Deep learning-based relation extraction from construction safety regulations for automated field compliance checking

Xiyu Wang[1] and Nora El-Gohary, A.M.ASCE[2]

[1]Graduate Student, Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana-Champaign, 205 North Mathews Ave., Urbana, IL 61801. E-mail: xiyuw2@illinois.edu
[2]Associate Professor, Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana-Champaign, 205 North Mathews Ave., Urbana, IL 61801. E-mail: gohary@illinois.edu

ABSTRACT

Information extraction provides an opportunity to automatically extract safety requirements from construction safety regulations to support automated safety compliance checking for detecting field non-compliances with these regulations. However, previous efforts on automating the safety compliance checking process fall short in their scalability and ability to automatically extract safety requirements, due to the complexity in unstructured text. Therefore, this paper proposes a deep learning-based information extraction method for extracting relations that link fall protection-related entities extracted from construction safety regulations for supporting automated field compliance checking. The proposed method uses an attention-based convolutional neural network model for recognizing and classifying relations. The proposed method was implemented and tested on two selected Occupational Safety and Health Administration (OSHA) sections related to fall protection. It has achieved a weighted precision, recall, and F-1 measure of 82.7%, 81.1%, and 81.3%, respectively, which indicates good relation extraction performance.

INTRODUCTION

Field compliance checking aims to detect violations to construction safety regulations. Traditionally, this process is conducted manually by an experienced safety manager on site, which cannot guarantee that noncompliance is identified and resolved in a timely manner (Tang et al. 2020) to prevent potential safety incidents proactively. Many research efforts have thus been devoted to automating the field compliance checking process. For example, computer vision techniques have been widely used to monitor site conditions by detecting the existence of certain protection items such as hard hats and personal fall protection systems (Fang et al. 2018; Fang et al. 2019; Nath et al. 2020), tracking and predicting the trajectory of site objects such as workers and equipment (Tang et al. 2019; Roberts et al. 2020), and recognizing workers' operations and interactions with their environment (Teizer 2015; Zhang et al. 2015b; Park and Brilakis 2016; Tang et al. 2020). Despite these efforts to collect and analyze site information, compliance decisions are usually made in a rough way, without sufficiently considering different situations and/or exceptions as described in construction safety regulations.

Extracting safety requirements and knowledge from regulatory documents requires much manual work, which is expensive and error-prone (Salama and El-Gohary 2013; Zhang and El-Gohary 2013; Zhang and El-Gohary 2019). For example, current efforts that aim to model hazard or risk knowledge from industry safety best practice reports (e.g., Lu 2015) can develop safety checking ontology(ies) by classifying concepts and specifying relations between concepts manually. Such methods fall short in their scalability, because of the need to consider a large number of safety regulatory documents from different jurisdictional levels, the large number of provisions in each of these documents, and the potential change to the documents as safety knowledge and site practices improve over time. Hence, significant manual effort is required for initial extraction as well as future updates.

Information extraction offers a potential solution to automatically extract information from unstructured text. However, extracting information from construction safety regulations is rather difficult for two reasons. First, it is difficult to extract information without (or with limited) human involvement. For example, rule-based information extraction efforts can depend heavily on human interpretation and predefined extraction rules (Zhang and El-Gohary 2013; Zhou and El-Gohary 2017). Traditional machine learning-based methods require highly engineered features obtained through trial and error (Liu and El-Gohary 2017). To further reduce human assistance, an end-to-end method that does not rely on feature engineering is desired. Second, it is difficult to achieve good performance with unsupervised methods due to the complexity in unstructured text. Such complexity can include various descriptions about the same semantic information element, nested conditions in describing a scenario, and ambiguities in the text itself. To achieve good performance given such complex text, the characteristics of the text such as syntactic and semantic features, context, and discourse need to be taken into account. Therefore, there is a need to develop an information extraction method to extract safety requirements from construction safety regulations, with less human assistance and good performance, to support automated field compliance checking for detecting field non-compliances with these regulations.

To address this gap, this paper proposes a deep learning (DL)-based information extraction method to automatically extract relations (e.g., is above, cause, and greater or equal to) that link fall protection related entities (e.g., employee, scaffold, and toeboards), with the latter extracted from construction safety regulations in a previous study.

