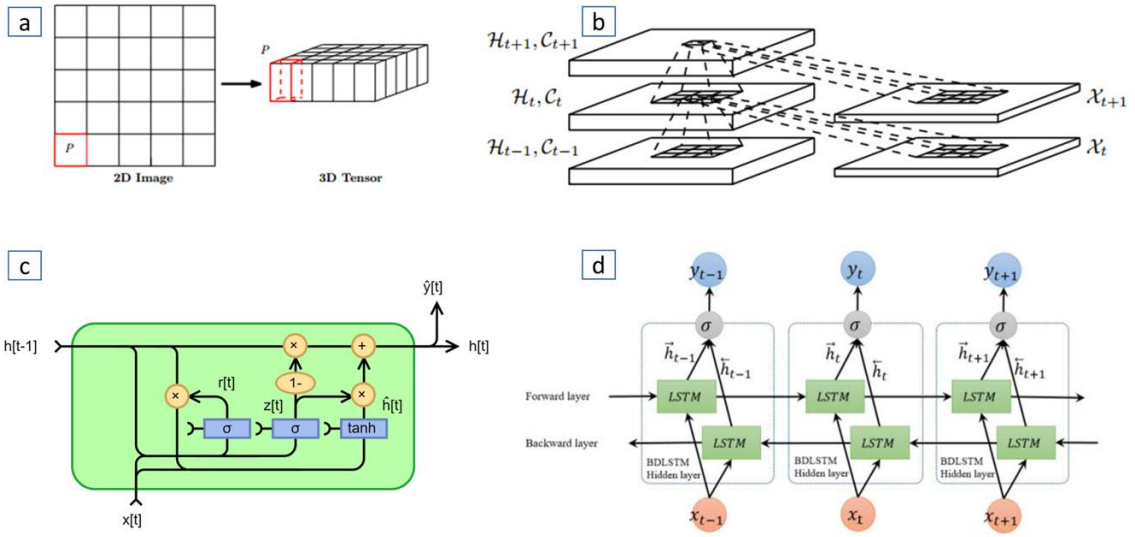


Figure 3-2 Structure of the ConvLSTM (Shi et al., 2015) model with (a) transforming the 2D image into 3D tensor; and (b) its inner structure, and the structure of (c) the GRU (Jeblad, 2018) ; (d) BiLSTM (Cui, Ke, & Wang, 2018) models

LSTM is more frequently used than traditional RNN due to its longer memory capabilities. The key elements of LSTMs are the cell state, sigmoid activation, forget gate, input gate, and output gate and tanh activation (Figure 3-3) which control the relevant information through the network. At each processing step, the gates regulate the addition and removal of information to the cell state. Gates have sigmoid activation, which multiply values between 0 and 1 to derive the percentage of data that will be kept or removed. If the input multiplies with 0, the information is forgotten; if the input

multiplies with “1”, the information is remembered. Tanh activation delivers values between -1 and 1. Hochreiter and Schmidhuber (1997) provide the formulae (3-2~3-7) to describe the algorithms of a typical LSTM layer as follows:

$$i_t = \sigma\,(w_i[h_{t-1}, x_t] + b_i) \tag{3-4}$$

$$f_t = \sigma\,(w_f[h_{t-1}, x_t] + b_f) \tag{3-5}$$

$$o_t = \sigma\,(w_o[h_{t-1}, x_t] + b_o]) \tag{3-6}$$

$$g_t = \tanh\,(w_g[h_{t-1}, x_t] + b_g]]) \tag{3-7}$$

$$c_t = f_t{*}c_{t-1}{+}i_t{*}\,g_t \tag{3-8}$$

$$h_t = o_t{*}\tanh(c_t) \tag{3-9}$$

where x_t, c_t, h_t represents the input, cell state, output at time step t. f_t is forget gate, i_t is input gate, o_t is output gate, σ is sigmoid function, w and b are weight and bias respectively. g_t is a vector of new candidate value called cell activation, which adds with current cell state. f_t is a value between 0 and 1, which means the ratio of old information that will be passed to new cell state, i_t decides the ratio of each value in a sequence from g_t that will be preserved.

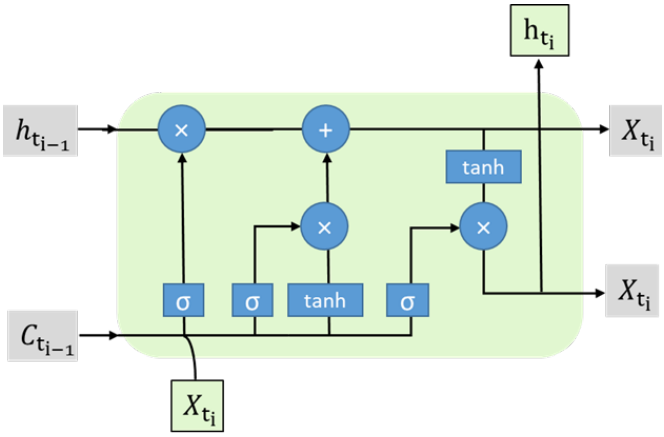


Figure 3-3 The graphical representation of an LSTM layer
(The idea of image is based upon Colah, 2015, the image was self-created)

3.3. Methodology

3.3.1. Study area

The City of Surrey, British Columbia, Canada is one of the fast-growing municipalities in the Metro Vancouver Region. Significant population growth occurred between 2007 and 2017, and the population is estimated to increase by over 262,000 inhabitants from 2018 to 2046 (City of Surrey, 2018). The increased population will mean considerable challenges related to urban residential development, the management of the lands and the natural environment. The study area of the City of Surrey covers 316.4 km² (City of Surrey, 2019) (Figure 3-4).

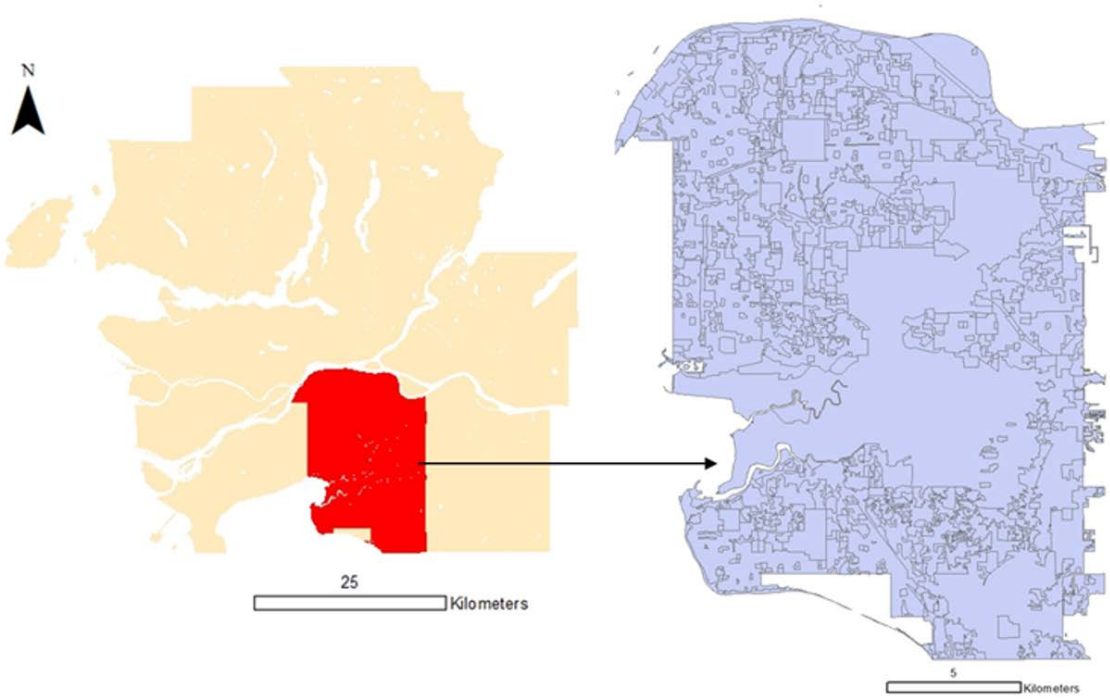


Figure 3-4 The City of Surrey located in the south of the Metro Vancouver Region.

