PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks

Wenwen Yu[*1], Ning Lu[*2], Xianbiao Qi[2], Ping Gong[1], Rong Xiao[2]

[1] Xuzhou Medical University, Xuzhou, China
[2] Ping An Property & Casualty Insurance Company, Shenzhen, China

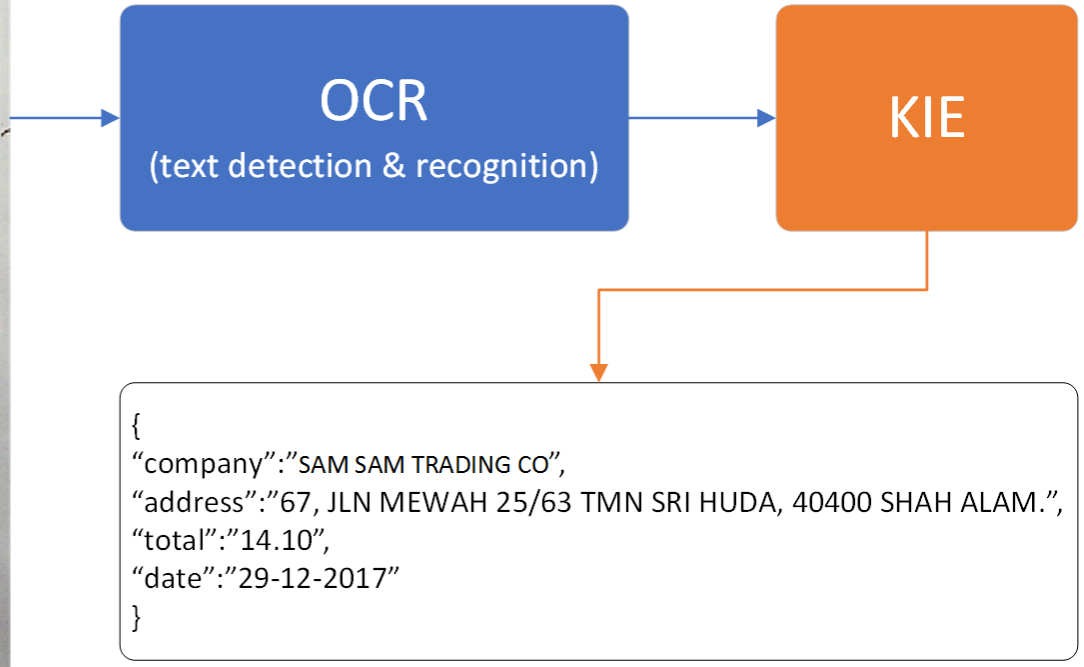
13 Jan 2021

Outline

- Task Definition
- Motivation
- Method
- Results
- Conclusion

Task Definition

- **Key Information Extraction (KIE)** from documents is the downstream task of OCR
- The **aim** of KIE is to extract a number of key fields from the given documents, and save the texts to structured documents.
- KIE is **essential** for a wide range of technologies such as efficient archiving, fast indexing, document analysis and so on.



Wenwen Yu, et al., PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks. *International conference on Pattern Recognition, 2020.*

Outline

- Problem Definition
- Motivation
- Method
- Results
- Conclusion

Motivation

- KIE is a **challenge** task because documents not only have **textual features** extracting from OCR systems, but also have **semantic visual features** that are not fully exploited, and it play a critical role in KIE.
- **Too little work** has been devoted to efficiently make full use of both textual and visual features of the documents.

