Ahmet Üstün, Rob van der Goot

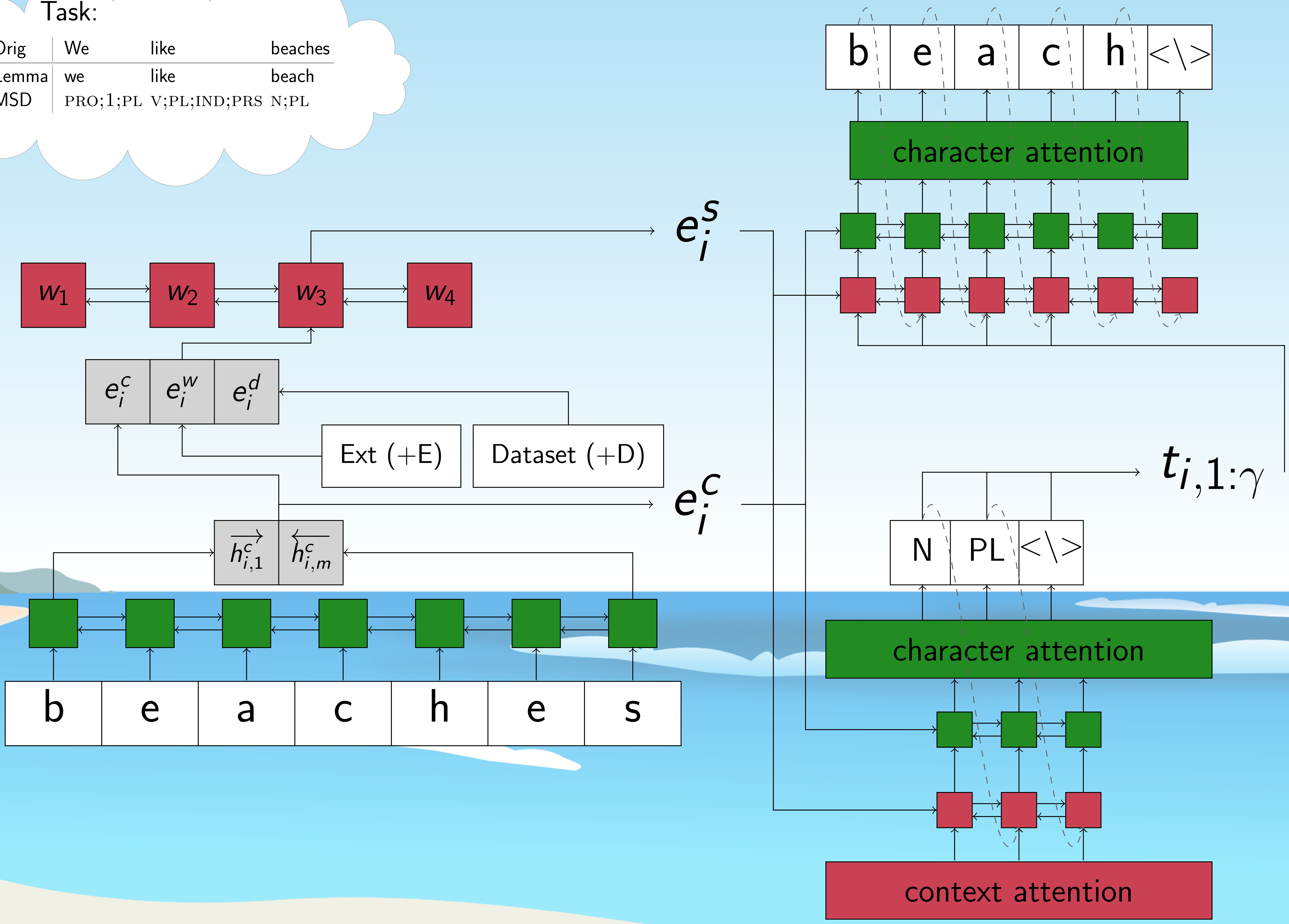
Gertjan van Noord, Gosse Bouma

Multi-Team: A Multi-attention, Multi-decoder Approach to Morphological Analysis

university of groningen

Task:

Orig	We	like	beaches
Lemma	we	like	beach
MSD	PRO;1;PL	V;PL;IND;PRS	N;PL

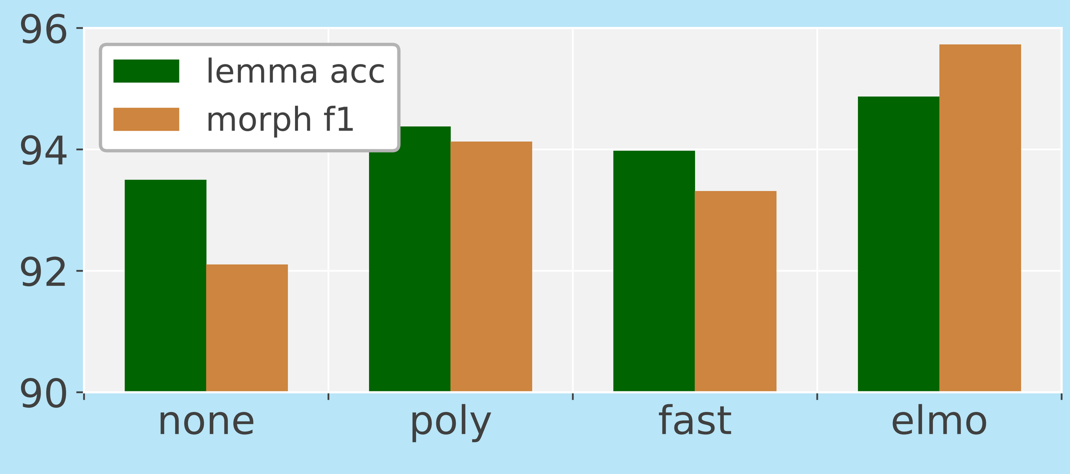


Analysis

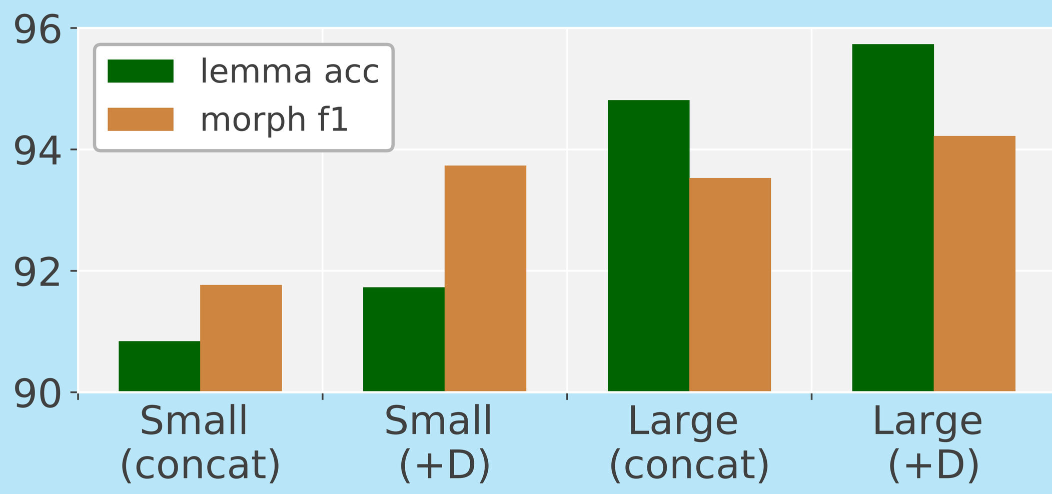
Overview of the datasets used for experimenting.

Dataset	Language Family	Sents	words	tag/word
en_ewt	IE,Germanic	13,297	204,857	1.95
en_pud	IE,Germanic	800	16,927	1.88
tr_imst	Turkic,Southwestern	4,508	46,417	3.58
tr_pud	Turkic,Southwestern	800	13,380	2.78
zh_cfl	Sino-Tibetan	360	5,688	1.00
zh_gsd	Sino-Tibetan	3,997	98,734	1.06
fi_pud	Uralic,Finnic	800	12,556	2.97
fi_ftb	Uralic,Finnic	14,978	127,536	3.07

Comparison of the different types of embeddings to generate word representations.

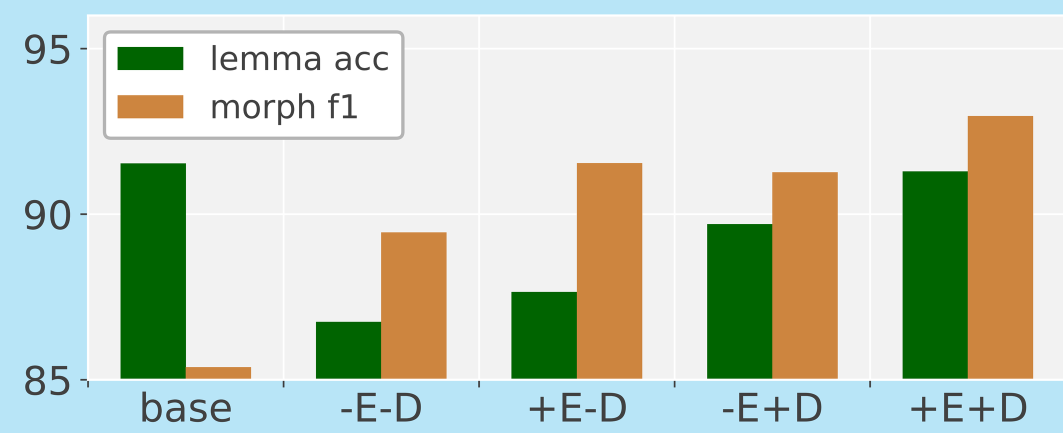


Comparison between the dataset embedding method and the naive approach to combine datasets for multilingual training.



Results

Results on the development data:



Results on the test data:

	Morph. tags		Lemma	
Models	Acc	F1	Acc	Lev
base*	73.16	87.92	94.17	0.13
-E	89.00	93.35	93.05	0.16
+E	90.61	94.57	93.94	0.15

* Baseline from Malaviya et al.(2019)

Conclusions

- Employing a multi-task achitecture having multiple levels of attention mechanism improved the morphological tagging over the baseline strategy.
- Pre-trained embeddings improved our scores for both tasks.
- A dataset embedding strategy also improved our scores, specifically for small datasets.
- Furthermore, these improvements are highly complementary: using dataset embeddings simultaneously with external embeddings leads to superior performance.

Source code

https://bitbucket.org/ahmetustunn/morphology_in_context/