3.3.2. Data preparation

The generalized LU data was obtained from the Metro Vancouver Open Data Catalogue (Metro Vancouver, 2011) for years 1996, 2001, 2006 and 2011. The road network data was obtained from the CanMap streetfiles (DMTI Spatial Inc., 2011). The

LU classes and road networks were rasterized at 10m spatial resolution and data processing was done within the ArcGIS desktop software (Esri, 2018).

The RNN models examined spatial and temporal features from 1996, 2001, 2006 and 2011 LU datasets and forecasted the 2016 LU pattern. Due to the different classification schemes, the LU data of each year has a different total number of LU classes. For example, the year 1996, 2001, 2006, and 2011 LU data have respectively 13, 12, 15, 22 LU classes, and specifically, 2011 have more varieties of LU classes. In order to create uniform datasets with the same group of LU classes, the LU data from 1996, 2006, 2011 were reclassified and merged based on the LU classes in 2001 LU data, and then similar types of residential classes (eg. rural, single, townhouse, high-rise) are merged into one LU class as residential. A total of 9 LU classes were considered and these are: transportation, communication and utilities; recreation and protected natural areas; industrial; open and undeveloped land; residential; lakes and water bodies; institutional; commercial; and agricultural. The 1996 and 2001 LU data contained no major road information and they were combined with rasterized road networks for the same year obtained from DMTI Spatial Inc.

RNNs usually process sequential inputs and can have multiple outputs. It has been proven that even if the data is not in the form of sequences, it can be formatted as sequences and be used to train RNN models (Karpathy, 2015). Consider a study area V consists of m×n raster cells (m=2437, n=1952), V={$c_{1,1}, c_{1,2}, c_{1,3}, \dots\ c_{i,j}, \dots\ c_{m,n},$}, with associated LU label L ={$l_{1,1}, l_{1,2}, l_{1,3}, \dots\ l_{i,j}, \dots, l_{m,n},$}, (i, j) indicates the raster cell at row i and column j. Since the LU class of each cell is influenced by its surrounding cell state, another two raster layers with the size of m×n were created. Each raster cell in one layer stores the most frequently occurring LU class in its adjacent 7×7 cells as Moore neighborhood, and each raster cell in another layer stores the second most frequently occurring LU class in its adjacent 7×7 cells as Moore neighborhood, represented as L_{moore}= {$l^1_{1,1}, l^1_{1,2}, l^1_{1,3}, \dots, l^1_{i,j}, \dots, l^1_{m,n}$} and L_{moore2}= {$l^2_{1,1}, l^2_{1,2}, l^2_{1,3}, \dots, l^2_{i,j}, \dots, l^2_{m,n}$} respectively.

3.3.3. Training and validation of RNNs

The training and validation of RNNs are similar to other neural networks. Through repeated forward-propagation and back-propagation, parameters are updated until the cost function is minimized. The validation process is part of training the model and

updating the parameters, which uses a small part of datasets to validate and update the model parameters after each training epoch. When performing a classification task, categorical cross entropy loss is usually used as a cost function. The key approach to ensure the model learning from data correctly is minimizing a cost function during the training and validation process. Supposing K categories are expected from the model. There is a certain sample x and its true label is represented as vector $[\bar{y}_1, \bar{y}_2, \dots, \bar{y}_i, \dots, \bar{y}_k]$, where $\bar{y}_i$ can be represented as equation (3-8):

$$\bar{y}_i = \begin{cases} 1, & if\ x \text{ belongs to the } i^{th} \text{ category} \\ 0, & if\ x \text{ not belongs to the } i^{th} \text{ category} \end{cases} \quad (3\text{-}8)$$

The output from the model y is a vector $[y_1, y_2, \dots, y_i, \dots, y_K]$, where y_i is the forecasted probability of sample x being the ith category. Cross Entropy Loss is defined in equation (3-9) (Rubinstein & Kroese, 2004):

$$C(y, \bar{y}) = -\sum_{i=1}^{K} \bar{y}_i \, log(y_i) \quad (3\text{-}9)$$

The softmax layer (equation 3-10) is used to transform the outputs (i.e. K dimensional vector) from last layer to vector $\bar{y}$ with each value ranging between 0 and 1, which shows the probability distribution of K categories (Bishop, 2016; Bridle, 1990).

$$\sigma(Z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad (3\text{-}10)$$

The original set of raster cells from the study area were split by the ratio of 8:2 according to Pareto principle (Box & Meyer, 1986). This is a common starting point for splitting training data sets, as there are no strictly defined rules for dataset splitting. Further investigation of split ratio were conducted by Guyon (1997). Usually, when the total number of training dataset is more than 100000, split ratio such as 7:3 (Gholami, Chau, Fadaie, Torkaman, & Ghaffari, 2015; Gholami et al., 2015) or 9:1 will have small impact on model accuracy. The ratio of 5:5 was not considered because it will be more suitable when cross validation method was used. While in this study, 80% of the raster cells were used for training of the model so the parameters of the model were updated during each training epoch to minimize Cross Entropy Loss. At the same time, the remaining 20% of the raster cells were used to evaluate the models after each training epoch by measuring the validation accuracy. The validation accuracy is the percentage of raster cells in the validation dataset that fit the model after each training epoch, it is also calculated based on Cross Entropy (equation 3-9). If the validation accuracy is low,

the Cross Entropy will be fed back to the model in the next training epoch and adjust the configuration of the model parameters.

3.3.4. LTSM implementation

Figure 3-5 outlines the flowchart of the proposed LUC model for short-term forecast based on the LSTM model and spatio-temporal data available for the study area.

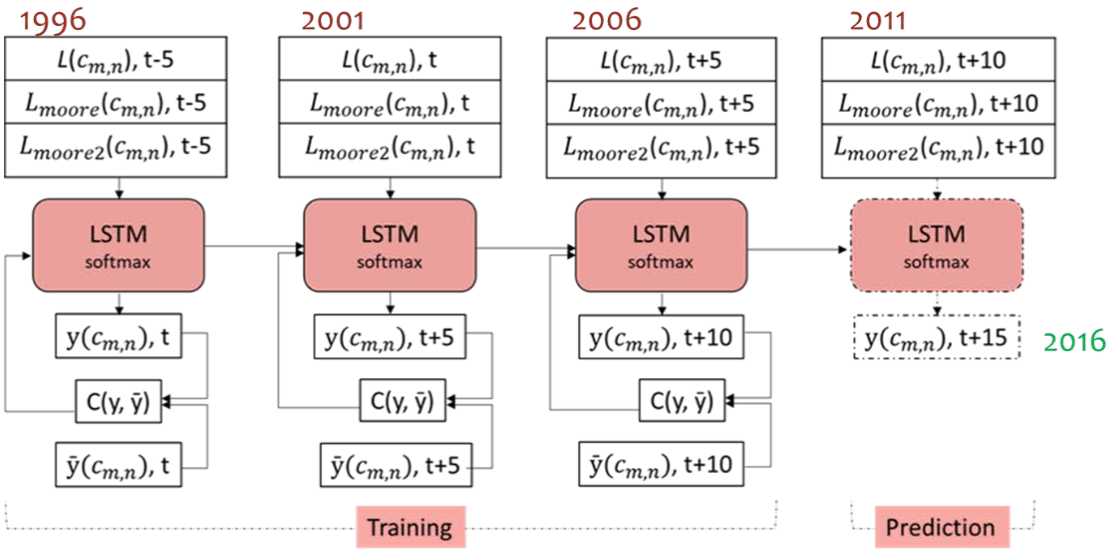


Figure 3-5 Flowchart of the proposed LSTM models for LUC forecast, where LU data were considered for years t-5=1996, t=2001, t+5=2006 and t+10=2011 for training and validation of the LSTM and then forecasted for t+15=2016.