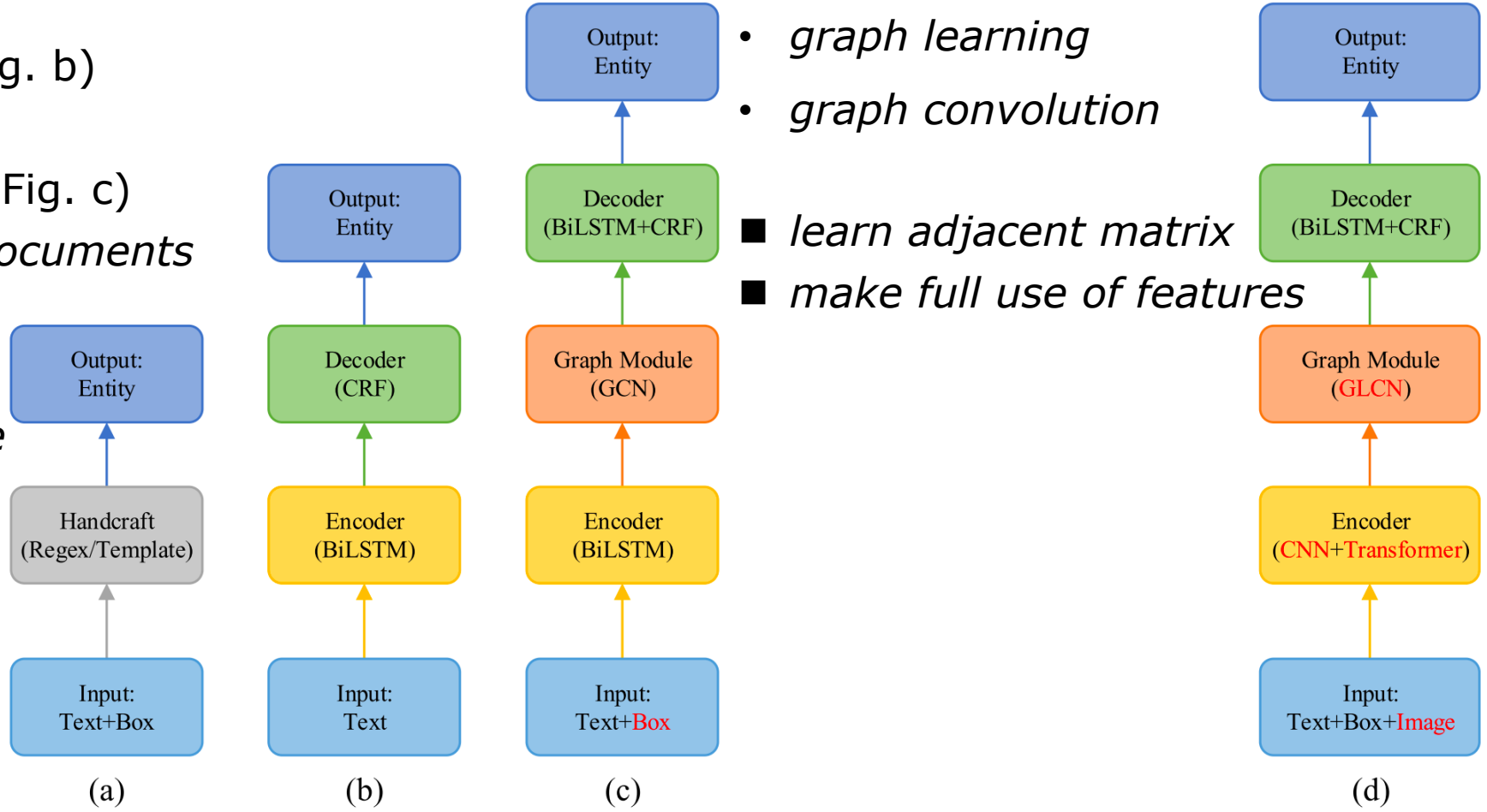
(a) Medical invoice (b) Train ticket (c) Tax receipt

Examples of documents with different layouts and types.

