Extractive Summarization of Long Documents by Combining Global and Local Context

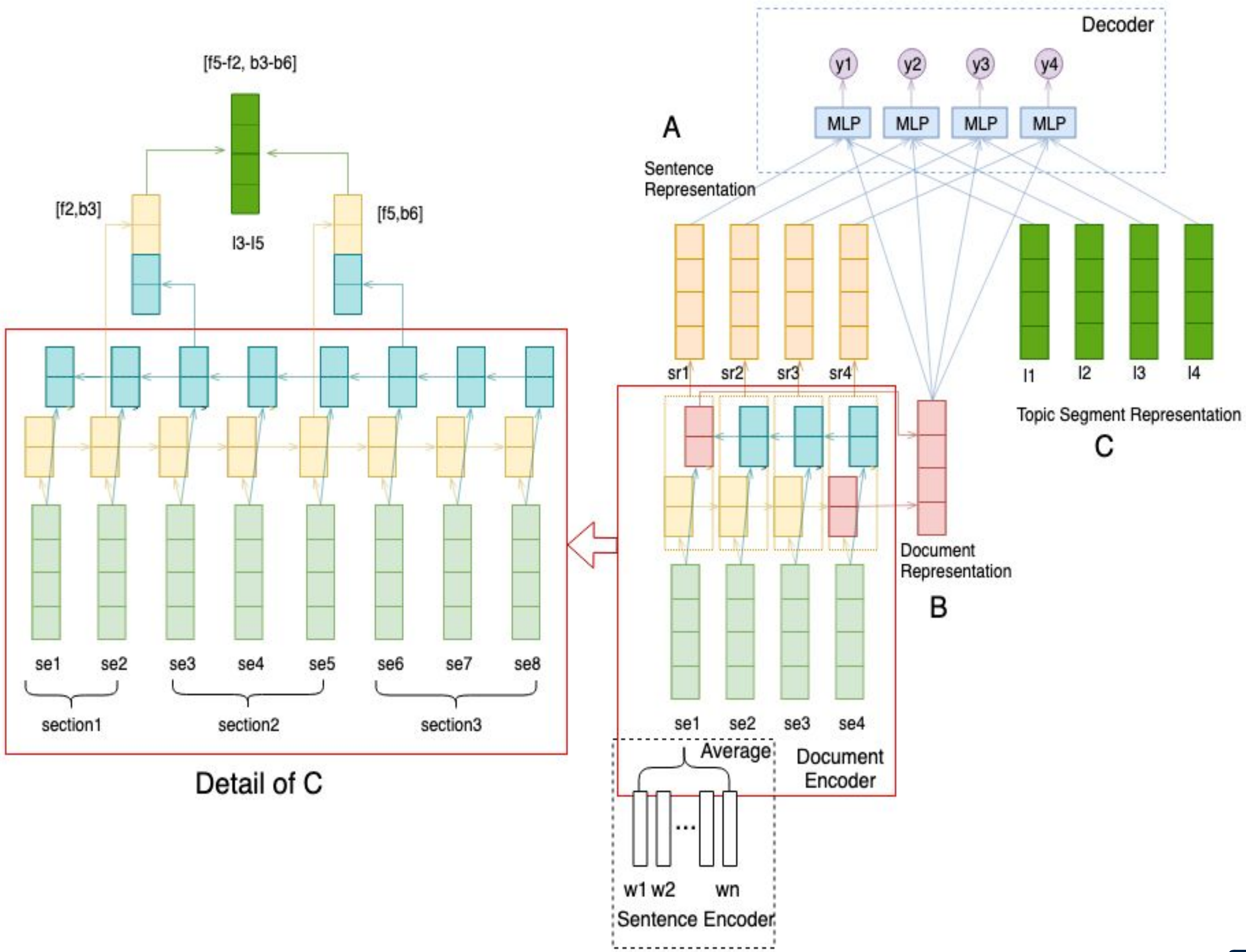
Wen Xiao and Giuseppe Carenini

University of British Columbia Department of Computer Science

xiaowen3@cs.ubc.ca

UBC-NLP Group

Neural Architecture Combining Global and Local Context



Summary of this paper (generated by our system)