BACKGROUND

To extract the relations that describe fall protection requirements, this study focuses on the relation extraction task which is one important aspect of information extraction. Relation extraction is the task of recognizing and classifying semantic relations from unstructured text into several predefined classes (Bach and Badaskar 2007; Nguyen and Grishman 2015). For example, in the sentence "Defective safety net components shall be removed from service", this study tries to classify the relation between "safety_net_components" and "defective" as "Is", and the relation between "safety_net_components" and "service" as "Keep_From", as is illustrated in Fig. 1. In this way, relation extraction helps identify the interconnections between different entities, for developing a semantically rich and structured representation of the extracted requirements.

Is(safety_net_components, defective)

<**attribute**>Defective<**/attribute**> <**equipment**>safety net components<**/equipment**> shall be removed from <**other entity**>service<**/other entity**>.

Keep_From(safety_net_components, service)

Figure 1. Example of the entity-relation triples after relation extraction.

Various relation extraction methods have been proposed over the past years, including rule-based and machine learning-based methods (Nebhi 2013; Lai et al 2018; Santus et al 2018). Most recently, DL-based methods have received growing popularity in relation extraction (Jiang et al. 2020). Depending on the types of supervision received, those DL-based methods can be further divided into two categories: distant supervised methods and fully supervised methods. Distant supervised methods learn from unlabeled data with the help of some external knowledge bases. For example, Mintz et al. (2009) used the Freebase (Bollacker et al. 2008), a semantic knowledge base, for distant supervised learning. In general, research on distant supervised methods attempts to experiment with different DL architectures or different knowledge bases for performance improvement. Fully supervised relation extraction methods are more suitable for construction applications, because (1) they do not require external knowledge bases, which are not available for the construction domain; and (2) customized relation classes can be easily incorporated through additional classes/labels.

A limited number of relation extraction efforts have been conducted in the construction domain. For example, Zhang and El-Gohary (2013) proposed a semantic rule-based natural language processing approach to automatically extract requirements including quantitative relations and comparative relations from building codes. Liu and El-Gohary (2021) proposed a dependency parsing framework to automatically extract dependency relations between bridge-related entities.

PROPOSED DEEP LEARNING-BASED APPROACH

Extracting the semantic information elements that describe the relations (e.g., is above, cause, and greater or equal to) between entities from construction safety regulations is formulated as a relation extraction task. A total of 48 relation classes were predefined based on a review of two OSHA sections related to fall protection. The predefined relation classes included all necessary relations involved without redundant expressions. For each predefined relation class, there are two directions associated with it that the model tries to capture: direction from head entity to tail entity, or direction from tail entity to head entity. For example, in the sentence “anticipated loads(1) caused by ice buildup(2) …”, the direction is that (2) causes (1). In the sentence “ladder deflection(1) cause the ladder(2) to …”, the direction is that (1) causes the (2). After predefining the relation classes and preprocessing the raw text, an Attention-based convolutional neural network (CNN) model was implemented to automatically recognize and classify the relations based on their syntactic and semantic features. To improve the model performance, the state-of-the-art continuous bag-of-words (CBOW) embedding (Mikolov et al. 2013) and part-of-speech (POS) embedding were used. The research methodology consisted of four primary steps: text preprocessing, feature preparation, relation extraction model training, and evaluation.

Text preprocessing

Two OSHA sections related to fall protection (1926.501 and 1926.502) were selected to create the dataset for the relation extraction task. The tagging scheme of SemEval-2010 dataset (Hendrickx et al. 2019) from the computational linguistic domain was followed for annotating the dataset. Due to the complex situations considered in each OSHA clause, one sentence can contain multiple entity-relation triples. A total number of 1,153 entity-relation triples were found in the dataset after the annotation. The annotated dataset was then divided into training and testing set using a 9:1 ratio. An example of the annotation is shown in Table 1.

Table 1. Example of an entity-relation triple annotated in the dataset.

Original sentence	Annotated sentence[1]	Relation type	Relation index[2]
Each employee on a walking/working surface shall be protected from objects falling through holes (including skylights) by covers .	Each <**e1**>employee</**e1**> on a <**e2**>walking/working surface</**e2**> shall be protected from objects falling through holes (including skylights) by covers .	Is_Located_At	41

[1] e1= head tag; e2=tail tag.

[2] odd number = relation direction is head to tail; even number: relation direction is tail to head.

