

CLASSIFICATION OF SETTLEMENT TYPES FROM TWEETS USING LDA AND LSTM

Rong Huang^{1,2}, Hannes Taubenböck³, Lichao Mou^{1,2}, Xiao Xiang Zhu^{1,2}

¹Signal Processing in Earth Observation, Technical University of Munich (TUM), Munich, Germany

²Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Wessling, Germany

³German Remote Sensing Data Center (DFD), DLR, Wessling, Germany

ABSTRACT

Land use reflects the interrelation between the physically built environment and the activity patterns of people. It is indispensable information for decision-makers, but up-to-date and accurate land use information is often absent. Unlike approaches that make use of remote sensing data, in this work, we are interested in a novel data source, tweets, and explore its potential for land use classification in urban areas. Specifically, we propose a general framework for classifying settlement land-use types by extracting location, time, quantity and text features of twitter data. To do so, we apply latent Dirichlet allocation (LDA) and long short-term memory (LSTM) and then combines those features with spatial-temporal feature using Fused SVM and a two-stream convolutional neural network (CNN) for classification. For the case of classifying individual tweets by the land-use classes relevant in this study - residential, non-residential and mixed usage -, we reach overall accuracy (OA), average accuracy (AA), and Kappa coefficient with 72.35%, 73.76%, and 58.43%, respectively. As for the case of classifying block settlement types, we reach 61.90%, 63.33%, and 42.84%, respectively.

Index Terms— Land use classification, latent Dirichlet allocation (LDA), long short-term memory (LSTM), two-stream convolutional neural network (CNN)

1. INTRODUCTION

It is predicted that 66% of the world population will inhabit urban areas by the year of 2050 [1], and living in cities will be the major lifestyle. However, the overloaded population and increased human activities pose great challenges for urban governors and planners. To avoid uncontrolled development and improve quality of living environments or maintain at current levels, adequate information is required on different correlated aspects for urban planning. Land use, as a crucial part of those aforementioned aspects, has become indispensable for urban decision-making [2]. Nevertheless, land use is always dynamically changing due to human activities. Therefore, it is of great importance to characterize up-to-date usage of land areas accurately and reliably.

Many studies have shown that remote sensing data is capable of offering an instant and accurate monitoring of urban environment [3]. Nevertheless, external, natural and, physical properties of ground components can be represented by features extracted from remote sensing images, whereas land use types often have a strong correlation with indoor human activities, which can only be inferred indirectly from remote sensing images [4, 5]. Even with other proxies such as population density [6], building volumes [7], or employment data [8], specific land use characteristics of these proxies are still ignored. Meanwhile, it is notable that the land utilization is fleetly changing with the change of human life and work styles. Although remote sensing data can be used to produce land use classification results with a smaller time interval (typically of years), compared to methods like questionnaires, it is still far from offering up-to-date information. To this end, an easy-accessed and up-to-date data source, geo-tagged tweets from social networks, are exploited to provide a better solution.

Previous studies on land use classification using tweets mainly focus on analysis of spatial and temporal patterns of tweets generated from a geographic area [9, 10, 11]. Although these aforementioned studies provided promising results, there are remaining challenges to define an object-based classification scheme and to explore the potential of utilizing text information in land use classification. In this work, the potential of using tweets for land use classification is investigated. Since urban areas are the main region of our interest, we consider three categories: residential (R), non-residential (NR), and mixed-use (MU).

Therefore, in this paper, we aim to explore the feasibility of applying text contents of tweets in land use classification and to propose a general framework of jointly combining spatial-temporal information (i.e., geo-tagged and time-stamped) and text content information to achieve land-use classification at block level.

2. METHODOLOGY

2.1. Framework

The framework of this work is illustrated in Figure 1. The purpose of our study is to classify urban land use types by fusing multiple features from tweets and open street map (OSM). We apply a two-step procedure to identify land use types in each urban block footprint which is a polygon provided by OSM identifying a regional area with similar functionalities. Firstly, in tweet instance, two natural language processing methods, namely latent Dirichlet allocation (LDA) [12] and long short-term memory (LSTM) [13], are applied on tweet texts to investigate the potential of using text content. Then, a temporal feature, namely hourly changes in the number of tweets during each day per week, is extracted from our tweet dataset for each block. Afterwards, we combine text features and temporal features by applying a fused support vector machine (SVM) and a two-stream convolutional neural network (CNN) to achieve our final results - a land-use classification consisting of three classes.