S1: *We also ran an ablation study to assess the relative contribution of the global and local components of our approach.*

S2: *In this paper, we propose a novel extractive summarization model especially designed for long documents, by incorporating the local context within each topic, along with the global context of the whole document.*

S3: *Our approach integrates recent findings on neural extractive summarization in a parameter lean and modular architecture.*

S4: *We evaluate our model and compare with previous works in both extractive and abstractive summarization on two large scientific paper datasets, which contain documents that are much longer than in previously used corpora.*

S5: *Our model not only achieves state-of-the-art on these two datasets, but in an additional experiment, in which we consider documents with increasing length, becomes more competitive for longer documents.*

S6: *Furthermore, in an ablation study to assess the relative contributions of the global and the local model we found that, rather surprisingly, the benefits of our model seem to come exclusively from modeling the local context, even for the longest documents.*

S7: *This is a very challenging task, because it arguably requires an in-depth understanding of the source document, and current automatic solutions are still far from human performance.*

*Top 7 sentences with the highest confidence scores

Overall results on the Pubmed and arXiv datasets

Pubmed

Model	*Rouge-1*	*Rouge-2*	*Rouge-L*	*Meteor*
Best Traditional Ext *	33.85	10.73	28.99	-
Best Neural Abs*	35.80	11.05	**31.80**	-
Baseline	42.91	16.65	28.53	21.35
Cheng & Lapata	42.24	15.97	27.88	20.97
SummaRuNNer	42.81	16.52	28.23	21.35
Ours-attentive context	**43.58**	**17.37**	**29.30**	**21.71**
Ours-concat	**43.62**	**17.36**	**29.14**	**21.78**
Lead	33.66	8.94	22.19	16.45
Oracle	53.88	23.05	34.90	24.11

arXiv

Model	*Rouge-1*	*Rouge-2*	*Rouge-L*	*Meteor*
Best Traditional Ext*	39.19	13.89	34.59	-
Best Neural Abs*	38.93	15.37	**35.21**	-
Baseline	44.29	19.17	30.89	20.56
Cheng & Lapata	43.89	18.53	30.17	20.34
SummaRuNNer	43.89	18.78	30.36	20.42
Ours-attentive context	**44.81**	**19.74**	**31.48**	**20.83**
Ours-concat	**44.85**	**19.70**	**31.43**	**20.83**
Lead	35.63	12.28	25.17	16.19
Oracle	55.05	27.48	38.66	23.60