Feature preparation

Three additional features were used to further improve the performance of relation extraction: CBOW, POS, and position embeddings. CBOW embedding is one of the state-of-the-art word embeddings pre-trained on Google News dataset (about 100 billion words). It represents the semantics of each word as well as its context in the form of continuous and dense feature vectors, so that words similar in meaning are closer to each other in their embedding space. POS embedding aims to encode the POS tag of each word, which indicates the lexical category of that word, such as noun, verb, and adjective. A total of 15 POS categories were considered and obtained using the Stanford CoreNLP Toolkit (Manning et al. 2014). Position embedding is used to differentiate the importance of each word due to its location in the sentence. This is because usually words closer to the given entities are more informative. Position information is thus calculated with reference to the head entity. For example, in the sentence “All <e1>fall protection</e1> required by

<e2>1926.501</e2> shall …", the distance from the word 'required' to the head entity is 1, which is encoded in its POS embedding.

Relation extraction model training

The Attention-based CNN model contains four main types of layers: embedding layers, convolution layer, attention layer, and multi-layer perceptron layer. The embedding layers consist of three components that correspond to the above-mentioned three features, word, POS, and position embeddings. The word embedding layer uses the pre-trained CBOW embedding as a starting point, then adjusts to the domain-specific semantics during training. The outputs from the three embedding layers are concatenated before being fed into the CNN layers and the attention layer. The convolution process in the CNN layers aims to extract local features by applying different filters. The max-pooling layer in the CNN layers aims to keep the most important features for sentences with variable lengths. The outputs from the CNN layers are represented as sentence convolution vectors. For the attention layer, attention weights are calculated to quantitatively model the contextual relevance of the words. Then attention-based context vectors are calculated as a weighted sum of the words based on their attention weights. The outputs from both the CNN layers and the attention layer, namely sentence convolution vectors and attention-based context vectors, are concatenated together for a full representation of an input sentence. The multi-layer perceptron layer takes in all the concatenated vectors and transforms them into relation class tags as predictions. The model was implemented using PyTorch in Python 3. The detailed architecture of the Attention-based CNN model is shown in Fig. 2.

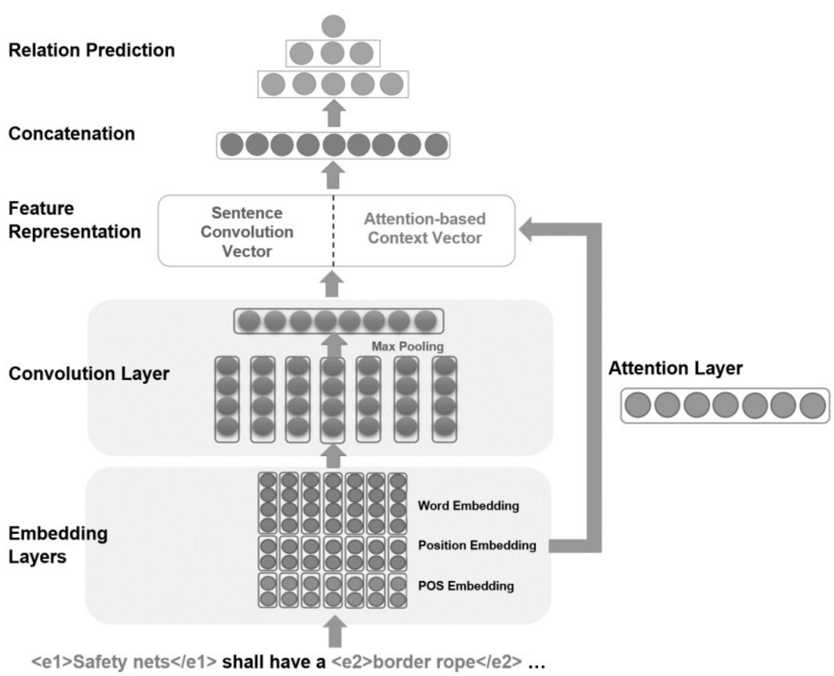

Figure 2. Detailed architecture of the Attention-based CNN model for relation extraction.

Evaluation

The relation extraction performance was evaluated by comparing the predictions from the Attention-based CNN model with the annotated gold standard developed during the text preprocessing step, using the following three metrics: precision, recall, and F-1 measure, as per Eqs. 1-3. Precision is defined as the number of correctly recognized relations divided by the total number of all recognized relations. Recall is defined as the number of correctly recognized relations divided by the total number of all relations in the document. A trade-off between precision and recall is measured by F-1 measure. Due to data imbalance, weighted averages of precision, recall, and F-1 measure were used to evaluate the relation extraction performance. This is because the model tends to perform well on more frequent relation classes. Therefore, less frequent relation classes were given a higher weight, to encourage the model to perform well in those cases.

