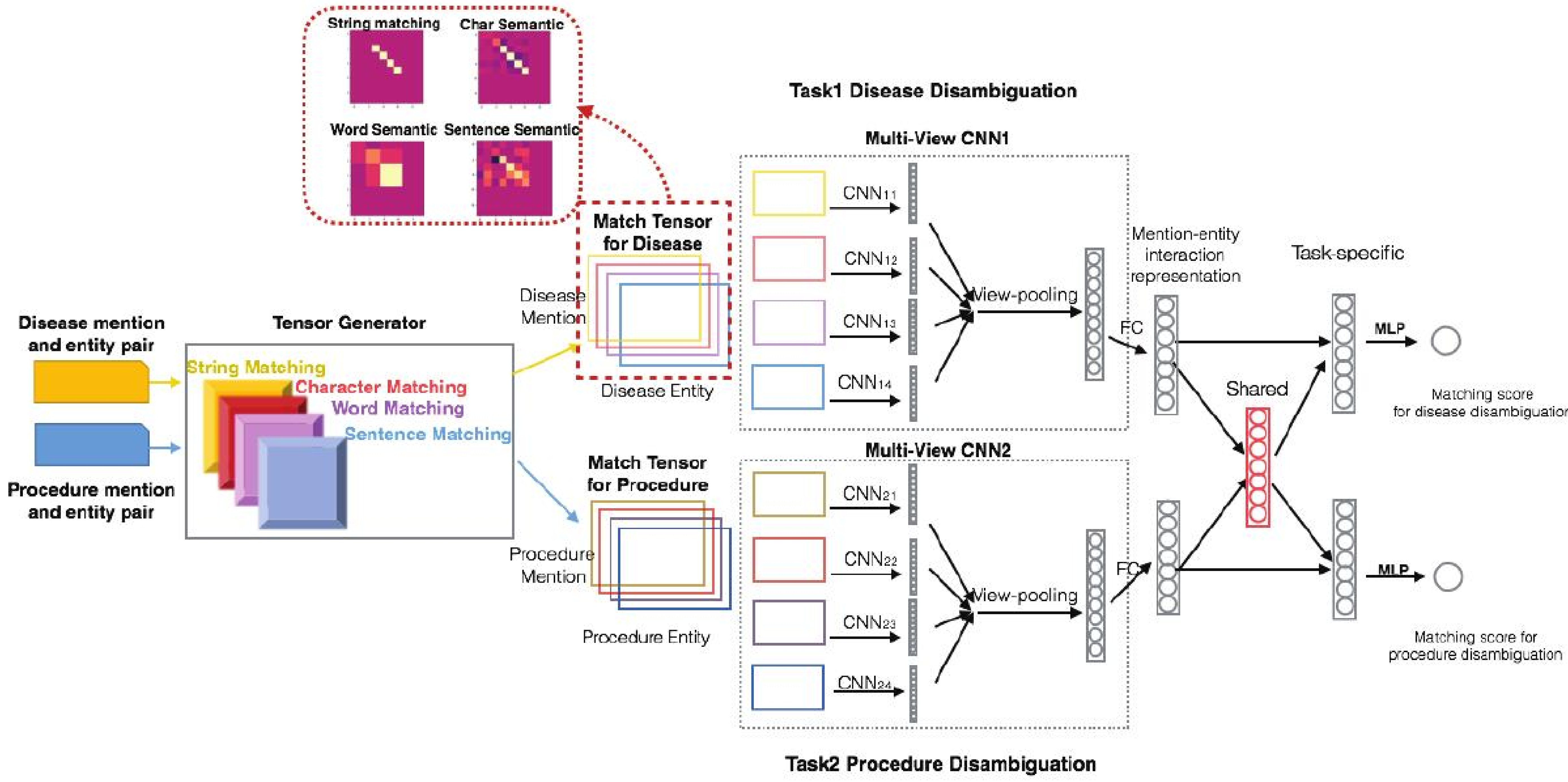

Figure 2: Our proposed architecture for multi-task medical concept normalization. Two tasks have their own matching tensor and multi-view CNN to extract features. Shared structure is used to reinforce the normalization result of both tasks.

ID	Chinese Words	Translation	ID	Chinese Words	Translation
C1a	左肱骨上端骨折	Fracture of humerus upper-end	C6a	间质性肾炎	Interstice nephritis
C1b	肱骨上端骨折	Fracture of upper-end humerus	C6b	间质性肾炎	Interstitial nephritis
C2a	腿	Leg	C7a	慢性乳腺炎	Chronic mastitis
C2b	下肢	Lower limbs	C7b	急性乳腺炎	Acute mastitis
C3a	胃部分切除术	Partial removal of stomach	C7c	乳腺炎	Mastitis
C3b	恶性淋巴瘤	Malignant lymphoma	C1a-1	肱骨	Humerus
C3c	胃恶性淋巴瘤	Stomach malignant lymphoma	C4a-1	扭伤	Sprain
C4a	腿扭伤后遗症	Sequela of sprain in leg	C4b-1	损伤	Injury
C4b	下肢损伤后遗症	Sequela of injury of lower limb	C4b-2	伤后疼痛	Sequela with pain
C5a	[illegible]	Benign tumor of bone	C5d-1	指骨	Finger
C5b	[illegible]	Excision of lesion of finger	C7a-1	慢性	Chronic
C5c	[illegible]	Benign tumor of hand	C7b-1	急性	Acute
C5d	[illegible]	Benign tumor of finger			

Table 1: Chinese words and translations in our paper

Model Formulation

The main idea of this work is based on introducing a tensor generator, and then embedding the multi-view architecture and the multi-task framework to a deep network. Given two mention-entity pairs for disease and procedure, the tensor generator yields two matching tensors separately. Then for each task, interaction representation vector is produced by multi-view CNN. Finally a matching score for normalization is generated in the multi-task module utilizing both shared information and task-specific features. The proposed framework is shown in Figure 2 and all the Chinese words in this paper are translated in Table 1. The details of the model are shown as follows:

- **Matching Tensor**. To tackle short-text problem, for both tasks, a matching tensor is formulated to model interaction between mention-entity pair from both string and semantic aspects in character, word and sentence levels. Particularly, to incorporate context information and solve word-order problem, Bi-LSTM is utilized to integrate sentence level semantics into character vectors.
- **Multi-view CNN model**. We aim to do semantic matching to address non-standard expression problem. CNN is capable of capturing higher level of meaningful matching patterns such as n-grams when convolving across matching matrix (Pang et al. 2016). In our model, four matrices in matching tensor represent different views of matching patterns rather than channels of a picture, where a single CNN can hardly capture all the information sufficiently. Therefore, we adopt multi-view CNN idea to first extract and then effectively aggregate matching signals from four views with a view-pooling strategy.
- **Multi-task learning framework**. Disease and corresponding procedure name for each patient could provide useful information such as body parts to the two related tasks which single task learning may fail to capture. To gain insights from heterogeneous data sources, we design multi-task architecture with constraints to combine the commonalities and differences between medical names in the clinical record.

Matching tensor

The design of matching tensor aims to enrich short-text comparison into string and semantic matching. It resembles human judgement when matching text pairs. Intuitively string matching relying on morphological features is the first to consider. Besides, in Chinese, the meaning of a word is correlated with its composing characters and for unknown token, we may even infer its meaning from the meanings of its