

## Introduction

### Problem

- The state-of-the-art NER methods based on LSTM fail to fully exploit GPU parallelism.
- Although a novel NER method based on Iterated Dilated CNNs can accelerate network computing, it tends to ignore the word-order feature and semantic information of the current word.

### Contributions

- Compared with the ID-CNNs-CRF, our method obtains improvements of 5.95%, 7.48% and 7.08% in Precision, Recall, and F1-score, respectively.
- The model we proposed is 22% faster than the Bi-LSTM-CRF.

## Methods(Network)

- Position Embedding is utilized to fuse word-order information.
- ID-CNNs architecture is used to rapidly extract global semantic information.
- Attention mechanism is employed to pay attention to the local context

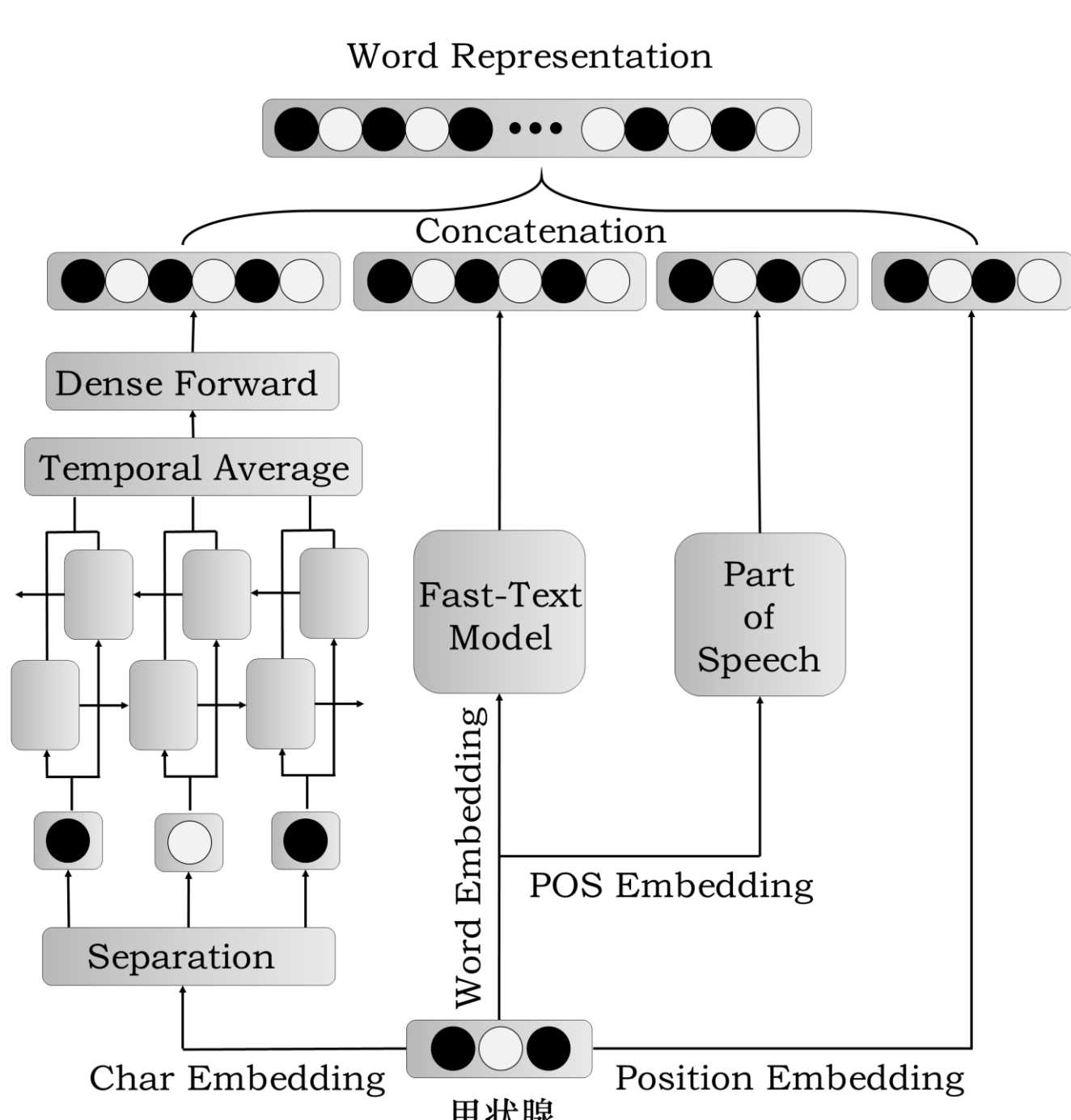


Fig.1

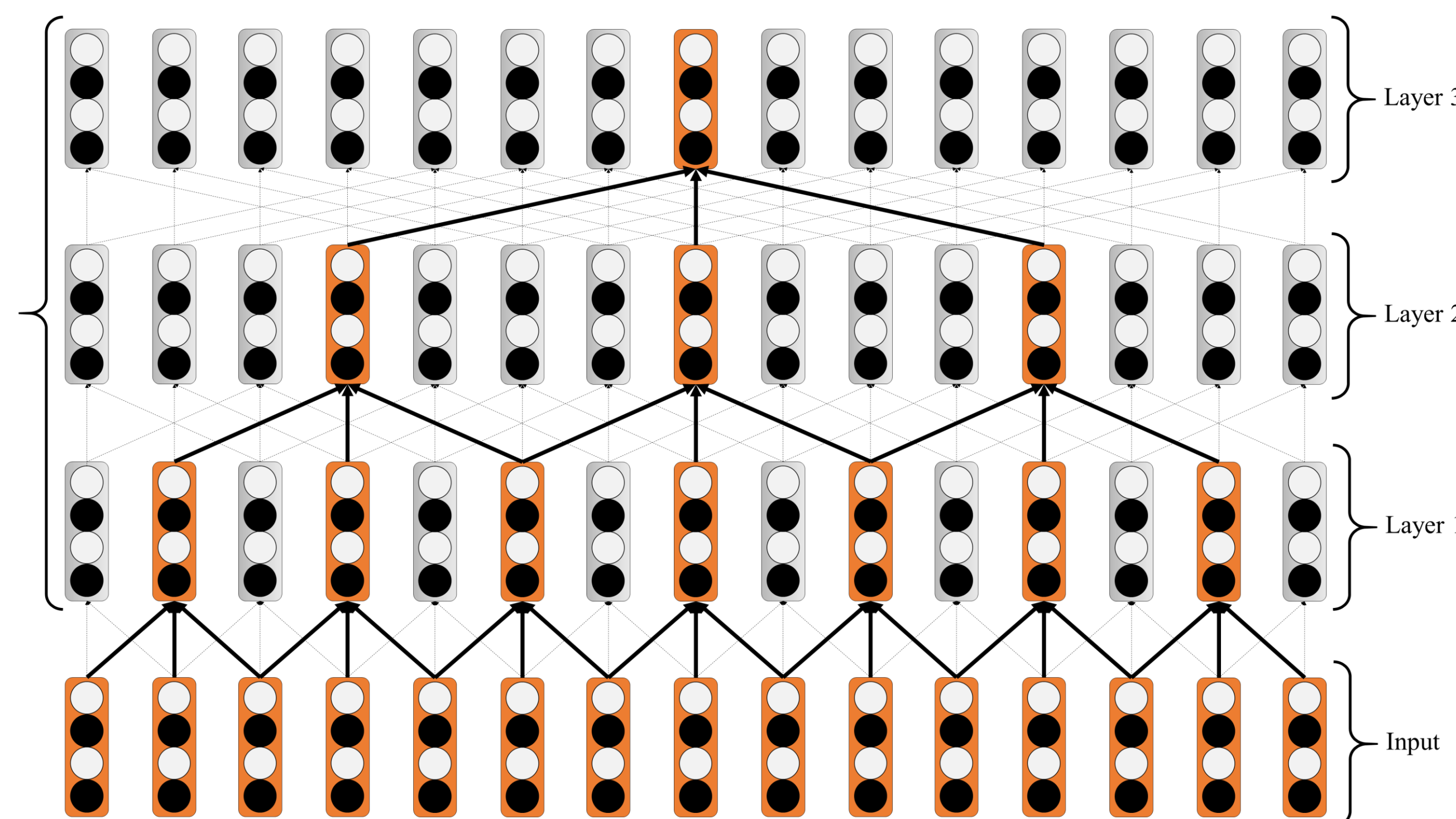


Fig.2

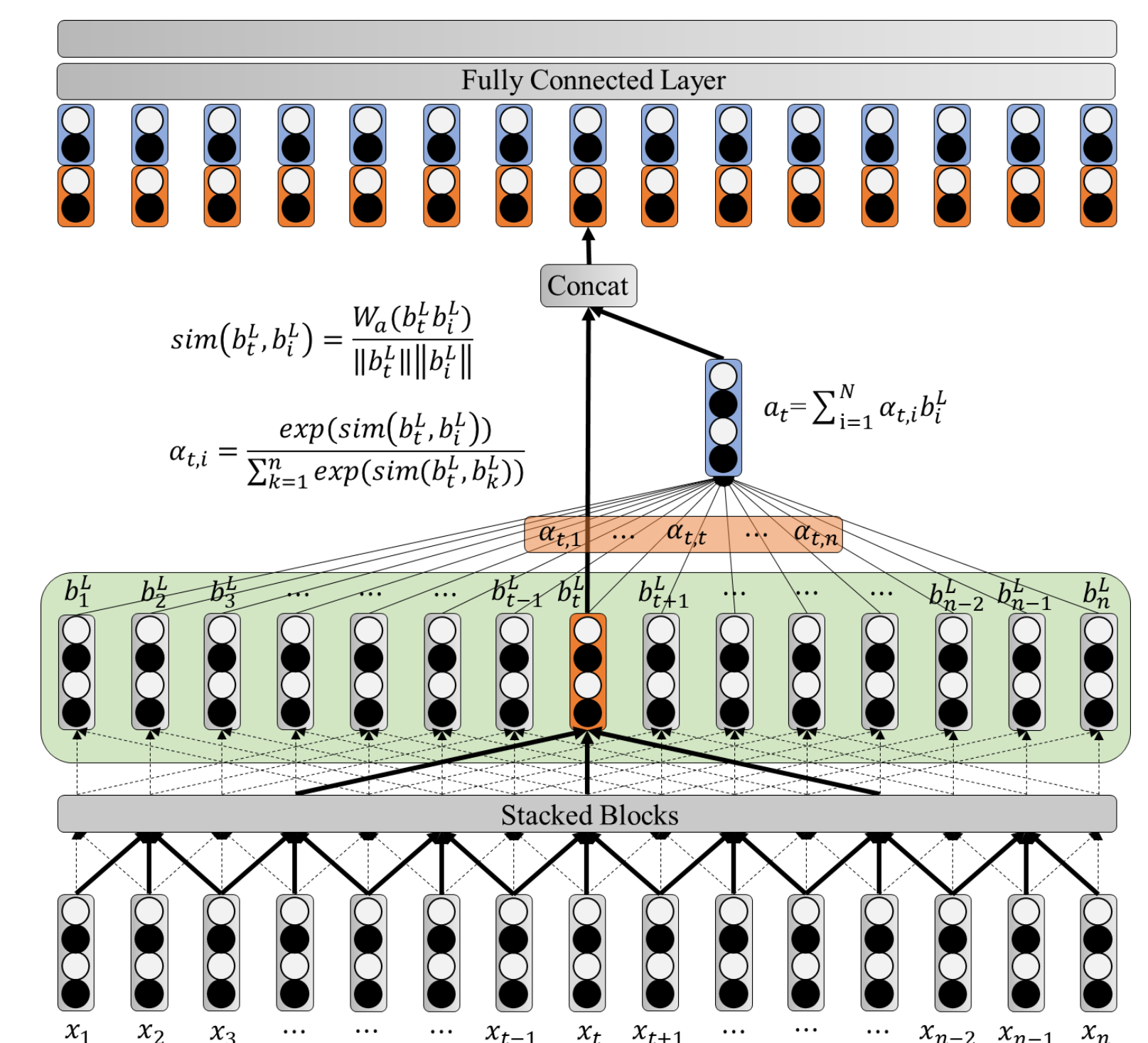