In this study, four variants of RNN models were tested; specifically the LSTM, GRU, BiLSTM, and ConvLSTM models. The LSTM, GRU, and BiLSTM models have similar training methods where all raster cells are considered independent. The inputs to the models were encoded as 3D vectors with the shape of [samples, time steps, features], where samples are equal to the number of raster cells for training and validation; time steps refers to years of 1996, 2001 and 2006; features refer to L, L_{moore} and L_{moore2} of different Moore neighborhoods. Only a small part of all raster cells has changed their LU classes in the period from 1996 to 2011. In order to incorporate information regarding changed raster cells while training the RNN, two groups of datasets were used for every model for training and for validation. Group one

sample set consisted of only raster cells that have changed from 1996 to 2011, and group two sample set contained all the raster cells. The inputs to the ConvLSTM layer had the shape of [samples, timesteps, rows, cols, features], where rows and columns represented the sample size for ConvLSTM. This means every rows x columns of raster cells are grouped as one sample for training and validation so that LU information of nearby raster cells were considered.

3.3.5. Testing the forecasted results

In order to check the accuracy of the forecasted LUC for the City of Surrey, 2016 orthophoto images with 10cm resolution obtained from Surrey Open Data catalog (City of Surrey, 2017) were used for reference. There are about three million raster cells whose LU classes have been forecasted, however, checking the correctness of every raster cell was time-consuming and complicated. A total of 604 sample points were randomly chosen on the orthophotos from the datatsets to reduce the computational workload but still achieve a high confidence level. The necessary sample size was decided based on Z-score from equation (3-11) (Isreal, 1992):

$$Necessary\ Sample\ Size = (Z-score)^2 * StdDev * (1 - StdDev) \,/\, (margin\ of\ error)^2 \quad \text{(Isreal, 1992)} \tag{3-11}$$

where margin of error is 4%, StdDev is 0.5, Z-score is 1.96, and confidence level is 95%. Each point i (i∈[1,2,3,…,604]) has corresponding location S_i on the orthophoto and the ones forecasted LUC for the year 2016, Ytrue(S_i) represented the manually classified LU class of sample point i on the orthophoto, and Ypred(S_i) represented the forecasted LU class by LSTM of location S_i. Therefore the LU forecast for the year 2016 was tested by comparing Ytrue(S_i) and Ypred(S_i) of 604 sample points, and if Ytrue(S_i)=Ypred(S_i) the forecast accuracy is considered 1, otherwise is 0.

The total accuracy indicator is used for the evaluation of the performance of the LU forecast model and calculated based on Cohen's Kappa coefficient (Cohen, 1960):

$$\text{Total accuracy} = \frac{TurePositive + TrueNegative}{TurePositive + TrueNegative + FalsePostive + FalseNegative} \tag{3-12}$$

where true positive (TN) = correctly forecasted, false positive (FP) = incorrectly forecasted, true negative (TN) = correctly rejected, false negative (FN) = incorrectly rejected raster cells. The total number of samples used for calculating the kappa coefficient is 604. A confusion matrix (Stehman, 1997) is typically used for describing the performance measurement for classification models and from which kappa coefficient can be calculated.

The implementation of the proposed methodology and RNN models was done in the MATLAB software (The MathWorks Inc., 2018) for data preprocessing. The Python programming language (Python Software Foundation, 2016) was used for implementing and training RNNs, and the Keras API (Fran, 2015) was used for constructing the RNN models. ArcGIS (Esri, 2018) was used to create LU output maps.

3.4. Results

Table 3-1 provides the obtained values for model accuracy for six different scenarios. Most of the scenarios of RNN models provided accuracies above 0.86 except for LSTM1 where the accuracy was only 0.62. The total number of raster cells were split into a training set and validation set in an 8:2 proportion. As indicated in Table 3-1, scenario 1 and 2 both used the LSTM model with the same configuration. Scenario 1 (LSTM 1) used only raster cells whose LU classes have changed during 1996 and 2011 for training and validation while scenario 2 (LSTM 2) used all raster cells from the study area for training and validation. Scenario 3 used the GRU model and scenario 4 used the BiLSTM model. Scenario 4 and 5 used ConvLSTM where ConvLSTM 1 received input data with shape of 10x10 raster cells, ConvLSTM 2 received input data with a shape of 5x5 raster cells. Scenarios (2-6) used all raster cells from the study area for training and validation. When only the changed raster cells were used for LSTM training (scenarios 1), the overall accuracy is lower than the LSTM trained by all raster cells both changed and unchanged (scenarios 2). Scenarios 2-4 used the LSTM, GRU, and BiLSTM models respectively and have obtained comparative accuracy when using the training data containing all the raster cells from the study area. In the cases of the ConvLSTM models from scenarios 5 and 6, there is no obvious difference in accuracy when the sample size is different. Training accuracy represents the percentage of total training set that forecasted LU label equal the actual LU label. Validation accuracy

represents the percentage of total validation set that forecasted LU label equal the actual LU label.

Figure 3-6 shows the obtained LU for the City of Surrey for the year 2016, generated by short-term forecast using LSTM 2 which was trained by using all raster cells from the study area. The percentages of each LU class as forecasted are: 10.71% transportation, communication and utilities, 9.25% recreation and protected natural areas, 2.87% industrial, 2.13% open and undeveloped, 22.99% residential, 0.02% lakes and water bodies, 0.92% institutional, 0.99% commercial, and 16.01% agricultural. As forecasted, only 4.5% of raster cells have changed their LU classes compared with the 2011 LU. Figure 3-7 shows the changed raster cells in 2011 and 2016, it can be seen that some industrial in northwest part of the study area became transportation, and some open area become industrial (figure 3-7a). Some of the natural and protected areas were predicted as transportation LU class (figure 3-7b). While some of the agricultural areas were forecasted as residential (figure 3-7c). Based on the prediction, the increased LU classes during 2011 and 2016 were transportation and residential while the other LUs were decreasing. Specifically, 28.1% of changed raster cells changed from recreational and protected natural LUs to transportation LUs; 11.9% of changed cells with agricultural LUs turned into residential LUs. 18.6% undeveloped and open LUs became residential LUs. 6.1% of commercial LUs became industrial LUs.

Figure 3-8 presents the Confusion Matrix calculated based on Ytrue(S_i) and Ypred(S_i). The rows indicated the target LU classes Ytrue(S_i) that are manually classified from sample location (S_i) on 2016 orthophoto. The columns indicated the predicted LU classes Ypred(S_i) at corresponding sample locations (S_i) (i=1,2,3…604) from the forecasted model outputs. Moreover, in the confusion matrix value of cell (i,j) (i,j=1,2,3…9), where 1-9 indicate number of LU classes, represents the percentage of LU class j that was forecasted as class i. The total accuracy of LU forecast is 87%, while the TN, TP, FN, FP of LU forecast in each LU class differs. Especially the true negative from LU classes such as “open and undeveloped land” and “lakes and water bodies”, “commercial” are low, close to 50%. Percentage of “water bodies” decreased while percentage of “residential” and “agricultural” increased as indicated in confusion matrix. Some misrepresentation of forecasted LU classes within the sample set are: 1.7% for “transportation” were forecasted as “residential”; 1.2 % of “residential” were forecasted as “open and undeveloped land”, this means the growth of urban residential areas were

likely overestimated. The rate of urbanization is driven by various factors related to human interactions and decision making.

Table 3-1 Training and validation accuracy of different RNN models and scenarios

Scenarios	RNN variants	The ratio of changed raster cells in training set	Training accuracy	Validation accuracy
(1)	LSTM 1	100%	0.62	0.62
(2)	LSTM 2	47%	0.87	0.87
(3)	GRU	47%	0.86	0.87
(4)	BiLSTM	47%	0.87	0.87
(5)	ConvLSTM 1, 10×10	47%	0.88	0.88
(6)	ConvLSTM 2, 5×5	47%	0.88	0.88

3.5. Discussion and conclusions

LUC is a spatiotemporal phenomenon and it can be correlated with various factors. Consequently, forecasting LUC is a challenging topic and extensive efforts have been dedicated to modeling LUC while only a few studies have explored the potential of DL models on this topic. While RNN has been shown to efficient to solve time-series data, the objective of this study was to test the feasibility of RNN based models for short-term LUC forecasting.

