

Geotagging Tweets to Landmarks using Convolutional Neural Networks with Text and Posting Time*

Kwan Hui Lim

Information Systems Technology and Design Pillar
Singapore University of Technology and Design
Singapore
kwanhui_lim@sutd.edu.sg

Shanika Karunasekera, Aaron Harwood,
Yasmeen George

School of Computing and Information Systems
The University of Melbourne
Australia
{karus,aharwood,georgey}@unimelb.edu.au

ABSTRACT

Geotagged tweets (and similar social media) serve many important applications, ranging from crisis/diasaster management to targetted advertising. However, only a small proportion of tweets are geotagged with their posting locations, thus limiting the effectiveness of these applications. In this paper, we propose a Convolutional Neural Network (CNN) architecture for geotagging tweets to landmarks, based on the text in tweets and other meta information, such as posting time and source. Using a dataset of Melbourne tweets, we evaluated our algorithm against various state-of-the-art baselines. Experimental results show that our algorithm outperformed these baselines by between 7.6% to 75.6% in terms of accuracy and 2.8% to 43.3% in terms of mean distance error.

CCS CONCEPTS

• **Human-centered computing** → **Social networking sites**;
• **Computer systems organization** → *Neural networks*; •
Information systems → *Location based services*.

KEYWORDS

Geotagging, Geolocation, Neural Networks, Twitter

*This paper is an extended version of a poster paper published at IUI'19 [18].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
IUI '19 Companion, March 17–20, 2019, Marina del Ray, CA, USA
© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6673-1/19/03...\$15.00
<https://doi.org/10.1145/3308557.3308691>

ACM Reference Format:

Kwan Hui Lim and Shanika Karunasekera, Aaron Harwood, Yasmeen George. 2019. Geotagging Tweets to Landmarks using Convolutional Neural Networks with Text and Posting Time. In *24th International Conference on Intelligent User Interfaces (IUI '19 Companion)*, March 17–20, 2019, Marina del Ray, CA, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3308557.3308691>

1 INTRODUCTION

Twitter is a popular and prevalent microblogging service, with as many as 500 million tweets being generated daily [12]. Most of these tweets are posted via mobile devices [31], where users have the option to tag their tweets with the location they are posting from. In turn, knowing the location of a tweet (and similar social media) helps to facilitate many important applications, ranging from crisis management [4, 11, 20, 27] to friendship and community detection [3, 8, 17, 26]. However, research has found that only 0.85% of tweets are geo-tagged with their posting location [28], i.e., explicitly labelled with their latitude and longitude coordinates. This low proportion of geo-tagged tweets severely limits the effectiveness of the above-mentioned applications and we aim to address this problem by developing algorithms for geotagging tweets.

Research Problem. We aim to address the problem of geotagging tweets to known landmarks, or more specifically, determine the landmark that an individual tweet is posted from without making use of GPS data.¹ This problem differs from many existing problems that attempt to geo-tag individual tweets to large regions like cities and countries (instead of fine-grained POIs), or those that geo-tag users to their home location (instead of individual tweets) [6, 10, 19].

Contributions. Our main contribution include proposing an algorithm for geotagging tweets to landmarks. This algorithm is based on a Convolutional Neural Network (CNN)

¹While we utilize tweets in this work, our proposed algorithm is generalizable to any social media message with similar characteristics.

architecture that uses tweet text represented as word embeddings, and meta information such as posting time and source. Our intuition behind posting time and source is that given certain landmarks, people are more likely to post tweets at specific timings (e.g., in the morning for cafes, in the evening for pubs) and using certain sources (e.g., via mobile phone app at train stations, via desktop website at university). Preliminary results show that our algorithm out-performs various state-of-the-art baselines by between 7.6% to 75.6% in terms of accuracy and 2.8% to 43.3% in terms of mean distance error.

