Initialization	Prec.	Recall	$F_{0.5}$
Random	51.90	12.59	31.96
Word2vec	52.80	12.80	32.49
fastText	51.08	13.63	32.97

Table 4: Results of different embedding initializations on the CoNLL-2013 test set.

ability of MLConv in capturing the context better, thereby favoring more corrections than copying of the source words.

Initialization with Pre-trained Embeddings

We assess various methods of initializing the source and target word embeddings. Table 4 shows the results of initializing the embeddings randomly as well as with *word2vec* and *fastText* on the CoNLL-2013 test set. We train skip-gram models with *word2vec* and use parameters identical to those we use for *fastText*. *fastText* embeddings have access to the character sequences that make up the words and hence are better suited to learn word representations taking morphology into account. We also find that initializing with *fastText* works well empirically, and hence we choose these embeddings to initialize our network when evaluating on benchmark test datasets.

Analysis and Discussion

We perform error type-specific performance comparison of our system and the state-of-the-art (SOTA) system (**?**), using the recently released ERRANT toolkit (**?**) on the CoNLL-2014 test data based on $F_{0.5}$. ERRANT relies on a rule-based framework to identify the error type of corrections proposed by a GEC system. The results on four common error types are shown in Figure 3. We find that our ensembled model with the rescorer (+EO+LM) performs competitively on preposition errors, and outperforms the SOTA system on noun-number, determiner, and subject-verb agreement errors. One of the weaknesses of SMT-based systems is in correction of subject-verb agreement errors, because a verb and its subject can be very far apart within a source sentence. On the other hand, even our single model ($\text{MLConv}_{\text{embed}}$) without rescoring is superior to the SOTA SMT-based system in terms of subject-verb agreement errors, since it has access to the entire source context through the global attention mechanism and to longer target context through multiple layers of convolutions in the decoder.

From our analysis, we find that a convolutional encoder-decoder NN captures the context more effectively compared to an RNN and achieves superior results. However, RNNs can give higher precision, so a combination of both approaches could be investigated in future. Improved language modeling has been previously shown to improve GEC performance considerably. We leave it to future work to explore the integration of web-scale LM during beam search and the fusion of neural LMs into the network. We also find that a simple preprocessing method that segments rare words into sub-words effectively deals with the rare word problem for GEC, and performs better than character-level models and complex word-character models.

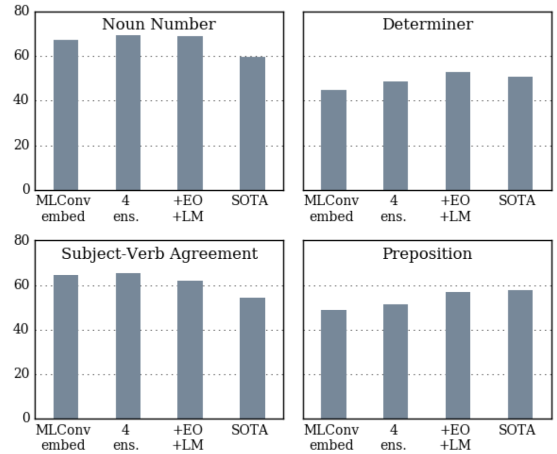

Figure 3: Performance of our models compared to the state-of-the-art system (**?**) on common error types evaluated on the CoNLL-2014 test set based on $F_{0.5}$.

Conclusion

We use a multilayer convolutional encoder-decoder neural network for the task of grammatical error correction and achieve significant improvements in performance compared to all previous encoder-decoder neural network approaches. We utilize large English corpora to pre-train and initialize the word embeddings and to train a language model to rescore the candidate corrections. We also make use of edit operation features during rescoring. By ensembling multiple neural models and rescoring, our novel method achieves improved performance on both CoNLL-2014 and JFLEG data sets, significantly outperforming the current leading SMT-based systems. We have thus fully closed the large performance gap that previously existed between neural and statistical approaches for this task. The source code and model files used in this paper are available at https://github.com/nusnlp/mlconvgec2018.

Acknowledgements

We thank the anonymous reviewers for their feedback. This research was supported by Singapore Ministry of Education Academic Research Fund Tier 2 grant MOE2013-T2-1-150. Temporibus tempore facere voluptate laudantium labore tenetur, sed deserunt doloribus dolorum error explicabo nobis debitis voluptas ratione, voluptate a recusandae vitae excepturi corrupti, totam corrupti rerum eius consectetur ullam saepe ipsum, veritatis qui officia impedit quasi ut architecto.Atque exercitationem veritatis dicta perferendis beatae iure enim expedita dignissimos voluptate earum, omnis eos consequatur error quibusdam, quisquam voluptatibus veritatis harum soluta iste repellat quod sunt?Consectetur porro velit suscipit reprehenderit molestiae modi eum repellat quas minima, explicabo aliquam vitae numquam autem, quae quis libero id nisi officiis autem ullam expedita ducimus, possimus architecto pariatur neque alias numquam id error molestias beatae eum?Hic sapiente neque dolorem, la-

boriosam delectus tempora consequuntur assumenda temporibus cumque nihil consectetur quisquam, enim recusandae