This study successfully tested several RNN based models to examine the LUC of the City pf Surrey. The LSTM, GRU, BiLSTM and ConvLSTM were trained by changed raster cells and persistent raster cells. Then the LU for the year 2016 was forecasted using LSTM, which was trained first by LU data for years 1996, 2001, 2006, and 2011 at 10-meter spatial and 5-year temporal resolutions. Due to the limitation of availability of classified LU data, the forecasted LU of the year 2016 was not fully validated in this study. Instead, it was compared with the actual 2011 LU data and changes of each LU class were analyzed. The LSTM can estimate how each LU will change based on “transition rules” which were learned from the 1996 to 2011 LU data.

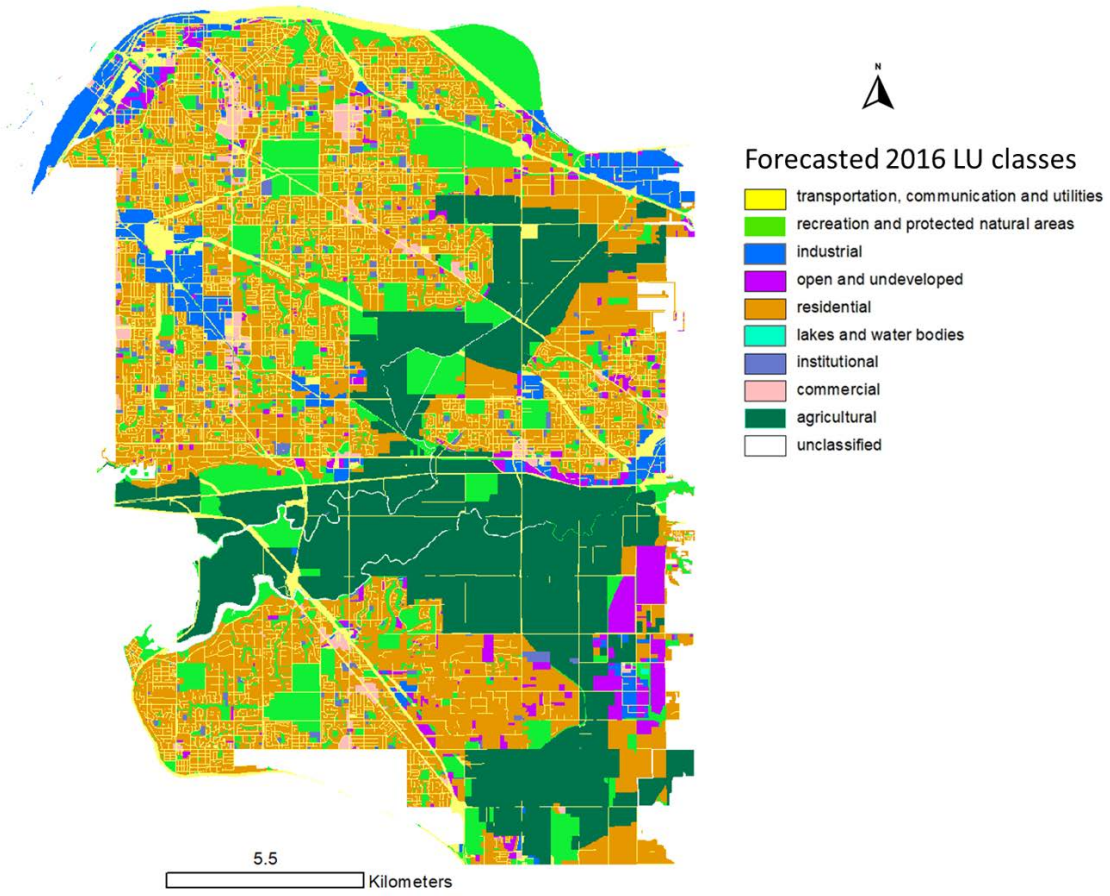


Figure 3-6 Forecasted LU for 2016 for the City of Surrey generated by LSTM 2 model.

The RNNs were successfully trained by LU data as well as being able to forecast the 2016 LU. The model accuracy turns out to be similar at about 86% while trained by both changed and persistent raster cells. Model accuracy was lower at about 62% while been trained by only changed cells which indicates the variants of RNN did not differ much when using the same LU datasets. The forecasted results indicated only 4.5% of the land in Surrey City have changed. The forecasted changes mainly occurred among industrial, recreational, transportation, open, agricultural and residential LUs. The results indicate the land was not overdeveloped between 2011 and 2016, while the City of Surrey experienced an increase in population and in transportation network.

Kappa statistic (Congalton, 1991) is often used as an assessment indicator to compare the similarity between observed and predicted results (Monserud & Leemans, 1992; Rußwurm & Körner, 2017). In this study, simple random sampling (SRS) was used

to select samples for evaluation from the study area. However, due to the fact that the majority of the raster cells remain unchanged, SRS is not sufficient to evaluate the overall predication performance of the RNN models. Instead, cluster sampling (CS), stratified random sampling (STRAT), mean of surface with non-homogeneity (MSN) method (Hashemian, Abkar, & Fatemi, 2004; Mu et al., 2015) are potential sampling methods to properly evaluate the results and models. If the appropriate data were available for year 2016 the evaluation of the accuracy of the obtained short-term forecasted LU could be performed with a variety of exiting methods for map comparisons (Stehman, 1997).

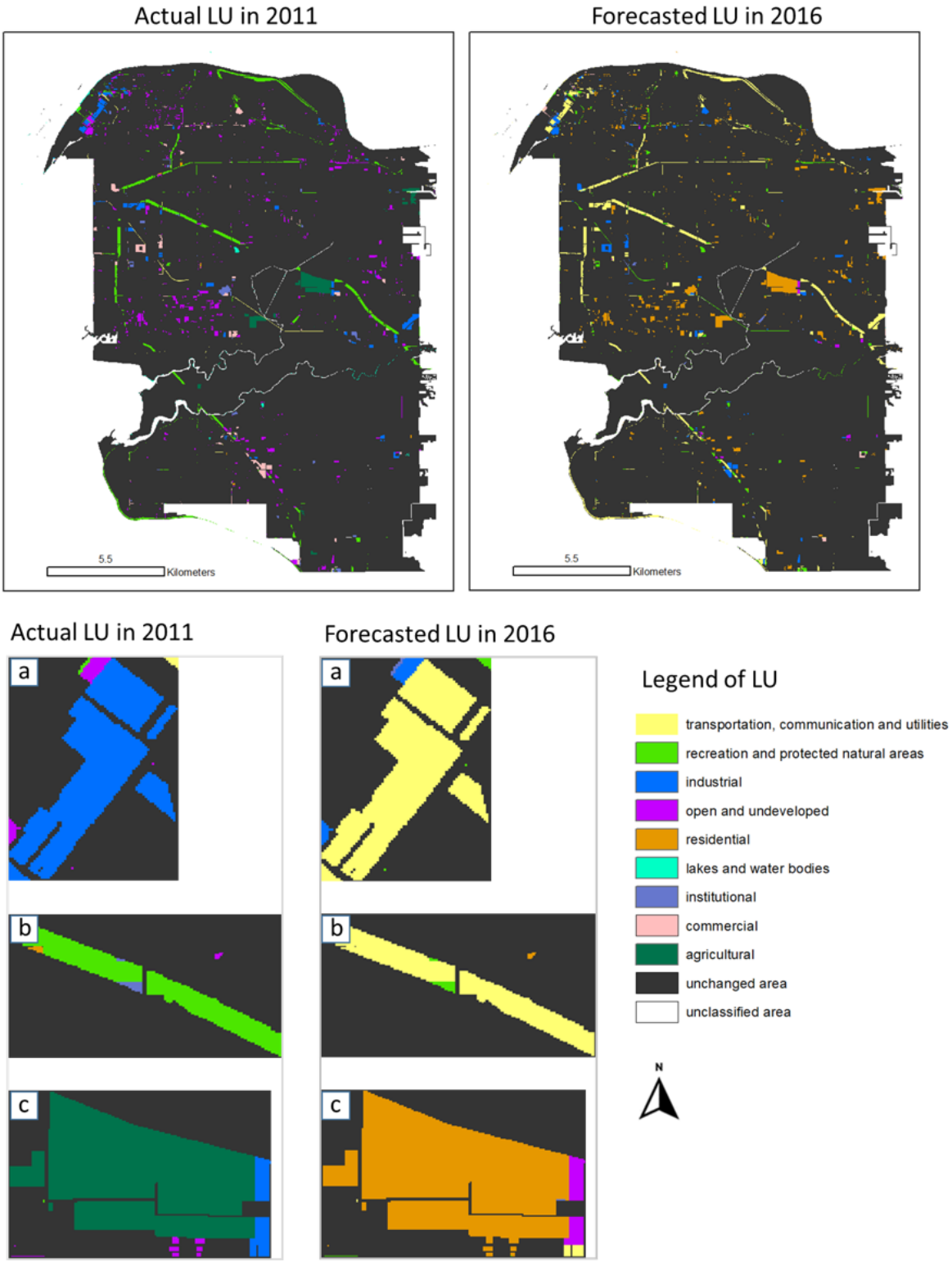


Figure 3-7 Changed LU raster cells in 2011 and 2016 as forecasted by LSTM with detailed subsections (a), (b), and (c)