Structure and Organization. The rest of the paper is organized as follows. Section 2 reviews the key related work in this area. Section 3 describes our proposed algorithm and illustrates the architecture of our model. Section 4 outlines our evaluation methodology, while Section 5 discusses our experimental results. Finally, Section 6 summarizes and concludes the paper.

2 RELATED WORK

Twitter geolocation has been well-studied in the recent years, with many works focussing on geolocating tweets to cities and countries, or geolocating users to their home location based on their history of tweets [6, 10, 19]. For this purpose, researchers have used features such as Location Indicative Words [9, 10], #hashtags and @mentions [6], dialect terms [24] and friendship networks via @mentions [25]. Others have studied the problem of geolocating tweets to landmarks, using variations of Multinomial Naive Bayes [7, 22]. For a more comprehensive discussion on the other aspects of location prediction on Twitter, we refer readers to the survey paper in [33].

Text classification is another research problem that is closely related to our problem of geolocating tweets, which comprises mainly text. For example, Kim [14] proposed one of the first application of CNN for text classification, alongside pre-trained and dynamic word embeddings. Extending the work of [14], Zhang and Wallace [32] empirically studied the effects of different parameters using their CNN model. Others like [13] utilized a CNN architecture with one-hot encoding of words, while [24] used a bag-of-words representation as input to a Multi Layer Perceptron.

3 NETWORK ARCHITECTURE

Various works have effectively utilized CNN for text classification tasks relating to sentiment analysis [14, 32]. In the same spirit as these earlier works, we propose a network architecture that utilizes textual features in a tweet, in addition to other meta-data such as tweeting time, posting source and media usage, for the task of geo-locating tweets to fine-grained locations. Figure 1 shows the overall architecture of our proposed CNN, which we describe in detail next.

Inputs to Model. Our model takes in a tweet comprising n words, where each i th word $w_i \in \mathbb{R}^d$ in the tweet is represented as an embedding vector of dimension d .² A tweet t is thus represented as an $n \times d$ matrix, denoted $t \in \mathbb{R}^{n \times d}$. In cases where a tweet t comprises less than n words, the tweet will be zero-padded to length n ($n = 50$ in our work). In addition, our model takes in input of tweeting time, posting source and media usage, which are in turn converted to one-hot encodings.

Convolutional and Pooling Operations. Each convolution operation involves applying a filter w of size $h \times d$ across the tweet $t \in \mathbb{R}^{n \times d}$, h words at a time. Let $t_{i:i+h-1}$ be a submatrix of tweet t from row i to row $i + h - 1$, a single application of the convolution operation with filter w on $t_{i:i+h-1}$ results in:

$$o_i = f(w \cdot t_{i:i+h-1} + b) \tag{1}$$

where $b \in \mathbb{R}$ denotes the bias term and f denotes the activation function applied to $w \cdot t_{i:i+h-1} + b$. We adopted the Rectified Linear Unit (ReLU) [21] as the activation function, which has shown good performance for CNNs [14, 32]. This convolution operation is then repeated for the remaining of the tweet, i.e., sliding filter w across the tweet striding one word at a time, resulting in an output $o = [o_1, o_2, \dots, o_{n-h+1}]$. Thereafter, we apply 1-max pooling [1] on the output o to select the most important feature $\hat{o} = \max\{o\}$.

The earlier mentioned procedures (applying convolution filters and 1-max pooling) is then repeated f times for each filter size (we use filters of size 1, 2 and 3), resulting in a vector of size f .

Combining Text and Other Features. Alongside the text of the tweet, we also incorporate other meta data, such as the user-specified location, tweeting time³ and posting source. The location data is processed using the same convolutional and pooling operations described in the previous section, while the tweeting time (hour and day of week) and source are converted to one-hot encodings. Following which, we concatenate the features based on the output of the convolutional and pooling operations (tweet text and location) and one-hot encodings (tweeting hour, day and source). This concatenated feature vector is then passed to a softmax layer to produce the probabilities for each POI, along with dropout being applied as a form of regularization [29]. This model is then trained using stochastic gradient descent with ADAM updating [15].

