

Deep learning-based named entity recognition from construction safety regulations for automated field compliance checking

XiYu Wang¹ and Nora El-Gohary, A.M.ASCE²

¹Graduate 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

²Associate 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

Automated safety compliance checking aims to detect field violations to construction safety regulations. Recent research and system development efforts have made good progress on automated tracking of labor and equipment towards improved violation detection and safety compliance. However, extracting and modeling safety requirements for supporting automated violation detection or safety alert systems remains highly manual. Towards addressing this gap, information extraction provides an opportunity to automatically extract safety requirements from regulatory documents for comparisons with field information to detect violations. However, existing information extraction methods fall short in their scalability and/or accuracy. To address this need, this paper proposes a deep learning-based information extraction method for extracting entities that describe fall protection requirements from construction safety regulations for supporting automated field compliance checking. The proposed method uses a hybrid bidirectional long short-term memory (BiLSTM) and convolutional neural network (CNN) model for recognizing the entities. The proposed method was implemented and tested on four selected Occupational Safety and Health Administration (OSHA) sections related to fall protection. It has achieved an average precision, recall, and F-1 measure of 81.5%, 80.3%, and 80.9%, respectively, which indicates good named entity recognition performance. The paper discusses the proposed method and experimental results, and outlines directions for further performance improvement.

INTRODUCTION

A large number of fall accidents happen on construction sites, with field non-compliance being a major cause. Statistics shows that at least 60,000 fatalities occur yearly in the construction sector worldwide (Biggs and Biggs 2013). Among all accident types, falling accounts for more than 30% of all deaths (Jebelli et al. 2016). A review of the fatality reports reveals that most fall fatalities were due to field non-compliance. For example, according to the Construction FACE Database, 54% of the fall fatalities occurred because personal fall arrest systems (PFAS) were not available and 23% because the decedents were not using PFAS even when they had access to them (Dong et al. 2017). Other common frequent field non-compliance cases related to fall fatalities include damaged or no personal protective equipment (PPE), inoperative or inappropriate equipment, and wrong poses or work sequence (Chi and Lin 2018; Teo et al. 2005; Huang and Hinze 2003).

Automated safety compliance checking aims to detect field violations to construction safety regulations to prevent workers from potential safety incidents. Many research efforts have been undertaken to propose solutions for this purpose. For example, computer vision techniques have been used to collect site visual information for determining compliance. Such information

includes the existence of PPE, the operations of the workers, and the trajectory of labors and equipment (Fang et al. 2018; Fang et al. 2019; Nath et al. 2020; Tang et al. 2019; Roberts et al. 2020; Teizer 2015; Zhang et al. 2015b; Park and Brilakis 2016; Tang et al. 2020). However, despite the importance of those previous efforts, three main knowledge gaps still exist. First, in many efforts (e.g., Nath et al. 2020), violation decisions are usually made without considering different situations as described in the construction safety regulations. Construction safety regulations, on the other hand, provide a large amount of details about the existence or properties of fall protection with respect to different equipment, facilities, and operations, which can be used as reliable guidance for field compliance checking. Second, extracting and modeling safety requirements or knowledge remains highly manual. For example, many efforts use manual ontology-based approaches to model hazard or risk knowledge from fatality reports or industry safety best practice reports (e.g., Xing et al. 2019; Lu et al. 2015; Zhang et al. 2015a; Zhong et al. 2020; Chi et al. 2014; Ding et al. 2016). Considering the amount of requirements to be extracted from regulations, an automated approach is desired, in which case information extraction (IE) provides a solution. Third, existing IE methods fall short in their scalability and/or accuracy. Rule-based IE methods (Zhang and El-Gohary 2013; Zhou and El-Gohary 2017) require significant amount of human effort to discover the text patterns and develop the corresponding IE rules. Traditional machine learning (ML)-based methods do require less human involvement, but their performance depends on the quality of the engineered features. Deep learning (DL) approaches are promising in terms of reducing manual effort and improving performance.

To address these gaps, this paper proposes a DL-based method to automatically extract entities that describe fall protection requirements from construction safety regulations for comparisons with field information to detect violations. To the best of the authors' knowledge, it is the first attempt to automatically extract fall protection-related entities from construction safety regulations for field compliance checking purposes.

BACKGROUND

Depending on the semantic information elements to be extracted, IE from construction safety regulations can involve two tasks: (1) named entity recognition that aims to identify entities from unstructured text and classify them into pre-defined entity classes, and (2) relation extraction that aims to identify relations between entities and classify them into predefined relation types. To extract entities that describe fall protection requirements, this study focuses on the named entity recognition task.