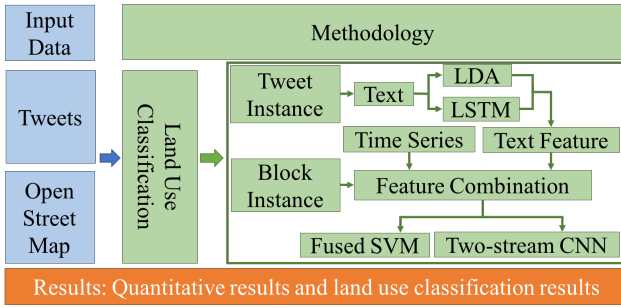


Fig. 1. Framework

2.2. Tweet Instance

In this section, we briefly introduce two natural language processing methods i.e., LDA and LSTM, which are applied to our tweet dataset.

2.2.1. LDA

LDA is a probabilistic latent space model for discrete data collections and proposed by Blei et al. The general idea of this model is that documents can be represented by random mixtures over latent topics, and each topic can be represented by a distribution of words. Based on the probability of each tweet on the latent topics, a classifier can be utilized to associate it with land use classification.

2.2.2. LSTM

LSTM, a special kind of recurrent neural networks (RNN), is capable of learning long-term dependencies. In LSTM struc-

ture shown in Figure 2, cell state is induced as C_t at step t to store internal memory. Meanwhile, with structures called gates, the amount of intermediate information to be stored by cell state C_t and to be passed to next hidden state h_t can be regulated. These gates are composed of a sigmoid neural net layer $\sigma(\cdot)$ and a point-wise multiplication operation.

Forget gate:

$$\mathbf{f}_t = \sigma_f(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f \mathbf{h}_{t-1} + \mathbf{b}_f) \quad (1)$$

Input gate:

$$\mathbf{i}_t = \sigma_i(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i \mathbf{h}_{t-1} + \mathbf{b}_i) \quad (2)$$

The cell state is updated:

$$\tilde{\mathbf{C}}_t = \tanh(\mathbf{W}_C \mathbf{x}_t + \mathbf{U}_C \mathbf{h}_{t-1} + \mathbf{b}_C) \quad (3)$$

$$\mathbf{C}_t = \mathbf{f}_t \odot \mathbf{C}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{C}}_t \quad (4)$$

Output gate:

$$\mathbf{o}_t = \sigma_o(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o \mathbf{h}_{t-1} + \mathbf{b}_o) \quad (5)$$

Finally the real output of this recurrent unit is:

$$\mathbf{h}_t = \mathbf{o}_t \tanh(\mathbf{C}_t) \quad (6)$$

where x_t is input data at step t , $W_{(\cdot)}$ and $U_{(\cdot)}$ are affine transformations and $b_{(\cdot)}$ is bias.

By doing so, LSTM is capable of deciding the amount of information to be passed to the next unit thus processing sequential data with more flexibility.

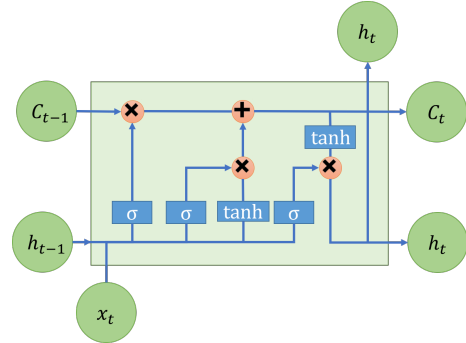


Fig. 2. Graphical model of a recurrent unit of LSTM.

2.3. Block Instance

The merits of tweet data is that it contains not only institutional text information but also important spatial and temporal patterns that can be associated with human activities. Text features can provide information about what people do in direct way. Temporal features provide when and where people post tweets, serving as an additional information for determining place types. In this work, we combine both features. Based on these input features we propose a method for land use classification, and methods in Section 2.3.1 are applied for comparison.