RNNs are an efficient method capable to perform feature extraction and classification separately. So far, few studies have exploited RNN for LUC forecasting. This study has demonstrated that RNNs have this potential although the performance of LU forecasting still needs strict validation and further study. In summary, in this study RNNs facilitated the full automation of the prediction process from available geospatial datasets due to the learning abilities of temporal correlations of RNN. Expert knowledge is not required to initialize the models and interpret the results. Overall, RNNs have the potential to capture the spatio-temporal patterns of LUC and provide consistent short-term forecast.

Figure 3-8 The Confusion Matrix based on the comparison of the forecasted LU and orthophotos for the year 2016.

This study has provided an effective way to generate informative LU data. Due to the fact that RNN is a data-driven model, the quality and quantity of the training data are important factors of the forecast accuracy. Thus, training data is also a challenge in this research. Incorporating proximate physical, local, demographic, socioeconomic and climatic factors into the training process, the RNNs can better learn the transition rules. However, not every type of data is available and open to the public. The 2016 LU data

from an official source is currently unavailable. If the RNNs are proven efficient at LUC modeling, they can be a useful tool for geographers to study LUC, which can further benefit scientists and decision-makers in land use decision making and management.

3.6. References

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). A land use and land cover classification system for use with remote sensor data (USGS Numbered Series No. 964). Retrieved from http://pubs.er.usgs.gov/publication/pp964

Arel, I., Rose, D. C., & Karnowski, T. P. (2010). Deep Machine Learning—A New Frontier in Artificial Intelligence Research. IEEE Computational Intelligence Magazine, 5(4), 13–18. https://doi.org/10.1109/MCI.2010.938364

Batty, M., & Xie, Y. (1994). From Cells to Cities. Environment and Planning B: Planning and Design, 21(7), S31–S48. https://doi.org/10.1068/b21S031

Batty, M., Xie, Y., & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. Computers, Environment and Urban Systems, 23(3), 205–233. https://doi.org/10.1016/S0198-9715(99)00015-0

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2), 157–166. https://doi.org/10.1109/72.279181

Bengio, Yoshua. (2009). Learning Deep Architectures for AI. Foundations and Trends® in Machine Learning, 2(1), 1–127. https://doi.org/10.1561/2200000006

Bishop, C. M. (2016). Pattern Recognition and Machine Learning. Springer New York.

Box, G. E. P., & Meyer, R. D. (1986). An Analysis for Unreplicated Fractional Factorials. Technometrics, 28(1), 11–18. https://doi.org/10.1080/00401706.1986.10488093

Bridle, J. S. (1990). Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition. In F. F. Soulié & J. Hérault (Eds.), Neurocomputing (pp. 227–236). Springer Berlin Heidelberg.

Brown, D. G., Page, S., Riolo, R., Zellner, M., & Rand, W. (2005). Path dependence and the validation of agent‐based spatial models of land use. International Journal of Geographical Information Science, 19(2), 153–174. https://doi.org/10.1080/13658810410001713399

Byeon, W., Breuel, T. M., Raue, F., & Liwicki, M. (2015). Scene labeling with LSTM recurrent neural networks. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3547–3555. https://doi.org/10.1109/CVPR.2015.7298977

Chaudhuri, G., & Clarke, K. (2013). The SLEUTH land use change model: A review. International Journal Of Environmental Resource Research, 1, 88–105.

Cheng, J., & Masser, I. (2003). Urban growth pattern modeling: A case study of Wuhan city, PR China. Landscape and Urban Planning, 62(4), 199–217. https://doi.org/10.1016/S0169-2046(02)00150-0

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1724–1734. https://doi.org/10.3115/v1/D14-1179

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. NIPS 2014 Workshop on Deep Learning, December 2014. Retrieved from https://nyuscholars.nyu.edu/en/publications/empirical-evaluation-of-gated-recurrent-neural-networks-on-sequen

City of Surrey. (2017). Imagery—City of Surrey Open Data Catalogue. Retrieved September 10, 2018, from http://data.surrey.ca/group/6878e307-9fec-4134-b042-d7e058310255?tags=orthophoto

City of Surrey. (2018). Population Estimates & Projections [Text/xml]. Retrieved May 24, 2019, from http://www.surrey.ca/business-economic-development/1418.aspx

City of Surrey. (2019). City of Surrey [Text/xml]. Retrieved May 17, 2019, from http://www.surrey.ca/default.aspx

Clarke, K C, Hoppen, S., & Gaydos, L. (1997). A Self-Modifying Cellular Automaton Model of Historical Urbanization in the San Francisco Bay Area. Environment and Planning B: Planning and Design, 24(2), 247–261. https://doi.org/10.1068/b240247

Clarke, Keith C., & Gaydos, L. J. (1998). Loose-coupling a cellular automaton model and GIS: Long-term urban growth prediction for San Francisco and Washington/Baltimore. International Journal of Geographical Information Science, 12(7), 699–714. https://doi.org/10.1080/136588198241617

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, 20(1), 37–46. https://doi.org/10.1177/001316446002000104

Colah. (2015). Understanding LSTM Networks. Retrieved June 17, 2019, from https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37(1), 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Cui, Z., Ke, R., & Wang, Y. (2018). Deep Stacked Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction. 12.

De Palma, A., Sanchez-Ortiz, K., Martin, P. A., Chadwick, A., Gilbert, G., Bates, A. E., … Purvis, A. (2018). Challenges With Inferring How Land-Use Affects Terrestrial Biodiversity: Study Design, Time, Space and Synthesis. In D. A. Bohan, A. J. Dumbrell, G. Woodward, & M. Jackson (Eds.), Advances in Ecological Research (pp. 163–199). https://doi.org/10.1016/bs.aecr.2017.12.004

DMTI Spatial Inc. (2011). CanMap Streetfiles. Retrieved May 24, 2019, from https://www.dmtispatial.com/canmap/

Du, G., Yuan, L., Shin, K. J., & Managi, S. (2018). Modeling the spatio-temporal dynamics of land use change with recurrent neural networks. Retrieved from https://arxiv.org/abs/1803.10915

Eck, D., & Schmidhuber, J. (2002). Learning the Long-Term Structure of the Blues. In J. R. Dorronsoro (Ed.), Artificial Neural Networks—ICANN 2002 (pp. 284–289). Springer Berlin Heidelberg.

Fan, B., Wang, L., Soong, F. K., & Xie, L. (2015). Photo-real talking head with deep bidirectional LSTM. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4884–4888. https://doi.org/10.1109/ICASSP.2015.7178899

Fernandes, B., Silva, F., Alaiz-Moretón, H., Novais, P., Analide, C., & Neves, J. (2019). Traffic Flow Forecasting on Data-Scarce Environments Using ARIMA and LSTM Networks. In Á. Rocha, H. Adeli, L. P. Reis, & S. Costanzo (Eds.), New Knowledge in Information Systems and Technologies (pp. 273–282). Cham: Springer International Publishing.

