End-to-End Hierarchical Relation Extraction for Generic Form Understanding

Tuan Anh Nguyen Dang, Duc Thanh Hoang, Quang Bach Tran,

Chih-Wei Pan, Thanh Dat Nguyen

ICPR 2020



Form understanding





Related works for entity labeling





Methods	Drawbacks
Heuristic methods	 Scalability/generalizability towards large datasets is limited Not generalization Quality of output depends much on the condition of erroneous output from earlier steps (text lines recognition and optical character recognition)
Deep learning-based segmentation *	Hard to model the relation linking between entities
Information extraction with Graph Neural Networks *	 Hard to incorporate both semantic and spatial information Quality of output depends much on the condition of erroneous output from earlier steps (text lines recognition and optical character recognition)

* End-to-End Information Extraction by Character-Level Embedding and Multi-Stage Attentional U-Net

* An Invoice Reading System Using a Graph Convolutional Network

Utilize the strengths of segmentation and graph based methods together





Before	After
Segmentation based methods has limited relational reasoning	Segmentation outputs used to construct input Graph for GNN
Graph is sensitive to misaligned line & merged/splitted error	Entities are re-segmented and constructed spatial relations, which overcomes layout errors.





Over MSAU-PAF architecture















PIF (Part Intensity Field)	• predicts the confidence score, the precise location and the size of the entity.
PAF (Part Association Field)	 predicts associations between entities, which is of a new composite structure, predict a confidence, two vectors to the two entities. refines entity endpoint locations from PIF with regressions, which would resolve joint locations of close-by text-lines (entities) precisely.

Visualizing entities and their associations in PIF/PAF







Before	Now
Limited in localizing entities with segmentation masks.	Part intensity field (PIF) detects and precisely localize entities with confidence maps
No way to represent relations between entities	Part association field (PAF) to associate all detected entities from PIF with composite fields
Accumulation errors due to link prediction is performed after entity detection.	Join supervision with both entity classification and link prediction.





There is no universal solution, so we gather several mechanisms to solve each of these

Corner pooling localizes the keypoint locations better by encoding explicit prior features in horizontally and vertically.



Deal with dense documents by coordinate convolution





- Using of extra coordinate channels, improve convolution's ability against perturbations of translation,
- These geometric information features provide more precisely location of each entities for relational reasoning

Experiment setup



Experiment setup:

Dataset	FUNSD (IBM Form understanding dataset), with tasks: * Node classification * Link prediction Number of samples: (English business forms) 149 train - 50 test
Train/Val split ratio	80/20
Metrics	F1 score, Precision, Recall
Baselines	GraphCNN+Link prediction, MSAU, MLP, Heuristic rules

DATE

DATE

SECTION/85





* https://guillaumejaume.github.io/FUNSD/

Results / Entity labeling



Methods	Metric (F1 score)
MLP (BERT embedding)	0.57
MSAU (one-hot character)	0.73
LayoutLM(base)	0.79
MSAU-PAF (one-hot character) (ours)	0.83



	TEST ARTICLE MORIST
TA . T07281	
SRINGER'S MANET LOTIT	ure
TEST METICIE IDENTIFICA	
tor . toresto	CONCILLION OF SUCHARINE
SILEND VIA:	
	are been Legal
D. OF CONDUMERS	•
CODES HETCHET	NET WEIGHT (if provided)
CONTRACTOR DATE	
CHARLEN INVESTIGN PERSONNEL	esen here refrintrated in darti
1. <u>1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1</u>	
· RECEIVED BY: Anida	and and
STATE SERVICE AND AND	- fav. David
SIGOUVE OF DEPARTMENT O	asian Harde Jane
03/02/90	V.
	성 성상 주변하는 것이야 할?