²We use pre-trained GloVe word vectors with 200 dimensions, which were trained on a Twitter corpus of 2B tweets [23]

³We conducted a one-way ANOVA, which showed a significant effect for landmark on tweeting time, which were: (i) in terms of tweeting hour, $F(237,266693)=24.508, p<.0001$; (ii) in terms of tweeting day of week, $F(237,266693)=35.217, p<.0001$.

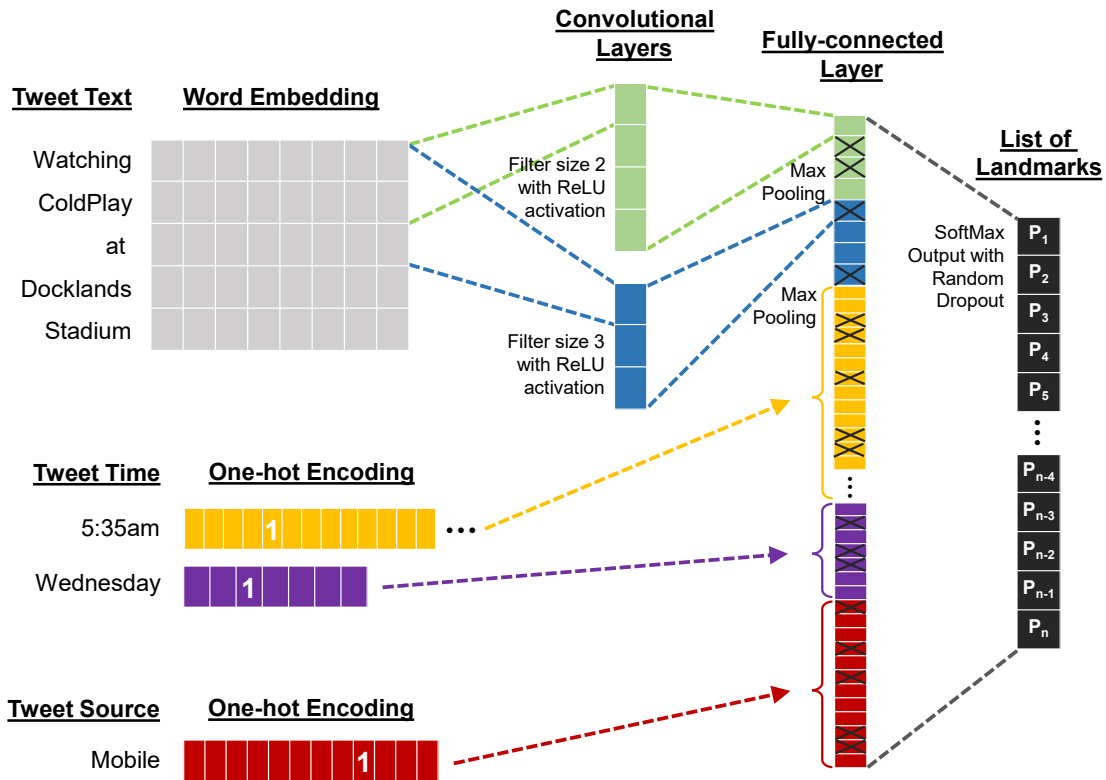


Figure 1: Architecture of Convolutional Neural Network for Geolocation (with a sample subset of tweet features).

4 EXPERIMENTAL SETUP

In this section, we describe our dataset and elaborate on our experimentation methodology in terms of the evaluation metrics and baseline algorithms.

Dataset

Our dataset comprises 266,931 geo-tagged tweets that were posted in Melbourne, which we collected using the Twitter API. Using an similar approach as [2, 5, 16, 30], we label a tweet as being associated with a landmark if the latitude and longitude coordinates of the tweet and landmark differs by less than 100m. There are a total of 242 landmarks in this dataset. Thereafter, 80% of the tweets were used for training and the remaining 20% for testing/validation.