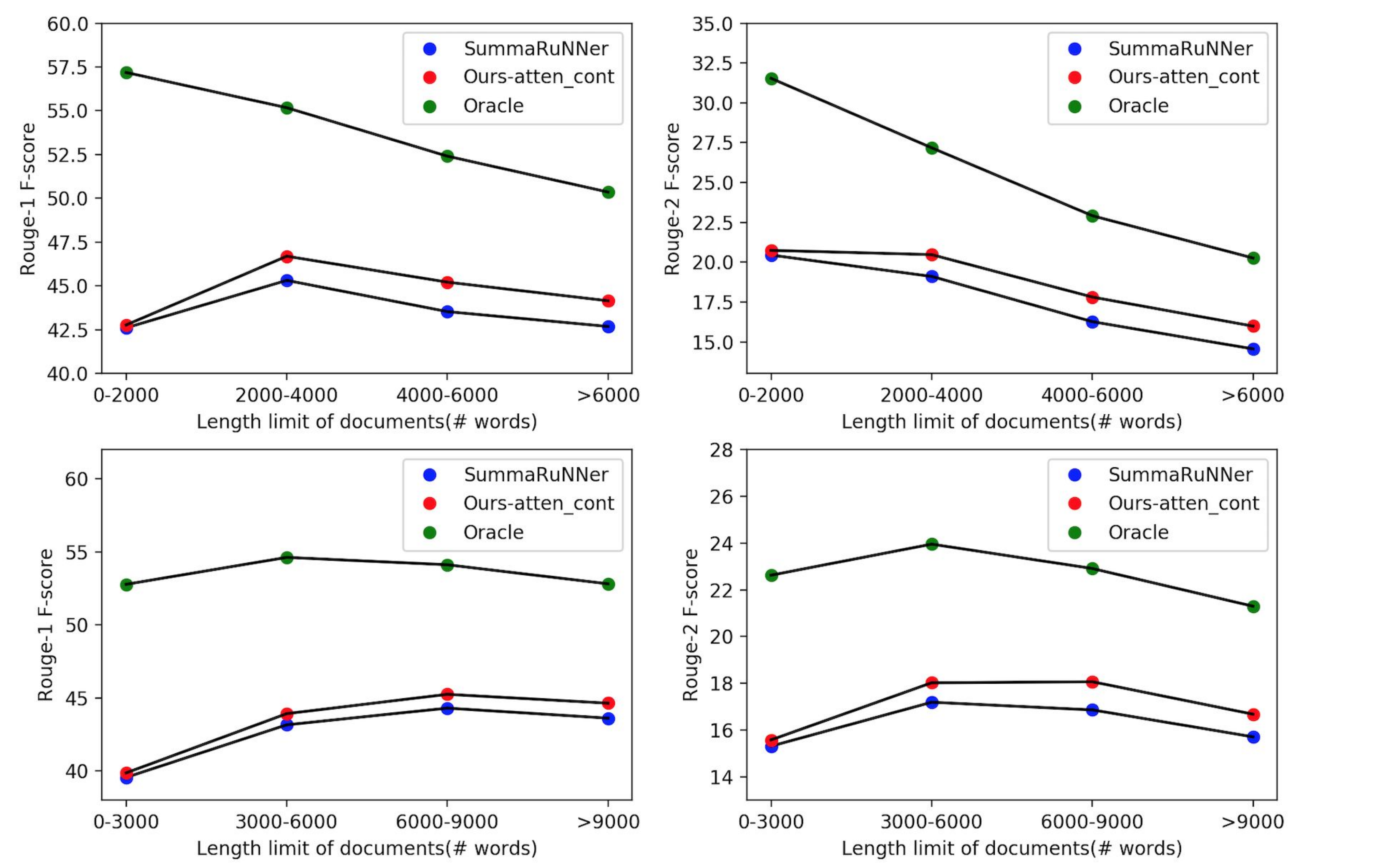
Results that are not significantly distinguished from the best systems are bold.
For models with an *, we report results from (Cohan et al., 2018).

Statistics of popular summarization datasets

Datasets	# docs	avg. doc. length	avg. summ. length
CNN	92K	656	43
Daily Mail	219K	693	52
NY Times	655K	530	38
PubMed	133K	3016	203
arXiv	215K	4938	220

*lengths are In terms of words

Higher Gains for Longer Documents



Upper: Pubmed, Bottom: arXiv

Ablation Study

arXiv

Model	*ROUGE-1(+l/+g)*	*ROUGE-2(+l/+g)*	*ROUGE-L(+l/+g)*
BSL	42.91 (na/na)	16.65 (na/na)	28.53 (na/na)
BSL+l	**43.57** (+.66/na)	**17.35** (+.7/na)	**29.29** (+.76/na)
BSL+g	42.90 (na/-.01)	16.58 (na/-.07)	28.36 (na/-.17)
BSL+l+g	**43.58** (+.68/+.01)	**17.37** (+.79/+.02)	**29.30** (+.94/+.01)
BSL	42.95 (na/na)	14.85 (na/na)	28.66 (na/na)
BSL+l	**44.01** (+1.06/na)	**15.95** (+1.1/na)	**29.68** (+1.02/na)
BSL+g	43.05 (na/+.1)	14.91 (na/+.06)	28.57 (na/-.09)
BSL+l+g	**44.17** (+1.12/+.16)	**16.01** (+1.1/+.06)	**29.72** (+1.15/+.04)

* Similar results on Pubmed

- On both datasets the performance significantly improves when local topic information (i.e. local context) added.
- The improvement is even greater when we only consider long documents.
- Adding a representation of the whole document (i.e. global context) never significantly improves performance.

Scan here for the whole paper

Whole Paper

[1] Arman Cohan, Franck Dernoncourt, Doo Soon Kim, Trung Bui, Seokhwan Kim, Walter Chang, and Nazli Goharian. 2018. A discourse-aware attention model for abstractive summarization of long documents. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), pages 615–621, New Orleans, Louisiana.