2.3.1. Majority Voting and Taking Average

Since tweets for each block have been classified, classes of blocks can be directly determined based on labels of tweets using the following two methods. One is using the label of the majority of classified tweets located in the block, and the other is to calculate the probability of each category by averaging probabilities of each category of classified tweets located in this block.

2.3.2. Fused SVM

To effectively improve the discriminative ability of extracted features, we propose to use a multi-feature combination strategy by fully considering temporal information. This allows to obtain a better classification performance by using more advanced classifiers. More specifically, the multi-features consist of text feature extracted by natural language processing technique and block-based temporal features, which can be simply stacked to be a new feature representation as a input of a kernel SVM classifier.

2.3.3. Two-Stream CNN

To further improve the classification performance, a two-stream CNN [14] is utilized, as illustrated in Figure 3. After crossing three 1-D convolutional blocks, a deep feature representation can be obtained. Subsequently a direct connection between the deep feature and label information via a softmax layer. The proposed network aims at automatically learning relationship between these two features, thereby achieve a better classification result.

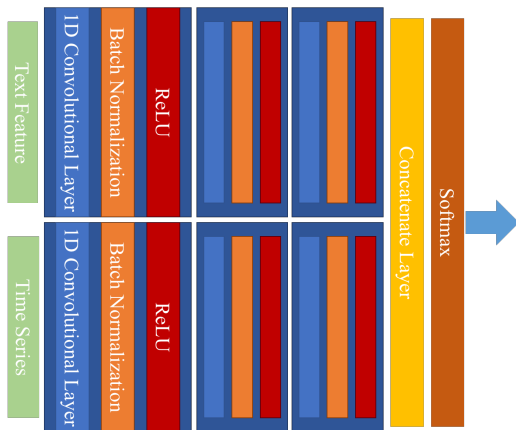


Fig. 3. Structure of Two-Stream CNN.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1. Dataset and Study Area

We obtain a 6-month sample of 122,577 geo-located tweets (11/09/2015-09/03/2016) from Twitter via a standardized application programming interface for the city of Munich, Germany.

3.2. Comparison between LDA and LSTM in tweet instance

The results of LSTM and LDA are listed in Table 1. Compared to LDA, LSTM increases the accuracy significantly by 9% for the OA, 9% for the AA, and 14% for the Kappa coefficient, respectively, in the case of English language; by 9% for the OA, 7% for the AA, and 13% for the Kappa coefficient, respectively, in German language. Besides, the land-use classes residential and non-residential are less prone to be misclassified when applying LSTM for text processing. Naturally, LSTM is a better choice for block-based instance classification.

Table 1. Classification accuracies comparison between LDA and LSTM in percentages in tweet instance (the best accuracy in each row is shown in bold).

(a) English			(b) German		
Category	LDA	LSTM	Category	LDA	LSTM
R	48.59	59.76	R	53.26	63.43
NR	54.79	66.84	NR	53.10	75.81
MU	79.14	78.33	MU	93.33	82.04
OA	60.84	67.75	OA	62.10	72.35
AA	58.84	68.31	AA	66.57	73.76
Kappa	38.51	51.61	Kappa	43.09	58.43

3.3. Comparison of Different Techniques in Block Classification

As displayed in Table 2, results of different classification techniques in block instance are compared in different aspects. The increasing accuracy of classification of Fused SVM compared with methods which only use text feature supports the promise of extending the feature space by temporal features. However, NR is easier to be misclassified as the other two classes with the addition of temporal features. As for the designed two-stream CNN, on our tweet dataset in Munich, our network is able to achieve an increment of 4.76%, 4.64% and 7.02% for OA, AA and the Kappa coefficient, respectively, compared with Fused SVM. Meanwhile, for each categories, they all reach better accuracies. In general, the two-stream CNN performs better than the fused CNN regarding to feature combination.

Table 2. Classification accuracies of different techniques in percentages in block instance (the best accuracy in each row is shown in bold).