Wenwen Yu, et al., PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks. *International conference on Pattern Recognition, 2020.*

Motivation

- Hand-craft based methods (Fig. a)
- NER based methods (Fig. b)
- Graph based methods (Fig. c)
  predefine a graph on documents
- *only use text and box*
- *task-specific knowledge*
- *human-designed rules*



- **Our method** (Fig. d)
  - *graph learning*
  - *graph convolution*
- *learn adjacent matrix*
- *make full use of features*

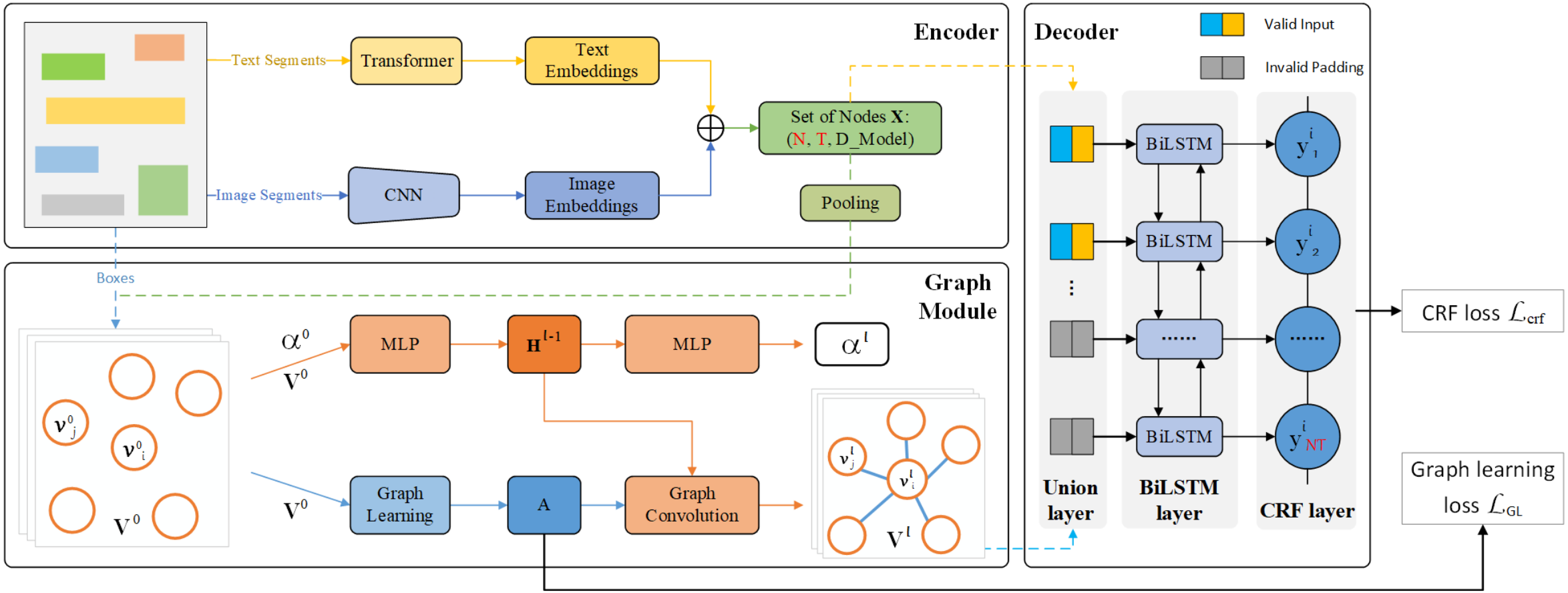
Wenwen Yu, et al., PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks. *International conference on Pattern Recognition, 2020.*