Evaluation and Metrics

Like many works in similar areas [6, 9, 10, 33], we evaluate the performance of our algorithm and the baselines using the following metrics:

- **Accuracy@k (Accuracy@k)**. The proportion of correctly predicted locations, based on the top- k predicted locations.
- **Distance Error (MeanDist/MedDist)**. The mean and median distance (in metres) between the predicted location and the ground truth location.

The various Accuracy@ k metrics aims to measure the proportion of correct predictions at varying values of k , while MeanDist and MedDist measure how far a predicted location is from its actual location.

Algorithms and Baselines

We compare our algorithm against various baselines, including traditional approaches based on Multinomial Naive Bayes and more recent approaches based on neural networks. These baselines are:

- **Multinomial Naive Bayes with Uni/Bi/Tri-grams (MNB-Ngram)**. Multinomial Naive Bayes (MNB) is a popular classifier used for geotagging tasks [6, 7, 10, 19, 22]. In the same spirit, we use baselines with a

Table 1: Comparison of our algorithm against various baselines in terms of mean accuracy@k (Accuracy@k) and mean/median distance errors (MeanDist/MedDist). Best results are highlighted in bold/blue, second best in italics/red.

Algorithm	Accuracy@1	Accuracy@5	Accuracy@10	Accuracy@20	MeanDist	MedDist
MNB-1gram	0.3863	0.6796	0.7509	0.8221	620.1	254.7
MNB-2gram	0.4628	0.7249	0.7836	0.8427	536.1	156.8
MNB-3gram	0.4982	0.7360	0.7905	0.8462	500.7	<i>50.2</i>
CNN-Rand	0.5648	0.7394	0.8044	0.8634	428.7	0.00
CNN-PreW	0.6045	<i>0.7727</i>	<i>0.8345</i>	<i>0.8854</i>	408.5	0.00
CNN-1Hot	<i>0.6308</i>	0.7689	0.8092	0.8543	<i>362.0</i>	0.00
MLP-BoW	0.6039	0.7635	0.8107	0.8572	408.1	0.00
Our Algo.	0.6785	0.8069	0.8493	0.8919	351.9	0.00

MNB classifier on unigrams, bigrams and trigrams as features for our geotagging task.

- **Convolutional Neural Network with Random Word Embedding (CNN-Rand)**. The basic version of the algorithm proposed in [14], using a Convolutional Neural Network with words in tweets represented by word embeddings that are randomly initialized.
- **Convolutional Neural Network with Pre-trained Word Embedding (CNN-PreW)**. Similar to **CNN-Rand** [14], except that this algorithm uses pre-trained GloVe word embeddings [23] that are trainable.
- **Convolutional Neural Network with One-hot Encoding (CNN-1Hot)**. A similar Convolutional Neural Network architecture as that in [13], where the input tweet is represented as one-hot encodings instead of word embeddings.
- **Multi Layer Perceptron (MLP-BoW)**. Similar to [24], this baseline represents tweets as a bag-of-words that is used as input to a Multi Layer Perceptron with one hidden layer, and optimized using ADAM [15].

5 RESULTS AND DISCUSSION

Table 1 shows an overall summary of the performance of our proposed algorithm against the various baselines, in terms of accuracy@k (for k = 1, 5, 10, 20), and mean and median distance errors (in metres).

Our algorithm provides the best overall performance in terms of all metrics. The Accuracy@1 results show improvements of 7.6% to 75.6% when comparing our algorithm against the baselines, and similar improvements for the Accuracy@5, Accuracy@10 and Accuracy@20 results. Similarly, our algorithm results in an improvement of 2.8% to 43.3% against the baselines in terms of MeanDist. In terms of MedDist, all neural network based algorithms out-perform all Multinomial Naive Bayes variants with a median distance error of 0, due to them being able to geotag more than half of the tweets correctly.