In recent years, DL-based methods have achieved significant improvement for tasks in many research domains, including construction. Various DL models and training techniques have been proposed. For example, Zhang and El-Gohary (2019) have proposed a long short-term memory (LSTM)-based method to extract building-code requirement hierarchies. Zhang and El-Gohary (2020) proposed an LSTM-based method to generate semantically enriched building-code sentences. For named entity recognition tasks, DL-based methods use stacked neural networks to extract patterns from unstructured text, and output entity class predictions in an end-to-end manner, thus requiring less human effort. Recent research efforts have proposed various DL-based methods to improve named entity recognition performance. Those methods include: (1) proposing new DL architecture based on existing ones, (2) creating a hybrid model by combining existing ones, and (3) developing state-of-the-art word embeddings. Those word embeddings, such as embeddings from language models (ELMo) (Peters et al. 2018) and the global vectors (GloVe) embeddings

(Pennington et al. 2014), are pre-trained on large amount of general social media text and thus can represent the semantics of words in the form of dense and rich feature vectors.

PROPOSED DEEP LEARNING-BASED METHOD

Extracting the semantic information elements that describe fall protection requirements from construction safety regulations for supporting compliance checking is defined as a named entity recognition task. To recognize and classify named entities, fourteen entity classes were first defined, based on a thorough review of related OSHA sections, and research work on safety knowledge modeling. The pre-defined entity classes are person, equipment, clause, hazard, facility, location, operation, material, property, date, other attribute, quantity value, quantity unit, and other entity. Examples of the most frequent entity names in each entity class are shown in Table 1. Then, a hybrid BiLSTM-CNNs model was used to capture the syntactic and semantic features from the noisy text and to produce entity class predictions. To further improve the named entity recognition performance, three types of features were used, including GloVe embedding, word-level features, and character-level features. The proposed method consists of four primary steps, as per Figure 1: data preprocessing, feature preparation, named entity recognition model training, and evaluation.

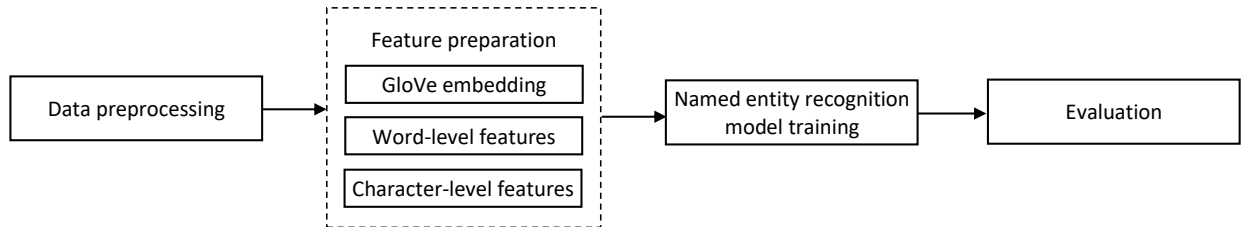


Figure 1. Proposed deep learning-based method for extracting entities from construction safety regulations

Table 1. Examples of the most frequent entity names in each pre-defined class.

Class	Examples
Person	Employer, employee, competent person, qualified person, engineer
Equipment	Scaffold, ladder, guardrail system, platform, rope
Clause	This section, paragraph (k) of 1926.502, other criteria
Hazard	Falling, damage, snag of clothing, tripping, failure
Facility	Beams, roof, passageway, metal decking, other structure
Location	Walking/working surface, work area, storage area
Operation	Hoisting operation, dropping, rigging, lifting
Material	Reinforcing steel, wood, mortar, debris
Property	Weight, modulus of elasticity, length, diameter, direction
Date	January 1, 1998
Other attribute	In use, center-to-center, under construction
Quantity value	6, 3.3, half, one
Quantity unit	Feet, times, meters, inches
Other entity	Fall protection plan, requirements

Data preprocessing

Four OSHA sections related to fall protection were selected to create the dataset for the named entity recognition task, which covers various topics such as general fall protection requirements, fall protection systems, scaffolds, and ladders. A total of 625 sentences were collected and annotated based on a BIO tagging scheme using Doccano (Nakayama et al. 2018). The BIO tagging scheme aims to distinguish adjacent entities that could have the same entity class tag. It uses “B” to denote the beginning of an entity, “I” to denote the inside of an entity, and “O” to denote the absence of an entity. Table 2 shows an example clause annotated using the BIO tagging scheme. To simplify the entity class tags, abbreviations were used, including person (PER), equipment (EQU), clause (CLA), hazard (HAZ), facility (FAC), location (LOC), operation (OPE), material (MAT), property (PRO), date (DAT), other attribute (ATT), quantity value (QUA), quantity unit (UNI), and other entity (ENT).