Outline

- Problem Definition
- Motivation
- Method
- Results
- Conclusion

Method

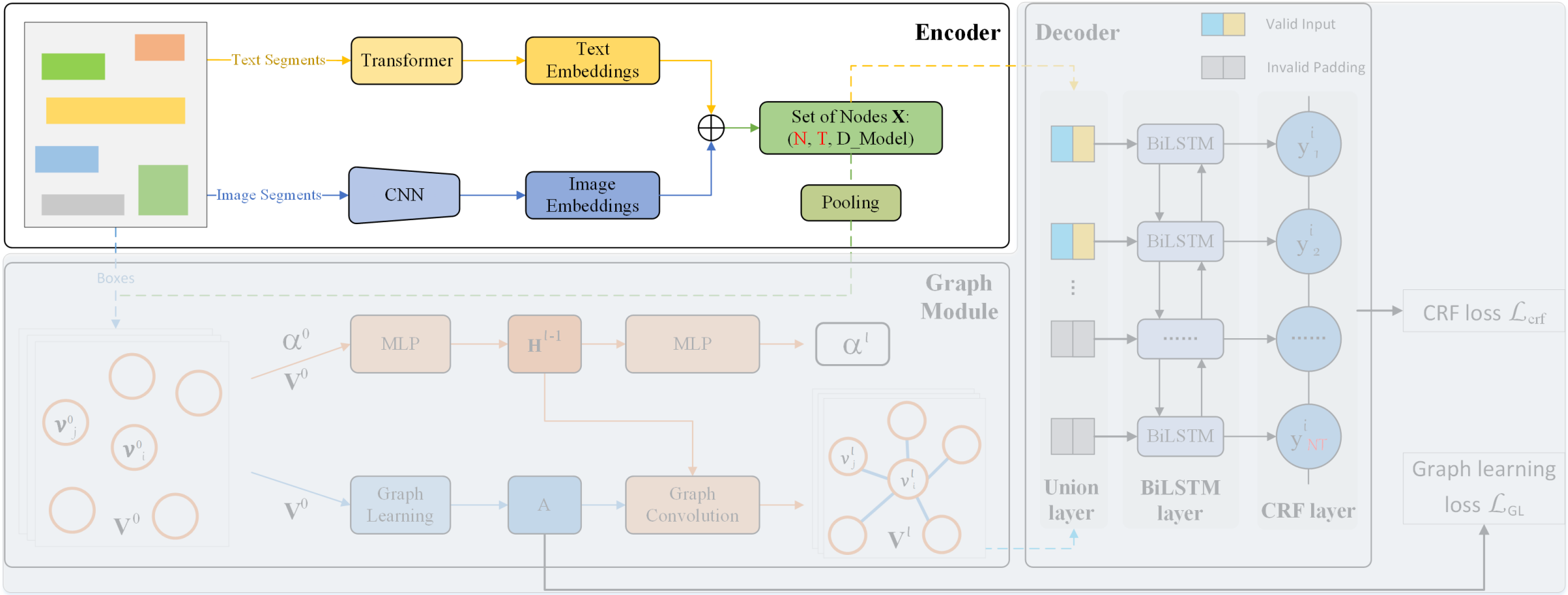
- The overall architecture contains 3 modules: Encoder, Graph Module, Decoder



Wenwen Yu, et al., PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks. *International conference on Pattern Recognition, 2020.*