$$P = \frac{\text{number of correctly recognized relations}}{\text{total number of all recognized relations}} \tag{1}$$

$$R = \frac{\text{number of correctly recognized relations}}{\text{total number of all relations in the document}} \tag{2}$$

$$F = \frac{2\times\mathrm{P}\times\mathrm{R}}{P+R} \tag{3}$$

EXPERIMENTAL RESULTS AND DISCUSSION

The proposed method achieved a weighted precision, recall, and F-1 measure of 82.7%, 81.1%, and 81.3% respectively, which indicates a good relation extraction performance. An error analysis was conducted to identify the sources of errors. First, ambiguity is a major error source, especially when the relations are indicated using prepositions only. For example, in the sentence "<u>employee</u>(1) in a <u>controlled access zone</u>(2)", the actual relation class is "Is_Located_At", with a direction from (1) to (2), since (2) is the location of (1). However, in the phrase "a <u>body belt</u>(1) in a <u>positioning device system</u>(2)", the actual relation class is "Is_Part_Of", with a direction from (1) to (2), since (1) is a component of (2). In both cases, there is only one preposition of "in" that can provide information for predictions, hence the difficulty to distinguish such cases.

Second, frequent omission is another source of error, in which case there is no sufficient information for the model to make the correct predictions. For example, in the phrase "leaving <u>both hands</u>(1) <u>free</u>(2)", the actual relation class is "Is", with a direction from (1) to (2), since (2) is an attribute of (1). However, there are no other words near the given entities supporting such prediction due to omission. Similarly, in the phrase "<u>one-eighth</u>(1) the <u>working length</u>(2)", words for indicating relations between the given entities are omitted, which makes it difficult to predict the correct relations.

Third, a lack of domain knowledge can also cause incorrect predictions. For example, in the sentence "When the <u>employee</u>(1) is progressing up and/or down the <u>ladder</u>(2)", the actual relation class is "Use", with a direction from (1) to (2), since progressing is the action for (1) to use (2). Similarly, in the phrase "If the slope(1) is steeper than one vertical in eight horizontal(2)…", the actual relation class is "Greater_Or_Equal", with a direction from (1) to (2), since a steeper slope has a higher ratio. However, there is no sufficient context, background information, or term explanations related to each OSHA clause. It is, therefore, difficult for the model to make the desired predictions.

CONCLUSION AND FUTURE WORK

In this paper, the authors proposed a deep learning-based information extraction method to extract relations that connect fall protection-related entities from construction safety regulations for supporting automated field compliance checking. The proposed method uses an Attention-based CNN model to recognize and classify the relations based on their syntactic and semantic features and context. To improve the relation extraction performance, three types of features were used: CBOW embedding, POS embedding, and position embedding. The proposed method was tested on two OSHA sections related to fall protection. The experimental results (a weighted precision, recall, and F-1 measure of 82.7%, 81.1%, and 81.3%, respectively) indicate that the proposed method is potentially effective in extracting fall protection-related relations from construction safety regulations.

Two main limitations of the work are acknowledged. First, the size of the dataset used in this study (only two selected OSHA sections) is limited. Second, and as a result, it may not contain sufficient training samples for certain relation classes. More relation classes can be considered when the dataset size grows, especially the relations indicating workers' interactions and operations. This is because different OSHA sections focus on different topics, and thus operations can vary from one topic to another.

To address these limitations, in their future work, the authors plan to add more OSHA sections to the dataset, as well as more necessary relation classes, to further improve the relation extraction performance and generalizability. We will also explore and compare the use of different DL models, such as recurrent neural networks (RNN)-based models, which is another popular branch of models in the natural language processing domain.

REFERENCES

Bach, N., & Badaskar, S. (2007). "A review of relation extraction." *Language & Statistics II*, 2, 1-15.

Bollacker, K., Evans, C., Paritosh, P., Sturge, T., & Taylor, J. (2008, June). "Freebase: a collaboratively created graph database for structuring human knowledge." *Proc., 2008 ACM SIGMOD international conference on Management of data*, 1247-1250.

Fang, W., Ding, L., Luo, H., & Love, P. E. (2018). "Falls from heights: A computer vision-based approach for safety harness detection." *Automat. Constr.*, 91, 53-61.

Fang, W., Zhong, B., Zhao, N., Love, P. E., Luo, H., Xue, J., & Xu, S. (2019). "A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network." *Adv. Eng. Inform*., 39, 170-177.

Hendrickx, I., Kim, S. N., Kozareva, Z., Nakov, P., Séaghdha, D. O., Padó, S., ... & Szpakowicz, S. (2019). "SemEval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals." *arXiv preprint arXiv*:1911.10422.