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics, 36(4), 193–202. https://doi.org/10.1007/BF00344251

Gholami, V., Chau, K. W., Fadaie, F., Torkaman, J., & Ghaffari, A. (2015). Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers. Journal of Hydrology, 529, 1060–1069. https://doi.org/10.1016/j.jhydrol.2015.09.028

Graves, A. (2012). Supervised Sequence Labelling with Recurrent Neural Networks. https://doi.org/10.1007/978-3-642-24797-2

Graves, A., Fernández, S., & Schmidhuber, J. (2005). Bidirectional LSTM Networks for Improved Phoneme Classification and Recognition. In W. Duch, J. Kacprzyk, E. Oja, & S. Zadrożny (Eds.), Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005 (pp. 799–804). Springer Berlin Heidelberg.

Graves, A., Jaitly, N., & Mohamed, A. (2013). Hybrid speech recognition with Deep Bidirectional LSTM. 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, 273–278. https://doi.org/10.1109/ASRU.2013.6707742

Graves, A., Mohamed, A., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 6645–6649. https://doi.org/10.1109/ICASSP.2013.6638947

Graves, A., & Schmidhuber, J. (2009). Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), Advances in Neural Information Processing Systems 21 (pp. 545–552). Retrieved from http://papers.nips.cc/paper/3449-offline-handwriting-recognition-with-multidimensional-recurrent-neural-networks.pdf

Green, K., Kempka, D., & Lackey, L. (1994). Using Remote Sensing to Detect and Monitor Land-Cover and Land-Use Change. 60(3), 331–337.

Guan, D., Li, H., Inohae, T., Su, W., Nagaie, T., & Hokao, K. (2011). Modeling urban land use change by the integration of cellular automaton and Markov model. Ecological Modelling, 222(20–22), 3761–3772. https://doi.org/10.1016/j.ecolmodel.2011.09.009

Guyon, I. (1997). A Scaling Law for the Validation-Set Training-Set Size Ratio. AT & T Bell Laboratories.

Han, Y., Wang, C., Ren, Y., Wang, S., Zheng, H., & Chen, G. (2019). Short-Term Prediction of Bus Passenger Flow Based on a Hybrid Optimized LSTM Network. ISPRS International Journal of Geo-Information, 8(9), 366. https://doi.org/10.3390/ijgi8090366

Hashemian, M., Abkar, A., & Fatemi, S. (2004). STUDY OF SAMPLING METHODS FOR ACCURACY ASSESSMENT OF CLASSIFIED REMOTELY SENSED DATA.

Hegazy, I. R., & Kaloop, M. R. (2015). Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt. International Journal of Sustainable Built Environment, 4(1), 117–124. https://doi.org/10.1016/j.ijsbe.2015.02.005

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Huang, B., Xie, C., Tay, R., & Wu, B. (2009). Land-use-change modeling using unbalanced support-vector machines. Environment and Planning B-Planning & Design, 36(3), 398–416. https://doi.org/10.1068/b33047

Huang, B., Zhao, B., & Song, Y. (2018). Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. Remote Sensing of Environment, (214), 73–86.

Ienco, D., Gaetano, R., Dupaquier, C., & Maurel, P. (2017). Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks. IEEE Geoscience and Remote Sensing Letters, 14(10), 1685–1689. https://doi.org/10.1109/LGRS.2017.2728698

Isreal, G. D. (1992). Determining Sample Size (No. PEOD-6; pp. 1–5). University of Florida: Institute of Food and Agricultural Sciences (IFAS).

Jeblad. (2018). Gradient Recurrent Unit, fully gated version. Based on example in Recurrent Neural Network (RNN) – Part 5: Custom Cells. Retrieved from https://commons.wikimedia.org/wiki/File:Gated_Recurrent_Unit,_base_type.svg

Karpathy, A. (2015, May 21). The Unreasonable Effectiveness of Recurrent Neural Networks. Retrieved March 21, 2019, from Andrej Karpathy blog website: http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Kleemann, J., Baysal, G., Bulley, H. N. N., & Fürst, C. (2017). Assessing driving forces of land use and land cover change by a mixed-method approach in north-eastern Ghana, West Africa. Journal of Environmental Management, 196, 411–442. https://doi.org/10.1016/j.jenvman.2017.01.053

Kumar, S., Radhakrishnan, N., & Mathew, S. (2014). Land use change modelling using a Markov model and remote sensing. Geomatics, Natural Hazards and Risk, 5(2), 145–156. https://doi.org/10.1080/19475705.2013.795502

Li, S., Li, W., Cook, C., Zhu, C., & Gao, Y. (2018, March 13). Independently Recurrent Neural Network (IndRNN): Building A Longer and Deeper RNN. https://doi.org/10.1109/CVPR.2018.00572

Li, X., & Yeh, A. G.-O. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. International Journal of Geographical Information Science, 16(4), 323–343. https://doi.org/10.1080/13658810210137004

Lin, H., Lu, K. S., Espey, M., & Allen, J. (2005). Modeling Urban Sprawl and Land Use Change in a Coastal Area—A Neural Network Approach (No. 19364). Retrieved from American Agricultural Economics Association (New Name 2008: Agricultural and Applied Economics Association) website: https://ideas.repec.org/p/ags/aaea05/19364.html

Lipton, Z. (2015). A Critical Review of Recurrent Neural Networks for Sequence Learning.

Liu, J., Shahroudy, A., Xu, D., Kot, A. C., & Wang, G. (2018). Skeleton-Based Action Recognition Using Spatio-Temporal LSTM Network with Trust Gates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(12), 3007–3021. https://doi.org/10.1109/TPAMI.2017.2771306

Liu, Y., Zheng, H., Feng, X., & Chen, Z. (2017). Short-term traffic flow prediction with Conv-LSTM. 2017 9th International Conference on Wireless Communications and Signal Processing (WCSP), 1–6. https://doi.org/10.1109/WCSP.2017.8171119

Luo, W., Liu, W., & Gao, S. (2017). Remembering history with convolutional LSTM for anomaly detection. 2017 IEEE International Conference on Multimedia and Expo (ICME), 439–444. https://doi.org/10.1109/ICME.2017.8019325

Lv, Z., Xu, J., Zheng, K., Yin, H., Zhao, P., & Zhou, X. (n.d.). LC-RNN: A Deep Learning Model for Traffic Speed Prediction. 7.

Lyu, H., Lu, H., & Mou, L. (2016). Learning a Transferable Change Rule from a Recurrent Neural Network for Land Cover Change Detection. Remote Sensing, 8(6), 506. https://doi.org/10.3390/rs8060506

Mahoney, M. (2017). Large Text Compression Benchmark. Retrieved May 24, 2019, from http://www.mattmahoney.net/dc/text.html#1218

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics, 5(4), 115–133. https://doi.org/10.1007/BF02478259

Metro Vancouver. (2011). Open Data Catalogue [Government]. Retrieved December 21, 2018, from Metro Vancouver website: http://www.metrovancouver.org/data

Monserud, R. A., & Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. Ecological Modelling, 62(4), 275–293. https://doi.org/10.1016/0304-3800(92)90003-W

Mou, L., & Zhu, X. X. (2018). A Recurrent Convolutional Neural Network for Land Cover Change Detection in Multispectral Images. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 4363–4366. https://doi.org/10.1109/IGARSS.2018.8517375

Mu, X., Hu, M., Song, W., Ruan, G., Ge, Y., Wang, J., … Yan, G. (2015). Evaluation of Sampling Methods for Validation of Remotely Sensed Fractional Vegetation Cover. Remote Sensing, 7(12), 16164–16182. https://doi.org/10.3390/rs71215817

Mukherjee, S., Shashtri, S., Singh, C. K., Srivastava, P. K., & Gupta, M. (2009). Effect of canal on land use/land cover using remote sensing and GIS. Journal of the Indian Society of Remote Sensing, 37(3), 527–537. https://doi.org/10.1007/s12524-009-0042-6

Muller, M. R., & Middleton, J. (1994). A Markov model of land-use change dynamics in the Niagara Region, Ontario, Canada. Landscape Ecology, 9(2), 151–157. https://doi.org/10.1007/BF00124382

National Research Council. (2014). Advancing Land Change Modeling: Opportunities and Research Requirements. https://doi.org/10.17226/18385

Nelson, D. M. Q., Pereira, A. M., & Oliveira, R. A. de. (2017). Stock market's price movement prediction with LSTM neural networks. 2017 International Joint Conference on Neural Networks (IJCNN), 1419–1426. https://doi.org/10.1109/IJCNN.2017.7966019