Table 2. Example clause annotated using the BIO tagging scheme and entity classes.

Original sentence	Annotated sentence
Each employee on a walking/working surface shall be protected from objects falling through holes by covers.	<O>Each</O> <B-PER>employee</B-PER> <O>on</O> <O>a</O> <B-LOC>walking/working</B-LOC> <I-LOC>surface</I-LOC> <O>shall</O> <O>be</O> <O>protected</O> <O>from</O> <B-ENT>objects</B-ENT> <O>falling</O> <O>through</O> <B-ENT>holes</B-ENT> <O>by</O> <B-EQU>covers</B-EQU> <O>.</O>

Feature preparation

To further improve the named entity recognition performance, three types of features were used: GloVe embedding, word-level features, and character-level features. The GloVe embedding is one of the state-of-the-art word embeddings pre-trained on 6 billion words of Wikipedia text. Since the GloVe embedding does not contain capitalization information, which is important in preserving the original meaning, additional word-level and character-level features were used in this study. The following word-level features were used: (1) if the word is all lower-cased, (2) if the word is all upper-cased, (3) if the word contains numbers, and (4) if the word contains capital letters. The used character-level features included: (1) if the character is punctuation, (2) if the character is a digit, (3) if the character is uppercase, and (4) if the character is lowercase.

Named entity recognition model training

The hybrid BiLSTM-CNNs model was used to capture the syntactic and semantic features from unstructured construction safety regulations and to produce entity class predictions in an end-to-end manner. The hybrid model was chosen for better performance, because it is expected to combine the benefits of both types of models that are designed with different strengths. BiLSTM is better at capturing context and long dependency, while CNN is better at capturing character-level information, both which are important for making correct predictions. This hybrid model consists of three main types of layers: embedding layers, BiLSTM layer, and multi-layer perceptron (MLP) layer. The embedding layers have three components: GloVe embedding layer, word-level and character-level feature embedding layer, and a CNN-extracted feature embedding layer. The outputs from the embedding layers are concatenated before being fed into the BiLSTM layer. The BiLSTM layer then computes the feature values for each word by considering its context. The MLP layer predicts entity class tags as the final output by transforming the feature values from the BiLSTM layer into tag scores. To optimize the hybrid model, cross-entropy loss was used. The

model was implemented using Keras in Python 3. The architecture of the BiLSTM-CNNs model is shown in Figure 2.

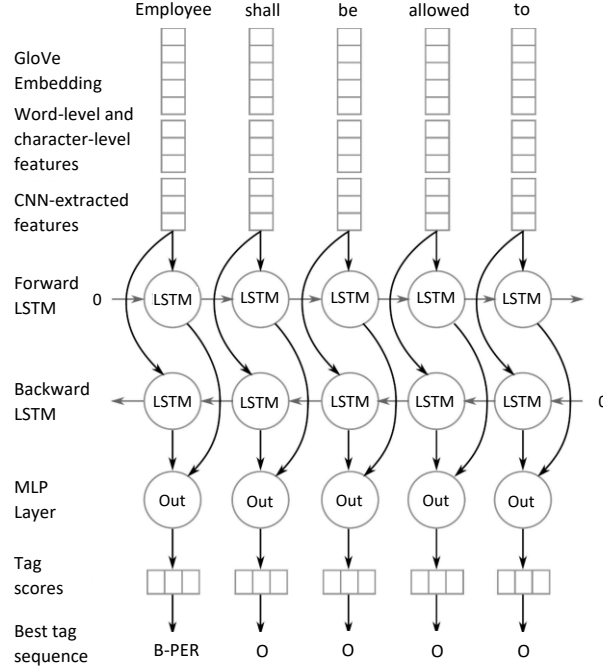


Figure 2. Architecture of the BiLSTM-CNNs model for recognizing named entities

Evaluation

The evaluation of the named entity recognition performance was conducted by comparing the predicted results from the hybrid BiLSTM-CNNs model with the annotated gold standard. Three metrics were used for the evaluation, as per Eqs. 1-3: precision, recall, and F-1 measure. For a specific entity class, the precision was calculated as the number of correctly recognized entities divided by the total number of all recognized entities for that entity class. The recall was calculated as the number of correctly recognized entities divided by the total number of entities in the document for that entity class. Macro average of precision and recall over all entity classes were used for the evaluation because of data imbalance. Macro average, rather than micro average, computes the metrics independently for each entity class, and takes the average without considering the frequency of each entity class. Therefore, macro average can be a better metric in our case. F-1 measure is the harmonic mean of both precision and recall.