Jiang, H., Bao, Q., Cheng, Q., Yang, D., Wang, L., & Xiao, Y. (2020). "Complex Relation Extraction: Challenges and Opportunities." *arXiv preprint arXiv*:2012.04821.

Liu, K., & El-Gohary, N. (2017). "Ontology-based semi-supervised conditional random fields for automated information extraction from bridge inspection reports." *Automat. Constr*., 81, 313-327.

Lai, S., Leung, K. S., & Leung, Y. (2018). "SUNNYNLP at SemEval-2018 Task 10: A support-vector-machine-based method for detecting semantic difference using taxonomy and word embedding features." *Proc., 12th international workshop on semantic evaluation* (pp. 741-746).

Liu, K., & El-Gohary, N. (2021). "Semantic neural network ensemble for automated dependency relation extraction from bridge inspection reports." *J. Comput. Civ. Eng*., 35(4), 04021007.

Lu, Y., Li, Q., Zhou, Z., & Deng, Y. (2015). "Ontology-based knowledge modeling for automated construction safety checking." *Saf. Sci*., 79, 11-18.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. (2014). "The Stanford CoreNLP natural language processing toolkit." *Proc., Association for Computational Linguistics (ACL)*, 55-60.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). "Efficient estimation of word representations in vector space." *arXiv preprint arXiv*:1301.3781.

Mintz, M., Bills, S., Snow, R., & Jurafsky, D. (2009). "Distant supervision for relation extraction without labeled data." *Proc., Joint Conference of Association for Computational Linguistics (ACL) and Natural Language Processing of the AFNLP*, 1003-1011.

Nath, N. D., Behzadan, A. H., & Paal, S. G. (2020). "Deep learning for site safety: Real-time detection of personal protective equipment." *Automat. Constr*., 112, 103085.

Nebhi, K. (2013). "A rule-based relation extraction system using DBpedia and syntactic parsing." *Proc., 12th International Semantic Web Conference*.

Nguyen, T. H., & Grishman, R. (2015). "Relation extraction: Perspective from convolutional neural networks." *Proc., Workshop on Vector Space Modeling for Natural Language Processing*, 39-48.

Park, M. W., & Brilakis, I. (2016). "Continuous localization of construction workers via integration of detection and tracking." *Automat. Constr*., 72, 129-142.

Roberts, D., Torres Calderon, W., Tang, S., & Golparvar-Fard, M. (2020). "Vision-based construction worker activity analysis informed by body posture." *J. Comput. Civ. Eng*., 34(4), 04020017.

Salama, D. A., & El-Gohary, N. M. (2013). “Automated compliance checking of construction operation plans using a deontology for the construction domain.” *J. Comput. Civ. Eng*, 27(6), 681-698.

Santus, E., Biemann, C., & Chersoni, E. (2018). “Bomji at semeval-2018 task 10: Combining vector-, pattern-and graph-based information to identify discriminative attributes.” *arXiv preprint arXiv*:1804.11251.

Tang, S., Golparvar-Fard, M., Naphade, M., & Gopalakrishna, M. M. (2019). “Video-based activity forecasting for construction safety monitoring use cases.” *J. Comput. Civ. Eng*., ASCE, Reston, VA, 204-210.

Tang, S., Roberts, D., & Golparvar-Fard, M. (2020). “Human-object interaction recognition for automatic construction site safety inspection.” *Automat. Constr*., 120, 103356.

Teizer, J. (2015). “Status quo and open challenges in vision-based sensing and tracking of temporary resources on infrastructure construction sites.” *Adv. Eng. Inform*., 29(2), 225-238.

Zhang, J., & El-Gohary, N. M. (2013). “Semantic NLP-based information extraction from construction regulatory documents for automated compliance checking.” *J. Comput. Civil Eng*., 30(2), 04015014.

Zhang, R., & El-Gohary, N. (2019). “A machine learning-based approach for building code requirement hierarchy extraction.” *Proc., 7th CSCE Int. Constr. Spec. Conf.,* CSCE, Montreal, Canada.

Zhang, S., Teizer, J., Pradhananga, N., & Eastman, C. M. (2015b). “Workforce location tracking to model, visualize and analyze workspace requirements in building information models for construction safety planning.” *Automat. Constr*., 60, 74-86.

Zhou, P., & El-Gohary, N. (2017). “Ontology-based automated information extraction from building energy conservation codes.” *Automat. Constr*., 74, 103-117.