Nemmour, H., & Chibani, Y. (2006). Multiple support vector machines for land cover change detection: An application for mapping urban extensions. ISPRS Journal of Photogrammetry and Remote Sensing, 61(2), 125–133. https://doi.org/10.1016/j.isprsjprs.2006.09.004

Otukei, J. R., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. International Journal of Applied Earth Observation and Geoinformation, 12, S27–S31. https://doi.org/10.1016/j.jag.2009.11.002

Pijanowski, B. C., Brown, D. G., Shellito, B. A., & Manik, G. A. (2002). Using neural networks and GIS to forecast land use changes: A Land Transformation Model. Computers, Environment and Urban Systems, 26(6), 553–575. https://doi.org/10.1016/S0198-9715(01)00015-1

Qing, X., & Niu, Y. (2018). Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy, 148, 461–468. https://doi.org/10.1016/j.energy.2018.01.177

Rubinstein, R. Y., & Kroese, D. P. (2004). The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning. Retrieved from https://www.springer.com/gp/book/9780387212401

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533. https://doi.org/10.1038/323533a0

Rußwurm, M., & Körner, M. (2017a). Multi-Temporal Land Cover Classification with Long Short-Term Memory Neural Networks. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42W1, 551–558. https://doi.org/10.5194/isprs-archives-XLII-1-W1-551-2017

Rußwurm, M., & Körner, M. (2017b, July 21). Temporal Vegetation Modelling using Long Short-Term Memory Networks for Crop Identification from Medium-Resolution Multi-Spectral Satellite Images. https://doi.org/10.1109/CVPRW.2017.193

Rußwurm, M., & Körner, M. (2018). Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders. ISPRS International Journal of Geo-Information, 7. https://doi.org/10.3390/ijgi7040129

Samardžić-Petrović, M., Dragićević, S., Bajat, B., & Kovačević, M. (2015). Exploring the Decision Tree Method for Modelling Urban Land Use Change. GEOMATICA, 69(3), 313–325. https://doi.org/10.5623/cig2015-305

Samardzic-Petrovic, M., Dragicevic, S., Kovacevic, M., & Bajat, B. (2016). Modeling Urban Land Use Changes Using Support Vector Machines. Transactions in Gis, 20(5), 718–734. https://doi.org/10.1111/tgis.12174

Samardzic-Petrovic, M., Kovačević, M., Bajat, B., & Dragićević, S. (2017). Machine Learning Techniques for Modelling Short Term Land-Use Change. ISPRS International Journal of Geo-Information, 6(12), 387. https://doi.org/10.3390/ijgi6120387

Sanders, L., Pumain, D., Mathian, H., Guérin-Pace, F., & Bura, S. (1997). SIMPOP: A Multiagent System for the Study of Urbanism. Environment and Planning B: Planning and Design, 24(2), 287–305. https://doi.org/10.1068/b240287

Schmidhuber, J. (2017, March). History of computer vision contests won by deep CNNs on GPUs. Retrieved May 24, 2019, from http://people.idsia.ch/~juergen/computer-vision-contests-won-by-gpu-cnns.html

Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing, 45(11), 2673–2681. https://doi.org/10.1109/78.650093

Sharma, A., Liu, X., & Yang, X. (2018). Land cover classification from multi-temporal, multi-spectral remotely sensed imagery using patch-based recurrent neural networks. Neural Networks, 105, 346–355. https://doi.org/10.1016/j.neunet.2018.05.019

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W., & WOO, W. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, & R. Garnett (Eds.), Advances in Neural Information Processing Systems 28 (pp. 802–810). Retrieved from http://papers.nips.cc/paper/5955-convolutional-lstm-network-a-machine-learning-approach-for-precipitation-nowcasting.pdf

Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. Remote Sensing of Environment, 62(1), 77–89. https://doi.org/10.1016/S0034-4257(97)00083-7

Stevens, D., & Dragićević, S. (2007). A GIS-Based Irregular Cellular Automata Model of Land-Use Change. Environment and Planning B: Planning and Design, 34(4), 708–724. https://doi.org/10.1068/b32098

White, R., & Engelen, G. (1997). Cellular Automata as the Basis of Integrated Dynamic Regional Modelling. Environment and Planning B: Planning and Design, 24(2), 235–246. https://doi.org/10.1068/b240235

White, R., Engelen, G., & Uljee, I. (1997). The Use of Constrained Cellular Automata for High-Resolution Modelling of Urban Land-Use Dynamics. Environment and Planning B: Planning and Design, 24(3), 323–343. https://doi.org/10.1068/b240323

Wu, F., & Webster, C. J. (1998). Simulation of Land Development through the Integration of Cellular Automata and Multicriteria Evaluation. Environment and Planning B: Planning and Design, 25(1), 103–126. https://doi.org/10.1068/b250103

Yang, Y., Zhou, J., Ai, J., Bin, Y., Hanjalic, A., Shen, H. T., & Ji, Y. (2018). Video Captioning by Adversarial LSTM. IEEE Transactions on Image Processing, 27(11), 5600–5611. https://doi.org/10.1109/TIP.2018.2855422

Yao, Y., Li, X., Liu, X., Liu, P., Liang, Z., Zhang, J., & Mai, K. (2017). Sensing Spatial Distribution of Urban Land Use by Integrating Points-of-interest and Google Word2Vec Model. Int. J. Geogr. Inf. Sci., 31(4), 825–848. https://doi.org/10.1080/13658816.2016.1244608

Yao, Y., Liang, H., Li, X., Zhang, J., & He, J. (2017). Sensing Urban Land-Use Patterns By Integrating Google Tensorflow And Scene-Classification Models. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2-W7. https://doi.org/10.5194/isprs-archives-XLII-2-W7-981-2017

Yuan, Z., Zhou, X., & Yang, T. (2018). Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 984–992. https://doi.org/10.1145/3219819.3219922

Zen, H., & Sak, H. (2015). Unidirectional long short-term memory recurrent neural network with recurrent output layer for low-latency speech synthesis. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4470–4474. https://doi.org/10.1109/ICASSP.2015.7178816

The MathWorks Inc. MATLAB R2018a. Natick, MA. 2018.

Python Software Foundation. Python Language Reference, version 3.6. Available at http://www.python.org. 2016.

Fran, C. Keras. Available at https://keras.io/. 2015.

ESRI. ArcGIS Desktop version 10.6. Redlands, CA. 2018.

Chapter 4.

Conclusion

4.1. Overall conclusions

LUC is a complex and dynamic phenomenon which is in tight relation with the environment, climate, and society. Typically, the methodological approaches and tools used for studying LU are from the field of RS and GIS given the fact that recently LU data is vastly available. The focus of this thesis is to explore the potential of the DL models on LUC classification and short-term forecasting. The Community of Cloverdale and the City of Surrey have been chosen as the study areas for the implementation of various CNN and RNN models. The obtained results have demonstrated high accuracy from CNN-based models for LUC classification and RNN-based models of LUC short-term forecasting. The research in this thesis contribute to the incorporation of DL models into LUC studies for its better understanding.

The objectives of this thesis research were achieved through the development of two DL based models trained by historical GIS and RS data. Chapter 2 developed CNN-based LU classification models to identify LUC from historical orthophotos. The pre-trained CNN frameworks which have been developed by experts, can be used as off-the-shelf tools with simple retraining. Eight transferred CNN models were tested, and all have achieved more than 95% accuracy. The study area has some unique LU features that were not. Therefore, some images were manually sampled and labeled from the study area to augment the UC-Merced dataset. The purpose of combining the UC-Merced dataset and manually sampled images from the study area was to increase the suitability of the CNN models in the chosen study area.

Chapter 3 was based on RNN models for short-term LUC forecasting. Despite being spatially diverse, the LU also change over time under the influence of various drivers. Knowing the possible changes of LU in near future is beneficial for better land management practices. RNNs have been used mostly for time-sequence modeling. They are proven to be effective for processing sequential inputs that have inner correlations. By identifying the rules of how LUC happens in time-space, LUC can be forecasted based on the rules and ancillary data. Therefore, the second objective of this

thesis was to explore the possibility of RNN models for short-term forecasting of LUC basing on historical LU datasets. RNNs are DL models that are expected to automatically study and apply the “rules” of the LUC.