$$P_i = \frac{\text{number of correctly recognized entities for class } i}{\text{total number of all recognized entities for class } i} \quad (1)$$

$$R_i = \frac{\text{number of correctly recognized entities for class } i}{\text{total number of entities in the document for class } i} \quad (2)$$

$$F = \frac{2 \times P \times R}{P + R} \quad (3)$$

EXPERIMENTAL RESULTS AND DISCUSSION

The proposed method achieved an average precision, recall, and F-1 measure of 81.5%, 80.3%, and 80.9%, respectively. These results indicate good named entity recognition performance. The difference in the performance between general social media text (e.g., 91.6% of F-1 on CoNLL-2003 dataset) and construction safety regulatory documents can result from three main factors. First, domain-specific data are more complex than general social media text. For construction safety regulations, such complexity includes different descriptions of the same semantic information element, various patterns and arrangement of clauses, nested conditions for describing a scenario, and ambiguities in the text meaning itself. Second, the pre-trained word embedding may not align well with the semantics of domain-specific data. The GloVe embedding used in the embedding layer was pre-trained on general social media text, and is expected to adjust itself to domain-specific semantics during training. However, considering our domain-specific dataset size versus the Wikipedia text size, the embeddings may not be able to learn enough semantic patterns in the domain-specific data. Third, not enough domain-specific information is provided in the text for the model to make accurate predictions. A review of the results reveals that domain-specific compound nouns are often misclassified. For example, the model predicts the phrase “adjustable suspension scaffold” as an ATT (other attribute) plus an EQU (equipment) based on the syntactic and semantic features. However, in our domain, “adjustable suspension scaffold” is a specific kind of equipment. To make an accurate prediction, the model needs to consider various sources of information, including syntactic and semantic features, context and discourse, etc. However, since each clause from the OSHA regulations is addressing one specific scenario, not much context, background information, or term explanations can be utilized for the model to learn to make the predictions.

The most frequent misclassification occurred for the EQU (equipment) entity class. For example, 9 FAC (facility) entities were recognized as EQU, with a total of only 49 extracted FAC entities. 7 ATT (other attribute) entities were recognized as EQU, with a total of 77 extracted ATT entities. This misclassification is mostly due to the imbalance of the dataset. Since EQU is the most frequent class in the dataset, the DL model learns to predicts an entity more often as EQU to achieve minimal error.

Misclassifications were seen for other entity classes such as LOC (location), PRO (property), ATT (other attribute), and ENT (other entity). A major cause for those misclassifications can be word-sense ambiguities, where the same word has different meanings in different contexts. For example, the word “level” can be regarded as LOC in the phrase “above lower levels”, since it means some surface or position, but can be regarded as ATT in the sentence “footings should be level, sound and rigid”, since it means that the footings should have no slopes or bumps. Similarly, the word “direction” can be regarded as PRO in the phrase “in an upward direction”, since it means a path for movement, but can be regarded as ENT in the phrase “under the supervision and direction of a competent person”, since it means guidance or management. However, it is rather difficult for the model to determine which meaning applies for which case, when there are limited training samples and when there is no sufficient context for each clause.

CONCLUSION AND FUTURE WORK

In this paper, the authors proposed a deep learning-based information extraction method for extracting entities that describe fall protection requirements from construction safety regulations for supporting automated field compliance checking. The proposed method uses a hybrid BiLSTM-CNNs model with three types of features: GloVe embedding, word-level features, and

character-level features. The proposed method was tested on four selected OSHA sections related to fall protection. The results indicate that the proposed method can automatically recognize fall protection-related entities from construction safety regulations with good performance.

Two main limitations of this work are acknowledged. First, due to the variability in expressing the same entity in the text, the extracted requirements included redundant expressions and referential ambiguity. For example, “two-point adjustable suspension scaffold” and “two-point scaffold” were recognized as two different entities. However, since they appear in the same clause consecutively, they are actually referring to the same equipment. Such way of expression happens frequently in the document where the second mention of the same entity can have the detailed information omitted for simplicity. Second, omission can happen in the document, which can lead to missing information in the extracted requirements. For example, in the sentence “12 inch nylon, or equivalent, with a maximum length...”, the compliance checking subject after the word “equivalent” was omitted. Since missing information does not happen frequently in the documents, the authors did not conduct a separate DL-based ellipsis resolution task to complete the omitted information. Instead, the authors plan to solve this problem in future research.

