

Explicitly Capturing Relations between Entity Mentions via Graph Neural Networks for Domain-specific Named Entity Recognition

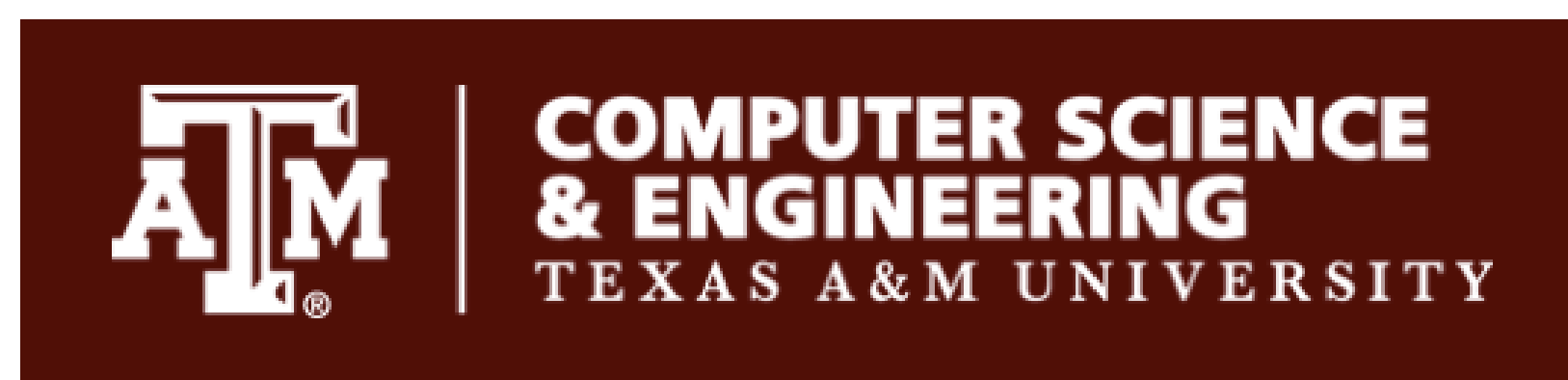
Pei Chen¹, Haibo Ding², Jun Araki², Ruihong Huang¹