The applications for CNNs were conducted using MATLAB software (The MathWorks Inc., 2018), and applications for RNNs were developed using Python software (Python Software Foundation, 2016) and Keras API (Fran, 2015). DL-based models have advanced the methods of classification and forecasting basing on RS and GIS datasets. They have also automated the process of LUC analysis. The developed and trained DL models can effectively learn the features from GIS and RS datasets. In addition, they can be transferred to some of the other study areas without complicated re-developing processes.

4.2. Limitations and future work

Due to the transferability of CNN and RNN models, the trained classification and forecasting DL based models in this thesis can be applied to other study areas and for different years in Canada where considerable urbanization is taking place. However, the proposed DL-based models still need additional consideration of composition, quality, and quantity of the training datasets to ensure the optimal performance. This thesis research can be improved in the several ways.

In Chapter 2, segmenting each LU scene apart has been a great challenge. From the observations of satellite images of the urban areas, it is noted that individual LU scene has irregular size and boundary. The CNN models were trained with an identical size of LU images. The orthophoto was divided uniformly into patches to be classified. However, this was not precise enough to distinguish each category. Features from different LU categories can overlap in a patch to be classified, while the CNN LU classifier can only assign one LU label to one patch. Similar to the study by Zhang et al (2019) future work should investigate the classification of LU on land parcels instead of raster cells as this would be closer to the real-world subdivisions of land. The parcel is a better representation of property boundaries with their range in size from a small lot to a large area. However, the random size of the parcel can be challenging. Therefore, a proper LU segmentation method needs to be developed before applying CNNs. Also, to identify the land features with different sizes, using multiscale LU images to train the

CNN models will be a potential solution for improvement in future studies. Finally, the kappa metrics of the obtained LU change classification results were not performed due to the lack of reference data. Kappa statistics (Congalton, 1991) or other metrics should be conducted in the future study once the sufficient data are collected.

In Chapter 3, only available LU data were used to train the RNNs due to the limited availability of data. Given that there are many factors that drive the LUC process, future work should consider more geographic, social or economic factors into training process. The improved design of training datasets would increase the reliability of the forecasted LUC. Moreover, finding appropriate datasets covering wider number of years and that contain information about different factors that drive the LUC would insure building a more comprehensive RNN models. Another limitation of is related to the sampling method used to evaluate the predicted results. Due to the fact that LU data were not available for 2016, the comparison of forecasted map output with reference data was performed in a specific manner. The 604 samples points were randomly used from the study area, which may not be an optimal method of taking samples over heterogeneous areas. A wide variety of sampling methods has been reported in RS literature such as simple random sampling (SRS), cluster sampling (CS), stratified random sampling (STRAT), mean of surface with non-homogeneity (MSN) method (Hashemian, Abkar, & Fatemi, 2004; Mu et al., 2015). They could potentially be applied in this study and a systematic approach on how different sampling design would affect the overall accuracy of obtained forecasted results should be conducted. According to Samardzic-Petrovic, Dragicevic, Kovacevic, & Bajat (2016), the Kappa value for random sampling is not a realistic estimate since the majority of raster cells stay unchanged. For example sample points could be chosen only from changed raster cells (Samardzic-Petrovic, Kovačević, Bajat, & Dragićević, 2017) to calculate the Kappa value separated from overall evaluation. Finally, if the appropriate data were available for year 2016 the evaluation of the accuracy of the obtained short-term forecasted land use could be performed with variety of exiting methods for map comparisons (Ahmed, Ahmed, & Zhu, 2013).

The issues surrounding data leakage and model overfitting problems, typical for ML and DL studies, were not considered in this thesis research. Data leakage (Gutierrez, 2014) is a problem that happens during the development of predictive models and when using cross-validation (Brownlee, 2016). It occurs when the training data

contain the information that are going to be predicted or is not from the real world. To avoid leakage, two stages are suggested by Kaufman, Rosset, & Perlich (2011): firstly, tagging every observation with legitimacy tags during collection; secondly, observing learn-predict separation. Data leakage may further result in model overfitting. Model overfitting (Hawkins, 2004) is a modeling error when a model learns the function that is too closely fit to a limited set of data points, so that cannot fit well with other data. Dropout is one of the simple way to prevent overfitting (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014), which means ignoring some random layers or neurons in ML models. Cross-validation (Schaffer, 1993), regularization (Leibbrandt, 1975), early stopping and feature reduction are some of the other possible solutions to alleviate the model overfitting problem.

4.3. Thesis contributions

The major contribution of this thesis is in bridging the gap between existing methods of classification and forecasting land use change and DL models that have the ability to extract features automatically from high-dimensional datasets. Another contribution is the convenience of quantifying and analyzing the LUC based on the outputs of DL-based models since the study areas were pre-processed and presented in regular patches or raster cells.

The DL-based models were capable of extracting features and transition rules automatically from historical data and then performing classification and short-term forecasting. DL models have greatly reduced the time and effort spent on performing the task without sacrificing efficiency and accuracy. This thesis also contributes to the fields of RS and GIS by providing advanced approaches based on DL models to understand, analyze and forecast LUC, and potentially benefit LU management practices and sustainable development.

4.4. Reference

Ahmed, B., Ahmed, R., & Zhu, X. (2013). Evaluation of Model Validation Techniques in Land Cover Dynamics. ISPRS International Journal of Geo-Information, 2(3), 577–597. https://doi.org/10.3390/ijgi2030577

Brownlee, J. (2016, August 2). Data Leakage in Machine Learning. Retrieved July 8, 2019, from Machine Learning Mastery website: https://machinelearningmastery.com/data-leakage-machine-learning/

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37(1), 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Gutierrez, D. (2014, November 26). Ask a Data Scientist: Data Leakage. Retrieved August 22, 2019, from InsideBIGDATA website: https://insidebigdata.com/2014/11/26/ask-data-scientist-data-leakage/

Hashemian, M., Abkar, A., & Fatemi, S. (2004). Study of sampling methods for accuracy assessment of classified remotely sensed data.

Hawkins, D. M. (2004). The Problem of Overfitting. Journal of Chemical Information and Computer Sciences, 44(1), 1–12. https://doi.org/10.1021/ci0342472

Kaufman, S., Rosset, S., & Perlich, C. (2011). Leakage in Data Mining: Formulation, Detection, and Avoidance. 8.

Leibbrandt, G. (1975). Introduction to the technique of dimensional regularization. Reviews of Modern Physics, 47(4), 849–876. https://doi.org/10.1103/RevModPhys.47.849

Mu, X., Hu, M., Song, W., Ruan, G., Ge, Y., Wang, J., … Yan, G. (2015). Evaluation of Sampling Methods for Validation of Remotely Sensed Fractional Vegetation Cover. Remote Sensing, 7(12), 16164–16182. https://doi.org/10.3390/rs71215817

Samardzic-Petrovic, M., Dragicevic, S., Kovacevic, M., & Bajat, B. (2016). Modeling Urban Land Use Changes Using Support Vector Machines. Transactions in Gis, 20(5), 718–734. https://doi.org/10.1111/tgis.12174

Samardzic-Petrovic, M., Kovačević, M., Bajat, B., & Dragićević, S. (2017). Machine Learning Techniques for Modelling Short Term Land-Use Change. ISPRS International Journal of Geo-Information, 6(12), 387. https://doi.org/10.3390/ijgi6120387

Schaffer, C. (1993). Selecting a classification method by cross-validation. Machine Learning, 13(1), 135–143. https://doi.org/10.1007/BF00993106

Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15, 1929–1958.

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., & Atkinson, P. M. (2019). Joint Deep Learning for land cover and land use classification. Remote Sensing of Environment, 221, 173–187. https://doi.org/10.1016/j.rse.2018.11.014

Fran, C. (2015). Keras. Available at https://keras.io/

Python Software Foundation. (2016). Python Language Reference, version 3.6. Available at http://www.python.org

The MathWorks Inc. (2018). MATLAB R2018a. Natick, MA