6 CONCLUSION AND FUTURE WORK

In this paper, we studied the problem of geotagging tweets to landmarks without the use of GPS information. To solve this problem, we proposed a CNN-based algorithm that uses word embeddings to represent tweet text and one-hot encodings for tweet meta information such as posting time and source. Using a dataset of geo-tagged tweets in Melbourne, preliminary results show that our algorithm out-performed state-of-the-art baselines by between 7.6% to 75.6% in terms of accuracy and 2.8% to 43.3% in terms of mean distance error.

While we focused on the use of tweet text and other meta information, future work can consider other forms of embedded media, such as images and videos, alongside image recognition techniques to geotag those media to known landmarks or localities. Similarly, we can augment our approach with the friendship network of a user, as mutual friendships are shown to display homogeneity in terms of their interests and visiting places.

ACKNOWLEDGMENTS

This research is partly supported by the Singapore University of Technology and Design under grant SRG-ISTD-2018-140, and Defence Science and Technology, Australia. The authors thank the anonymous reviewers for their useful comments on this work.

REFERENCES

- [1] Y-Lan Boureau, Jean Ponce, and Yann LeCun. 2010. A theoretical analysis of feature pooling in visual recognition. In *Proceedings of the 27th international conference on machine learning (ICML'10)*. 111–118.
- [2] Igo Ramalho Brilhante, Jose Antonio Macedo, Franco Maria Nardini, Raffaele Perego, and Chiara Renso. 2015. On planning sightseeing tours with TripBuilder. *Information Processing & Management* 51, 2 (2015), 1–15.
- [3] Chloë Brown, Vincenzo Nicosia, Salvatore Scellato, Anastasios Noulas, and Cecilia Mascolo. 2012. The importance of being placefriends: discovering location-focused online communities. In *Proceedings of the 2012 ACM workshop on Workshop on online social networks (WOSN'12)*. 31–36.