¹ Department of Computer Science and Engineering, Texas A&M University

² Bosch Research North America

{chenpei, huangrh}@tamu.edu

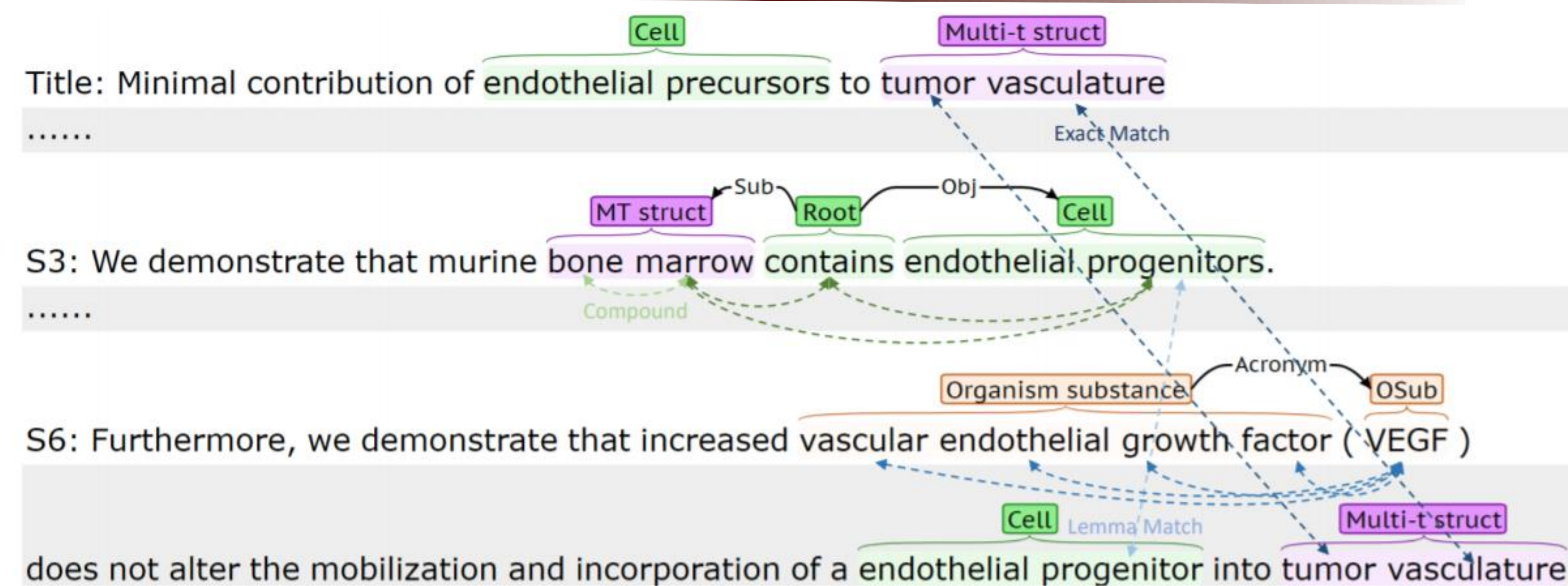
{Haibo.Ding, Jun.Araki}@us.bosch.com



Introduction

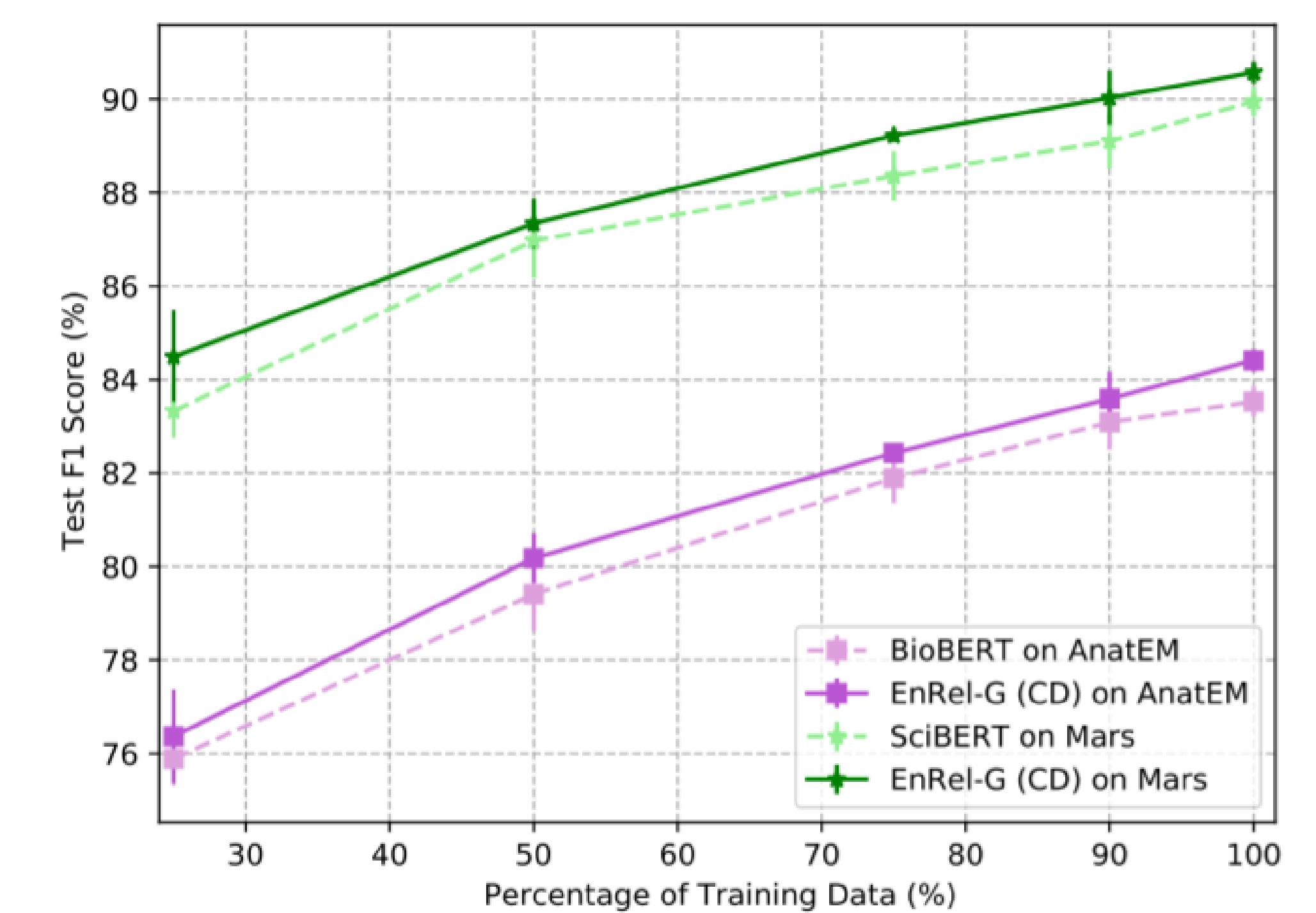
- The Named entity recognition (NER) performance is still moderate for **specialized domains** that have complicated contexts and jargonistic entity types.
- We hypothesize that the interactions of the related entity mentions (both **document-level coreference** and **sentence-level dependency**) will lead to better performance.
- Code for the system: <https://github.com/brickee/EnRel-G>