COM	PETITIVE PRODUCT INTRODUCTION PROGRESS REPORT
10 MRG. K. A. SPAIDO	
DATES STORES	THE OF RACKINGS . Full Firmer Box and Light Box
REPORTING PERIODS	wa see loo wa
	wind by the toth of the following months
TEOT MARKET GEDGRAMM	At of Region 7
PRICE POINT: FULL \$11.8	P/V \$ (Indicate Distributor's Cost Per Cartor)
SALES FORCE INVOLVEMENT	Unertransfering the sep tray of encourse strength and design and labeled
carton fedures in the Carnel section	on. Also olicing mital signs and temporary counter display
and the state of t	
DISTRIBUTORS - ACCEPTAN	NCE/INTRO TERMS/INTRO DEALS
DISTRIBUTORS - ACCEPTAN Product is being introduced to a	INCEANTRO TERMISIANTRO DEALO. D'IVICI Accounts III, d'IM Region, Accognitione in spony at this time.
DISTRIBUTORS - ACCEPTAN Product is being introduced to a	NCE INTRO TERMENTRO DEAL &
DISTRIBUTORS - ACCEPTAN Product is being introduced to al	KCE ANTHIO TOFARIA PARTINO DECAL Convert Accounties in the Region - Acceptance is report, at this time,
DISTRIBUTORS - ACCEPTAN Product is being incoduced to al DISTRIBUTORE INVOLVEMENT Assembly of processor produc	NOT EXAMPLISH THE REAL ADDRESS AND ADDRESS ADDRES ADDRESS ADDRESS ADDRESS ADDRESS ADDRESS ADDRESS ADDRESS ADDRESS ADDRESS ADDRESS ADDR
DISTRIBUTORS - ACCEPTAN Product Is being introduced to al perturbuting payor vessebut describing of procedures produ- occessing \$1016 product	INTERNETIO TEMANETO DEALO Const Actualita de Region - Actualitate la goda, al fila lina E E Interneti su anal d'attempos fond Actualita de Handbold Te-
DISTRIBUTORS - ACCEPTAN Product Is being introduced to al pretributing payor vessebut desiredly of procedure produ- occesing B101F product	-czanto tzanancho con Grand dasana za fagor, kospinsk kojen a ta ta B B skolataristi trada i televijsti televijsti ar sponorski
DISTRIBUTORS - ACCEPTAN Product is being involuted to al 	
BETRIBUTORS - ACCEPTANCE Distribution and the second of a construction and a second of a second Research of the second of protocol assessing BISH product. CHARMS - ACCEPTANCE/ARE Chara acceptions has been vary	
BISTRIBUTORS - ACCEPTANCEME Traduct II berg insoluced to d Distributors back vessesson Research of proceedings (proce Research of the food of CHAINS - ACCEPTANCEME Chain acceptance has been very	
DISTRIBUTORS - ACCEPTANCE ME Probat Is berg incodured in af Internet uncer New Years Automative provide and the Second Contraction and the Second Second Contraction and the Second Second Contraction and the Second Second Contraction and Second Second Second Contraction Second Second Second Second Contraction Second Second Second Second Contraction Second Second Second Second Second Contraction Second	
DistributIONS - ACCEPTAN DistributIONS - ACCEPTAN DistributIONS - ACCEPTANCE AND DistributIONS - DistributIONS - DistributIONS DistributIONS - DistributIONS - ACCEPTANCE AND Online - Conceptance in the Development of the Acceptance in the Development and the Acceptance in the Acceptance in the Acceptance and the Acceptance in the Acceptance	
ративитора – Ассеттан Палон I, вен и полосо и органито восс учена Палон I, вен и полосо учена Палон и полосо и очена Палон Сосруганска Она Ассеттанска И сосруга в Области и Сосруга И сосруга в Палон и Осласти И сосруга в Палон и Осласти	

Results / Entity linking



Methods	Metric (F1-score)
MLP (BERT embedding + Positional feat) (Ground-truth entities class)	0.04
GraphCNN with Link prediction (Ground-truth entities class)	0.13
Heuristic (with MSAU-PAF middle output)	0.64
Heuristic (with Ground-truth entities class)	0.80
MSAU-PAF (Joint classification and link prediction) (ours)	0.75

ر من	Symple No. <u>1194-90</u> Symple No. <u>1194-90</u> Symple of Cigaretty 100 mm Filter
COURSES - MAYERICK EPECIALS - PROBRESS REPORT	Criginal Propost Made De J.H. Beller 2/15/90
(acoutin)	- Rumpass-of-Gample Digarette Modification B-451 Sample Specifications Written By C.W. Lassiter
COLLY IF PARTINE, REGION CONTINUE INCOMES INCOMES	BLEED CASING PRALING PINAL PLAVOR MENTHOL
01/60%: 014 01000	Cigarettes Cilters
ground have encode have / M28	Malazi DiK. 8 Section A Oction B
BRITHMETON BRITIAND CANADA AND CA	Filter Length 37.mm Circumference 4.8 mm
100 07 100 100 00 07 100 000 000 00000000	Maight @8.0g/100 Rod Length @88 mm Paper @81-01-07 @Processure-Droop @20 mm
Triplanos Baddoutlan	• Min. Poper 6-04-07 • Circumference 4.45 mm

Results





16



Model	Entity-Labeling (F1-Score)	Entity-Linking (F1-Score)
MSAU_PAF	0.80	0.72
MSAU_PAF + CoordConv	0.82	0.74
MSAU_PAF + Corner Pooling	0.81	0.73
MSAU_PAF + Corner Pooling + CoordConv	0.83	0.75

Conclusion



- We have demonstrated:
 - A new problem formulation for general form understanding tasks, which can bring the possibility of general solution without any retrain.
 - Various improvements of MSAU-PAF
 - A new model which can solve jointly both entity segmentation and link prediction.
- Future works
 - Extend MSAU-PAF to other document structure parsing problems, i.e table structure recognition, document de-warping.
 - Using the output as part of interpretable Key-Value (to comply with the formulation of visual question answering).