In future work, the authors plan to address the issues of redundant expressions and referential ambiguity by conducting coreference resolution. The authors also plan to further extract the relations from construction safety regulations to add the needed interlinks to the isolated entities extracted in this work.

REFERENCES

- Biggs, H. C., & Biggs, S. E. (2013). “Interlocked projects in safety competency and safety effectiveness indicators in the construction sector.” *Saf. Sci.*, 52, 37-42.
- Chi, C. F., & Lin, S. Z. (2018). “Classification scheme and prevention measures for caught-in-between occupational fatalities.” *Appl. Ergon.*, 68, 338-348.
- Chi, N. W., Lin, K. Y., & Hsieh, S. H. (2014). “Using ontology-based text classification to assist Job Hazard Analysis.” *Adv. Eng. Inform.*, 28(4), 381-394.
- Chiu, J. P., & Nichols, E. (2016). “Named entity recognition with bidirectional LSTM-CNNs.” *Trans. Assoc. Comput. Linguist.*, 4, 357-370.
- Ding, L. Y., Zhong, B. T., Wu, S., & Luo, H. B. (2016). “Construction risk knowledge management in BIM using ontology and semantic web technology.” *Saf. Sci.*, 87, 202-213.
- Dong, X. S., Largay, J. A., Choi, S. D., Wang, X., Cain, C. T., & Romano, N. (2017). “Fatal falls and PFAS use in the construction industry: Findings from the NIOSH FACE reports.” *Accid. Anal. Prev.*, 102, 136-143.
- 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.
- Huang, X., & Hinze, J. (2003). “Analysis of construction worker fall accidents.” *J. Constr. Eng. Manag.*, 129(3), 262-271.
- Jebelli, H., Ahn, C. R., & Stentz, T. L. (2016). “Fall risk analysis of construction workers using inertial measurement units: Validating the usefulness of the postural stability metrics in construction.” *Saf. Sci.*, 84, 161-170.
- Lu, Y., Li, Q., Zhou, Z., & Deng, Y. (2015). “Ontology-based knowledge modeling for automated

- construction safety checking.” *Saf. Sci.*, 79, 11-18.
- Nakayama, H., Kubo, T., Kamura, J., Taniguchi, Y., & Liang, X. (2018). “doccano: Text annotation tool for human.” *Github*, <github.com/chakkiworks/doccano> (June. 26, 2020).
- 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.
- Park, M. W., & Brilakis, I. (2016). “Continuous localization of construction workers via integration of detection and tracking.” *Automat. Constr.*, 72, 129-142.
- Pennington, J., Socher, R., & Manning, C. (2014). “Glove: Global vectors for word representation.” *Proc., Empirical Methods in Natural Language Processing (EMNLP) Conf.*, 1532-1543.
- Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). “Deep contextualized word representations.” *arXiv preprint arXiv:1802.05365*.
- Roberts, D., Torres Calderon, W., Tang, S., & Golparvar-Fard, M. (2020). “Vision-based construction worker activity analysis informed by body posture.” *J. Comput. Civil Eng.*, 34(4), 04020017.
- Tang, S., Golparvar-Fard, M., Naphade, M., & Gopalakrishna, M. M. (2019). “Video-based activity forecasting for construction safety monitoring use cases.” *J. Comput. Civil 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.
- Teo, E. A. L., Ling, F. Y. Y., & Chong, A. F. W. (2005). “Framework for project managers to manage construction safety.” *Int. J. Constr. Proj. Manag.*, 23(4), 329-341.
- Xing, X., Zhong, B., Luo, H., Li, H., & Wu, H. (2019). “Ontology for safety risk identification in metro construction.” *Comput Ind.*, 109, 14-30.
- 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, R., & El-Gohary, N. (2020). “A Machine-Learning Approach for Semantically-Enriched Building-Code Sentence Generation for Automatic Semantic Analysis.” *Proc., Construction Research Congress (CRC) Conf.*, ASCE, Reston, VA, 1261-1270.
- Zhang, S., Boukamp, F., & Teizer, J. (2015a). “Ontology-based semantic modeling of construction safety knowledge: Towards automated safety planning for job hazard analysis (JHA).” *Automat. Constr.*, 52, 29-41.
- 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.
- Zhong, B., Li, H., Luo, H., Zhou, J., Fang, W., & Xing, X. (2020). “Ontology-based semantic modeling of knowledge in construction: classification and identification of hazards implied in images.” *J. Constr. Eng. Manag.*, 146(4), 04020013.
- Zhou, P., & El-Gohary, N. (2017). “Ontology-based automated information extraction from building energy conservation codes.” *Automat. Constr.*, 74, 103-117.