Relations between Entity Mentions



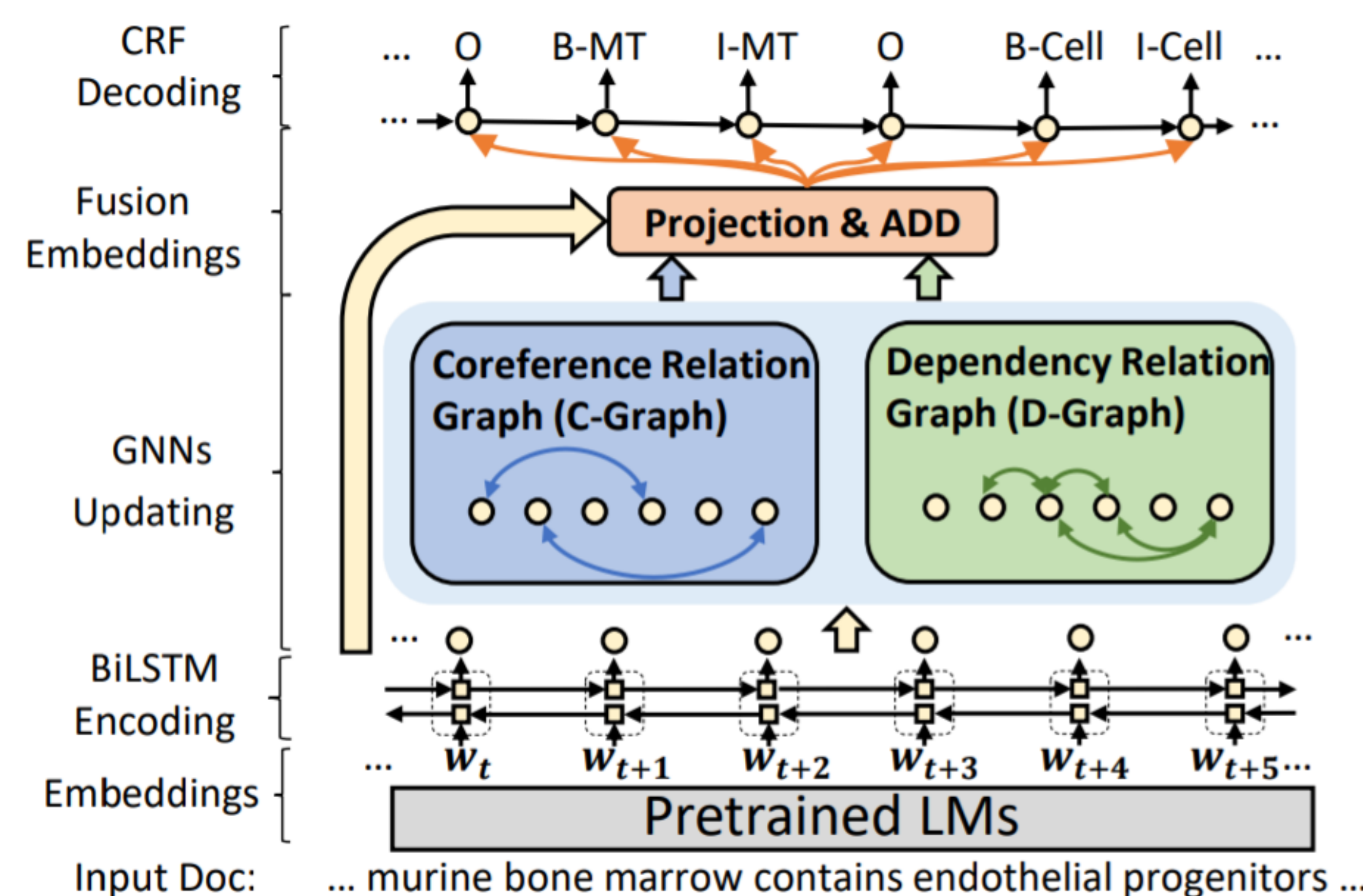
- The **Coreference Relation Graph** at document-level
- The **Dependency Relation Graph** at sentence-level