Category	Majority Voting	Taking Average	Fused SVM	Two-Stream CNN
R	54.84	54.84	60.00	62.07
NR	57.58	55.88	51.35	63.33
MU	60.00	57.89	64.71	60.00
OA	57.14	55.95	57.14	61.90
AA	57.47	56.21	58.69	63.33
Kappa	35.65	33.89	35.82	42.84

4. CONCLUSION AND OUTLOOK

In this work, a framework for block instance land-use classification is proposed, which tends to explore more potential from tweets and produce a more informative classification result. With this method, relatively reasonable classification accuracies can be achieved for land use classification of individual blocks. We also investigated two natural language processing methods (i.e., LDA and LSTM), and the latter LSTM performs better due to the capability of dealing with long-term sequential information. As for the feature combination, a two-stream CNN gives us a better classification accuracy than a fused SVM. In general, tweets are a valuable data source in supporting land-use classification.

In combination with remote sensing classifications the combination of both data sources is promising to improve timeliness and accuracies in land-use classifications.

5. ACKNOWLEDGEMENT

This work is jointly supported by the China Scholarship Council, the Helmholtz Association under the framework of the Young Investigators Group SiPEO (VH-NG-1018, www.sipeco.bgu.tum.de), and the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No. ERC-2016-StG-714087, Acronym: So2Sat).

6. REFERENCES

- [1] United Nations, "World urbanization prospects: The 2014 revision, highlights. department of economic and social affairs," *Population Division, United Nations*, 2014.
- [2] James Richard Anderson, *A land use and land cover classification system for use with remote sensor data*, vol. 964, US Government Printing Office, 1976.
- [3] Hannes Taubenböck, Thomas Esch, Andreas Felbier, Michael Wiesner, Achim Roth, and Stefan Dech, "Monitoring urbanization in mega cities from space," *Remote sensing of Environment*, vol. 117, pp. 162–176, 2012.
- [4] Justin D Paola and Robert A Schowengerdt, "A detailed comparison of backpropagation neural network and maximum-likelihood classifiers for urban land use classification," *IEEE Transactions on Geoscience and remote sensing*, vol. 33, no. 4, pp. 981–996, 1995.
- [5] Fabio Pacifici, Marco Chini, and William J Emery, "A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification," *Remote Sensing of Environment*, vol. 113, no. 6, pp. 1276–1292, 2009.
- [6] Dengsheng Lu and Qihao Weng, "Use of impervious surface in urban land-use classification," *Remote Sensing of Environment*, vol. 102, no. 1, pp. 146–160, 2006.
- [7] Hannes Taubenböck, Martin Klotz, Michael Wurm, J Schmieder, B Wagner, M Wooster, Thomas Esch, and Stefan Dech, "Delineation of central business districts in mega city regions using remotely sensed data," *Remote sensing of Environment*, vol. 136, pp. 386–401, 2013.
- [8] Hannes Taubenböck, Ines Standfuß, Michael Wurm, Angelika Krehl, and Stefan Siedentop, "Measuring morphological polycentricity-a comparative analysis of urban mass concentrations using remote sensing data," *Computers, Environment and Urban Systems*, vol. 64, pp. 42–56, 2017.
- [9] Ryong Lee, Shoko Wakamiya, and Kazutoshi Sumiya, "Urban area characterization based on crowd behavioral lifelogs over twitter," *Personal and ubiquitous computing*, vol. 17, no. 4, pp. 605–620, 2013.
- [10] Vanessa Frias-Martinez and Enrique Frias-Martinez, "Spectral clustering for sensing urban land use using twitter activity," *Engineering Applications of Artificial Intelligence*, vol. 35, pp. 237–245, 2014.
- [11] Martin Klotz, Michael Wurm, Xiaoxiang Zhu, and Hannes Taubenböck, "Digital deserts on the ground and from space," in *Urban Remote Sensing Event (JURSE), 2017 Joint. IEEE*, 2017, pp. 1–4.
- [12] David M Blei, Andrew Y Ng, and Michael I Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.
- [13] Sepp Hochreiter and Jürgen Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [14] Jingliang Hu, Lichao Mou, Andreas Schmitt, and Xiao Xiang Zhu, "Fusionet: A two-stream convolutional neural network for urban scene classification using polsar and hyperspectral data," in *Urban Remote Sensing Event (JURSE), 2017 Joint. IEEE*, 2017, pp. 1–4.