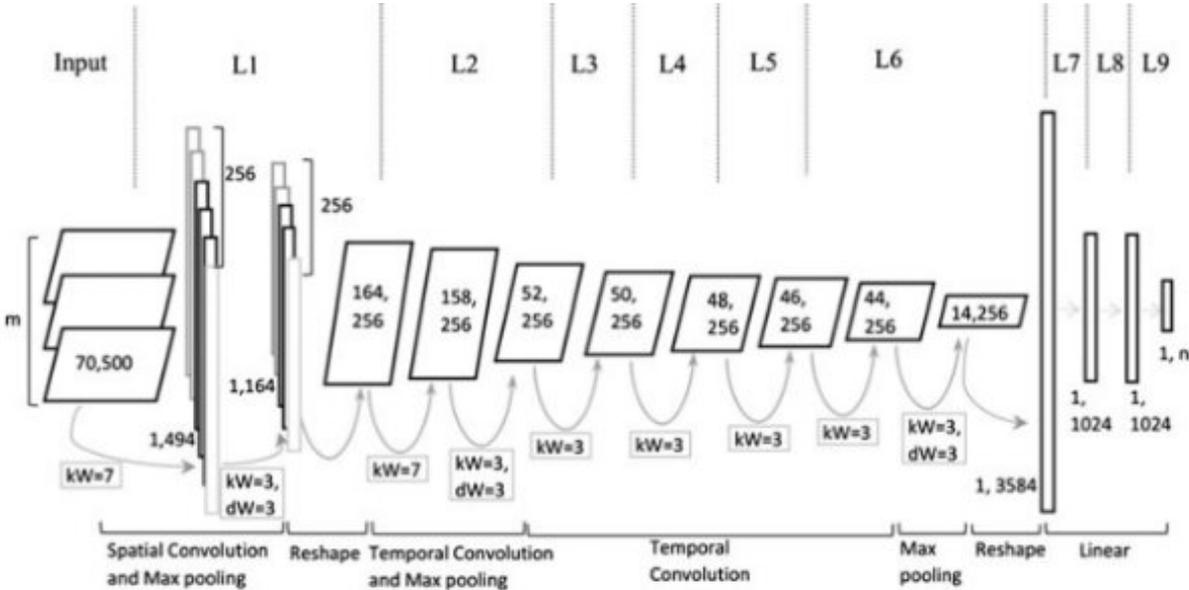


# A multi-representational convolutional neural network architecture for text classification

July 8 2019, by Ingrid Fadelli



Model architecture. Credit: Jin et al, Wiley Computational Intelligence journal.

Over the past decade or so, convolutional neural networks (CNNs) have proven to be very effective in tackling a variety of tasks, including natural language processing (NLP) tasks. NLP entails the use of computational techniques to analyze or synthesize language, both in written and spoken form. Researchers have successfully applied CNNs

to several NLP tasks, including semantic parsing, search query retrieval and text classification.

Typically, CNNs trained for text [classification](#) tasks process sentences on the word level, representing individual words as vectors. Although this approach might appear consistent with how humans process language, recent studies have shown that CNNs that process sentences on the character level can also achieve remarkable results.

A key advantage of character-level analyses is that they do not require prior knowledge of words. This makes it easier for CNNs to adapt to [different languages](#) and acquire abnormal words caused by misspelling.

Past studies suggest that different levels of text embedding (i.e. character-, word-, or -document level) are more effective for different types of tasks, but there is still no clear guidance on how to choose the right embedding or when to switch to another. With this in mind, a team of researchers at Tianjin Polytechnic University in China has recently developed a new CNN architecture based on types of representation typically used in text classification tasks.

"We propose a new architecture of CNN based on multiple representations for text classification by constructing multiple planes so that more information can be dumped into the networks, such as different parts of text obtained through a named entity recognizer or part-of-speech tagging tools, different levels of text embedding or contextual sentences," the researchers wrote in their paper.

The multi-representational CNN (Mr-CNN) model devised by the researchers is based on the assumption that all parts of written text (e.g. nouns, verbs, etc.) play a key role in classification tasks and that different text embeddings are more effective for different purposes. Their model combines two key tools, the Stanford named entity

recognizer (NER) and the part-of-speech (POS) tagger. The former is a method for tagging semantic roles of things in texts (e.g. person, company, etc.); the latter is a technique used to assign part of speech tags to each block of text (e.g. noun or verb).

The researchers used these tools to pre-process sentences, obtaining several subsets of the original [sentence](#), each of which contains specific types of words in the text. They then used the subsets and the full sentence as multiple representations for their Mr-CNN model.

When evaluated on text classification tasks with text from various large-scale and domain-specific datasets, the Mr-CNN model attained remarkable performance, with a maximum of 13 percent error rate improvement on one dataset and an 8 percent improvement on another. This suggests that multiple representations of text allow the network to adaptively focus its attention on the most relevant information, enhancing its classification capabilities.

"Various large-scale, domain-specific datasets were used to validate the proposed architecture," the researchers wrote. "Tasks analyzed include ontology document classification, biomedical event categorization, and sentiment analysis, showing that multi-representational CNNs, which learn to focus attention to specific representations of [text](#), can obtain further gains in performance over state-of-the-art deep neural network models."

In their future work, the researchers plan to investigate whether fine-grained features can help to prevent overfitting of the training dataset. They also want to explore other methods that could enhance the analysis of specific parts of sentences, potentially improving the model's performance further.

**More information:** Rize Jin et al. Multi-representational convolutional

neural networks for text classification, *Computational Intelligence* (2019).  
[DOI: 10.1111/coin.12225](https://doi.org/10.1111/coin.12225)

© 2019 Science X Network

Citation: A multi-representational convolutional neural network architecture for text classification (2019, July 8) retrieved 27 April 2024 from <https://techxplore.com/news/2019-07-multi-representational-convolutional-neural-network-architecture.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.