Learning Curves



System is effective when only a tiny amount of labeled data is available

Overall Architecture



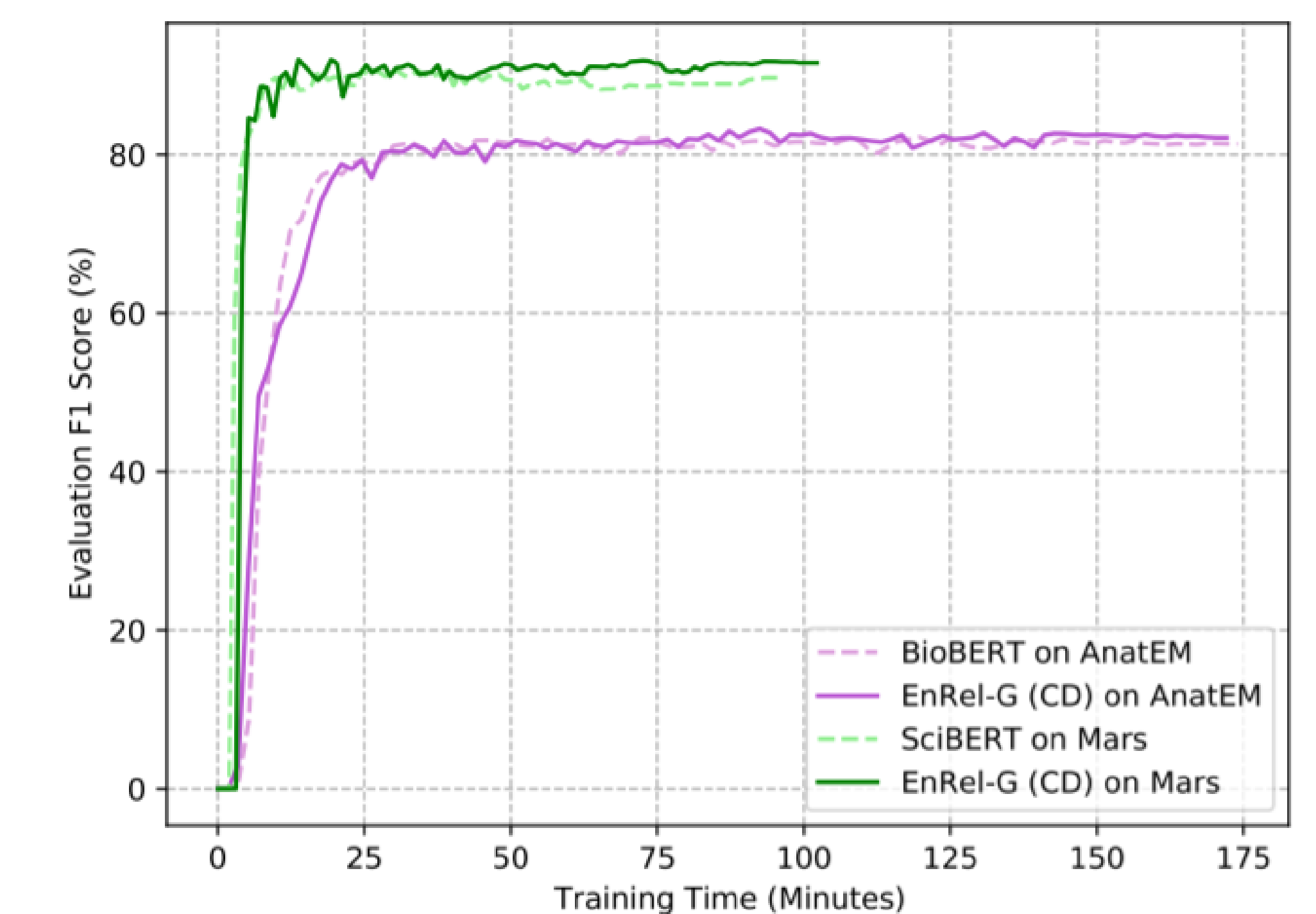
Use Graph Neural Networks to incorporate the relations between entity mentions.

Results

Methods	Datasets		
	AnatEM	Mars	
Wagstaff et al. (2018)	-	94.5 / 77.7 / 85.3	
NCRF++	83.40±0.34 / 76.96±0.46 / 80.05±0.12	91.28±1.08 / 80.57±0.55 / 85.59±0.23	
FLAIR	81.07±0.29 / 75.28±0.57 / 78.06±0.39	90.67±1.02 / 81.45±1.41 / 85.81±0.62	
Pooled FLAIR	82.11±0.50 / 77.55±0.40 / 79.76±0.34	87.79±1.31 / 86.57±1.10 / 87.17±0.17	
Tuning Bio/SciBERT	83.94±0.40 / 83.12±0.30 / 83.53±0.32	90.93±0.66 / 88.99±1.61 / 89.95±0.64	
EnRel-G (C)	84.65±0.67 / 83.69±0.31 / 84.17±0.41	91.21±1.05 / 89.35 ±1.76 / 90.27±0.45	
EnRel-G (D)	84.98 ±0.83 / 83.50±0.45 / 84.23±0.54	92.66 ±1.16 / 88.03±1.46 / 90.29±0.53	
EnRel-G (CD)	84.86±0.50 / 83.96 ±0.32 / 84.41 ±0.24	92.57±1.00 / 88.65±1.50 / 90.57 ±0.47	

- Our system with both the global entity coreference and local dependency relations performs the best among all the systems.
- Either the coreference or dependency relations can help to improve the NER performance.

Computation Cost



The lightweight approach does not increase the time cost compared to BERT models