- [4] Mark A. Cameron, Robert Power, Bella Robinson, and Jie Yin. 2012. Emergency situation awareness from twitter for crisis management. In *Proceedings of the 21st International Conference on World Wide Web (WWW'12)*. 695–698.
- [5] Dawei Chen, Cheng Soon Ong, and Lexing Xie. 2016. Learning Points and Routes to Recommend Trajectories. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management (CIKM'16)*. 2227–2232.
- [6] Lianhua Chi, Kwan Hui Lim, Nebula Alam, and Christopher J. Butler. 2016. Geolocation Prediction in Twitter Using Location Indicative Words and Textual Features. In *Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT'16)*. 227–234.
- [7] Wen-Haw Chong and Ee-Peng Lim. 2018. Exploiting User and Venue Characteristics for Fine-Grained Tweet Geolocation. *ACM Transactions on Information Systems* 36, 3 (2018), 26:1–26:34.
- [8] David J. Crandall, Lars Backstrom, Dan Cosley, Siddharth Suri, Daniel Huttenlocher, and Jon Kleinberg. 2010. Inferring social ties from geographic coincidences. *Proceedings of the National Academy of Sciences* 107, 52 (2010), 22436–22441.
- [9] Bo Han, Paul Cook, and Timothy Baldwin. 2012. Geolocation prediction in social media data by finding location indicative words. In *Proceedings of COLING 2012*. 1045–1062.
- [10] Bo Han, Paul Cook, and Timothy Baldwin. 2014. Text-based twitter user geolocation prediction. *Journal of Artificial Intelligence Research* 49 (2014), 451–500.
- [11] Muhammad Imran, Carlos Castillo, Fernando Diaz, and Sarah Vieweg. 2015. Processing social media messages in mass emergency: A survey. *Comput. Surveys* 47, 4 (2015), 67.
- [12] Internet Live Statistics. 2016. Twitter Usage Statistics. Internet. <http://www.internetlivestats.com/twitter-statistics/>.
- [13] Rie Johnson and Tong Zhang. 2015. Effective Use of Word Order for Text Categorization with Convolutional Neural Networks. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL'15)*. 103–112.
- [14] Yoon Kim. 2014. Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP'14)*. 1746–1751.
- [15] Diederik P. Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. In *arXiv preprint*. arXiv:1412.6980.
- [16] Kwan Hui Lim, Jeffrey Chan, Shanika Karunasekera, and Christopher Leckie. 2019. Tour Recommendation and Trip Planning using Location-based Social Media: A Survey. *Knowledge and Information Systems* (2019).
- [17] Kwan Hui Lim, Jeffrey Chan, Christopher Leckie, and Shanika Karunasekera. 2015. Detecting Location-centric Communities using Social-Spatial Links with Temporal Constraints. In *Proceedings of the 37th European Conference on Information Retrieval (ECIR'15)*. 489–494.
- [18] Kwan Hui Lim, Shanika Karunasekera, Aaron Harwood, and Yasmeen George. 2019. Geotagging Tweets to Landmarks using Convolutional Neural Networks with Text and Posting Time. In *Proceedings of the 24th International Conference on Intelligent User Interfaces Companion (IUI'19)*.
- [19] Jalal Mahmud, Jeffrey Nichols, and Clemens Drews. 2012. Where Is This Tweet From? Inferring Home Locations of Twitter Users. In *Proceedings of the Sixth International AAI Conference on Weblogs and Social Media (ICWSM'12)*. 511–514.
- [20] Stuart E. Middleton, Lee Middleton, and Stefano Modafferi. 2014. Real-time crisis mapping of natural disasters using social media. *IEEE Intelligent Systems* 29, 2 (2014), 9–17.
- [21] Vinod Nair and Geoffrey E. Hinton. 2010. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning (ICML'10)*. 807–814.
- [22] Ozer Ozdikis, Heri Ramampiaro, and Kjetil Norvag. 2018. Spatial Statistics of Term Co-occurrences for Location Prediction of Tweets. In *Proceedings of the 40th European Conference on Information Retrieval (ECIR'18)*. 494–506.
- [23] Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on Empirical Methods in Natural Language Processing (EMNLP'14)*. 1532–1543.
- [24] Afshin Rahimi, Trevor Cohn, and Timothy Baldwin. 2017. A Neural Model for User Geolocation and Lexical Dialectology. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL'17)*. 209–216.
- [25] Afshin Rahimi, Duy Vu, Trevor Cohn, and Timothy Baldwin. 2015. Exploiting Text and Network Context for Geolocation of Social Media Users. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL'15)*. 1362–1367.
- [26] Adam Sadilek, Henry Kautz, and Jeffrey P. Bigham. 2012. Finding your friends and following them to where you are. In *Proceedings of the fifth ACM international conference on Web search and data mining (WSDM'12)*. 723–732.
- [27] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. 2010. Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web (WWW'10)*. 851–860.
- [28] Luke Sloan, Jeffrey Morgan, William Housley, Matthew Williams, Adam Edwards, Pete Burnap, and Omer Rana. 2013. Knowing the tweeters: Deriving sociologically relevant demographics from Twitter. *Sociological Research Online* 18, 3 (2013), 7.
- [29] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 15, 1 (2014), 1929–1958.
- [30] Kendall Taylor, Kwan Hui Lim, and Jeffrey Chan. 2018. Travel Itinerary Recommendations with Must-see Points-of-Interest. In *Proceedings of the 2018 Web Conference Companion (WWW'18)*. 1198–1205.
- [31] World Street Journal. 2014. Data Point: Social Networking Is Moving on From the Desktop. Internet. <https://blogs.wsj.com/digits/2014/04/03/data-point-social-networking-is-moving-on-from-the-desktop/>.
- [32] Ye Zhang and Byron Wallace. 2017. A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (IJCNLP'17)*. 253–263.
- [33] Xin Zheng, Jialong Han, and Aixin Sun. 2018. A survey of location prediction on Twitter. *IEEE Transactions on Knowledge and Data Engineering* (2018).