Method

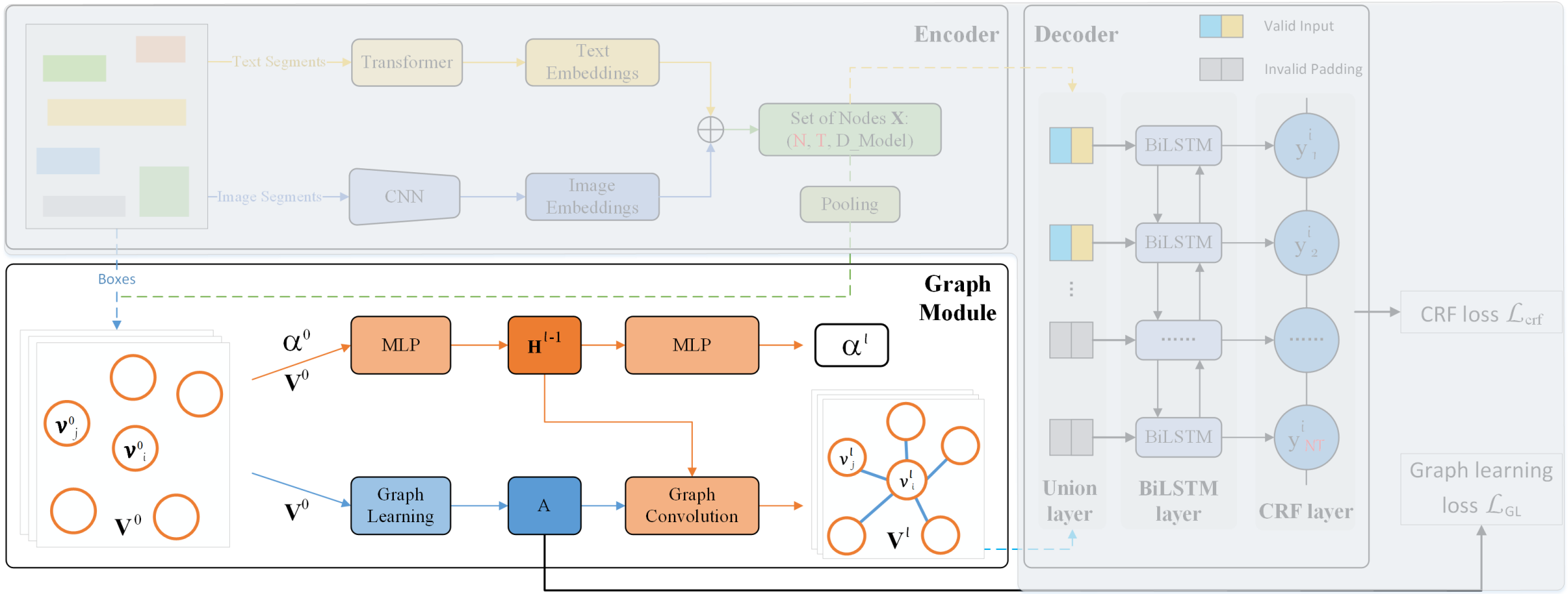
■ Encoder



Wenwen Yu, et al., PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks. *International conference on Pattern Recognition, 2020*.