Fig.3

- The extraction module of the word representation.
- A Dilated CNN block. The Input is a sequence of texts, each of which is the word representation of Fig.1.
- Attention-based ID-CNN-CRF architecture. We stack 4 Dilated CNN blocks, each as shown in Fig.2.

## Experimental Results

- Our attention-based ID-CNNs-CRF outperforms the prior methods for most metrics, as shown in Table 3.

Table 3. Results of different NER models on CCKS2017 and CCKS2018 data sets.

Model	CCKS2017			CCKS2018		
	Prec.(%)	Rec.(%)	F(%)	Prec.(%)	Rec.(%)	F(%)
CRF	89.18	81.60	85.11	92.67	72.10	77.58
LSTM-CRF	87.73	87.00	87.24	82.61	81.70	82.08
Bi-LSTM-CRF	94.73	93.29	93.97	90.43	90.49	90.44
ID-CNNs-CRF	88.20	87.15	87.47	81.27	81.42	81.34
Attention-based ID-CNNs-CRF (Ours)	94.15	94.63	94.55	91.11	91.25	91.17

- The model we proposed is 22% faster than the Bi-LSTM-CRF in the test time.

Table 4. Comparison of test time.

model (512 test data)	time(s)	speed
Bi-LSTM-CRF	15.62	1.0×
ID-CNNs-CRF	11.96	1.31×
Attention-ID-CNNs-CRF	12.81	1.22×

- For most categories, the model we proposed is significantly better than the Bi-LSTM-CRF. For example, it has a slightly high F1-score (2.59% on Body and 5.66% on Description in CCKS2018) than the Bi-LSTM-CRF model.

Table 5. Comparison of entity category on CCKS2017 and CCKS2018 data sets.

Model	CCKS2017					
	Body	Check	Disease	Signs	Treatment	Avg
Bi-LSTM-CRF	94.81	96.40	87.79	95.82	88.10	93.97
ID-CNNs-CRF	90.89	88.28	79.67	91.04	81.27	87.47
Attention-ID-CNNs-CRF	95.38	97.79	86.55	96.91	87.64	94.55

Model	CCKS2018					
	Body	Symptom	Operation	Drug	Description	Avg
Bi-LSTM-CRF	92.59	93.12	87.43	82.86	86.33	90.44
ID-CNNs-CRF	84.37	85.81	76.59	80.42	84.61	81.34
Attention-ID-CNNs-CRF	95.18	94.47	85.86	80.06	91.99	91.17