Method

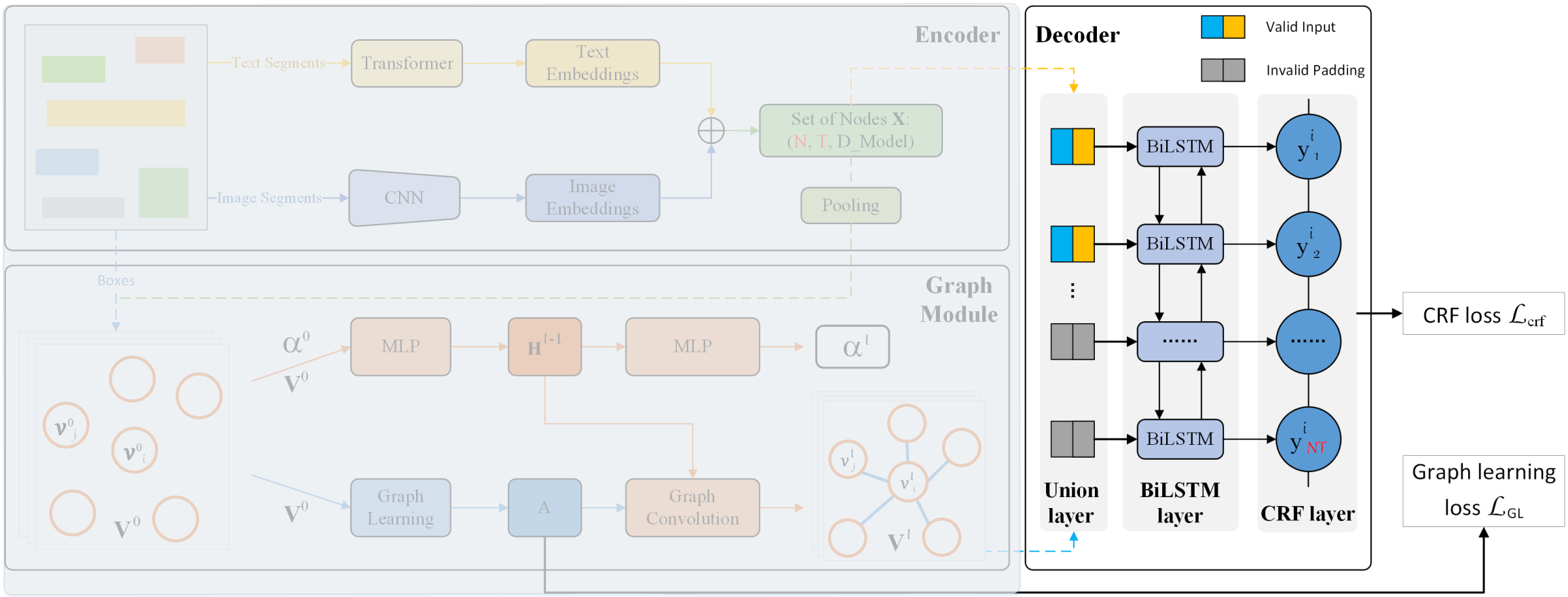
■ Graph Module



Wenwen Yu, et al., PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks. *International conference on Pattern Recognition, 2020.*

Method

■ Decoder



Wenwen Yu, et al., PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks. *International conference on Pattern Recognition, 2020.*

Outline

- Problem Definition
- Motivation
- Method
- Results
- Conclusion

Results

Table I
PERFORMANCE COMPARISON BETWEEN PICK (OURS) AND BASELINE METHOD ON MEDICAL INVOICE DATASETS. PICK IS MORE ACCURATE THAN THE BASELINE METHOD. **BOLD** REPRESENT THE BEST PERFORMANCE.

Entities	Baseline			PICK (Our)		
	mEP	**mER**	**mEF**	**mEP**	**mER**	**mEF**
Medical Insurance Type	66.8	77.1	71.6	85.0	81.1	**83.0**
Chinese Capital Total Amount	85.7	88.9	87.3	93.1	98.4	**95.6**
Invoice Number	61.1	57.7	59.3	93.9	90.9	**92.4**
Social Security Number	53.4	64.6	58.5	71.3	64.6	**67.8**
Name	73.1	73.1	73.1	74.7	85.6	**79.8**
Hospital Name	69.3	74.4	71.8	78.1	89.9	**83.6**
Overall (micro)	71.1	73.4	72.3	85.0	89.2	**87.0**

Table II
RESULTS COMPARISON ON SROIE AND TRAIN TICKET DATASETS. THE EVALUATION METRIC IS MEF.

Method	Train Ticket	SROIE
Baseline	85.4	-
LayoutLM [6]	-	95.2
PICK (Ours)	98.6	96.1

Table III
RESULTS OF EACH COMPONENT OF OUR MODEL. THE EVALUATION METRIC IS MEF.

Model	Medical Invoice	Train Ticket
PICK (Full model)	87.0	98.6
w/o image segments	↓0.9	↓0.4
w/o graph learning	↓1.6	↓0.7

Wenwen Yu, et al., PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks. *International conference on Pattern Recognition, 2020.*

Outline

- Problem Definition
- Motivation
- Method
- Results
- Conclusion

Conclusion

- we study the problem of how to improve KIE ability by automatically making full use of the textual and visual features
- We introduce the improved graph learning module into the model to refine the graph structure on the complex documents given visually rich context.
- This study provides a new perspective on structural information extraction from documents.

https://github.com/wenwenyu/PICK-pytorch

THANKS