

Table Detection in Invoice Documents by Graph Neural Networks

Pau Riba, Anjan Dutta, Lutz Goldmann, Alicia Fornés, Oriol Ramos, Josep Lladós

Computer Vision Center, omni:us

ICDAR, Sydney, Australia, 23rd September, 2019





Table Detection Framework

Experimental Validation

Conclusion

Outline

Introduction

Table Detection Framework

Graph Representation Network Architecture

Experimental Validation

Datasets and Statistics Node/Edge Classification Table Detection

Conclusions and Future Work





Table Detection Framework

Experimental Validation

Conclusion

Introduction Business Documents

- ► Information extraction: Finance, insurance, manufacturing...
- Manual extraction: Tedious and time consuming.
- ► Automatic extraction: Reduced time and improved quality.



Table Detection Framework

Experimental Validation

Conclusion

Introduction Semi-Structured Documents

- **Structured Documents**: Existing methods, high accuracy.
- Unstructured Documents: Human assistance and validation.
- Semi-structured Documents:
 - Without a fixed spatial layout.
 - Sharing a common set of components.

Marca Later (201), Bar (2010) Statistical Control (2010)	ALL HOME AND ALL AND		
---	---	--	--



Table Detection Framework

Experimental Validation

Conclusion

Introduction Invoice Documents

- Semi-structured documents with flexible layouts
- Spatial arrangement roughly perceived as a tabular layout
- ► Tables are commonly used to condense information

ARA Horas Carlos Carlo	



Table Detection Framework

Experimental Validation

Conclusion

Introduction Constraint

Industrial collaboration - Anonymized data





Introduction

Objective

Table Detection Framework

Experimental Validation

Conclusion

• Graph based representation: Exploit repetitive patterns





Table Detection Framework

Experimental Validation

Conclusion

Introduction Objective

- ► Graph based representation: Exploit repetitive patterns
- ▶ Classification: GNN classification for nodes and edges





Table Detection Framework

Experimental Validation

Conclusion

Introduction Objective

- ► Graph based representation: Exploit repetitive patterns
- Classification: GNN classification for nodes and edges
- Table detection: Group nodes into table regions







Experimental Validation

Conclusion

Graph Representation

- Commercial OCR (by the industrial partner)
- Textual attributes (numeric, alphabet or symbol)
- Visibility graph:
 - Nodes: Document regions
 - Edges: Visibility relations (vertical and horizontal)

TANJUT CERCAN IN VOICE	PhOE 1		
P.Del Bon P.Del Bon	no. Testini Testini Testini Testini		
MARDAT LOOKS CONDUCTION			
WORK LOUGLAND INC. ITY.			
BOOK ITTLE CRAMITLE	HAT AT A STATE		
I REAL PROPERTY DATE SUSAN			
THE COLOR	210-010-00 210-000-00	25.11	
and the set	11.499.00- 11.499.00- 11.499.00-		==
Income service of the service	10+100+10 92+100-22		
Net als 1-19 satisfate set satisfate set satisfate set satisfate set satisfate set	10.000.00 10.000.00 10.000.00	<u>=</u>	=:
LIBURE APPENDER (* 1943) OME APAG POINT LIDUM	101-00-00 101-000-00		
WHAT HER 2-28 WIGHTLATED RATE WIGHTLATED RATE WIGHTLATED RATE	10,110,000- 10,121,000- 10,121,000-	<u></u>	
	12204413		



Experimental Validation

Conclusion

Graph Representation

- Commercial OCR (by the industrial partner)
- Textual attributes (numeric, alphabet or symbol)
- Visibility graph:
 - Nodes: Document regions
 - Edges: Visibility relations (vertical and horizontal)

PARTY LENGTE IN V.O.	105	96 J	
anticasti vicia anticas alla titata anticatata one sun statut	A Del son 10 A DEL		
MARINE LOSAS CONFERNIALIAN			
gradion constant int. Ille			
BRIER THE CREATERS		CONTRACTOR OF T	
rem description	100.041	1811 18800001	
3 Pupuli misatun-sour parei suruni	• • • • • •		
EXTERN OF FURNER 12, 1945 THE ATTRACT FURN COLDS	de contrat	200100000	
PRAL RAX 1-24 MODULATIO ANTI ADDACT COMINSTIN	The unit	111100-02	
E Publies Read and a second cares devision			
Discuss spectrum, d. 1993 Discuss and visition Powe Course	20+140+29	92,486,00	
Paul Auto 1-10	10.000-000	10.000.000	
ABACTINTED BATT	12+174-00-	22,200,00	
August Classifies			
DESCRIPTION OF THE PART	101-000-00	181-800-00	
PART IN 1-08 PRA BAR 1-08 SOUTIATE ANT SCHOOL CONTINUES	34,110,000 13,137,580	10:00:00	
	1 120	6413	





Table Detection Framework

Experimental Validatio

Conclusion

Graph Neural Network

- ▶ Notation introduced in [2]
- A Graph intrinsic linear operators
- ρ Activation function (ReLU)
- θ Learnable parameters



$$x^{(k+1)} = G_C(x^{(k)}) = \rho\left(\sum_{B \in \mathcal{A}^{(k)}} Bx^{(k)}\theta_B^{(k)}\right)$$

[2] V. Garcia et. al., Few-shot learning with graph neural networks, in ICLR, 2018.



Table Detection Framework

Experimental Validation

Conclusion

Graph Neural Network Graph Adjacency Layer

- Importance of the neighbourhood connection
- MLP MultiLayer Perceptron
- σ Activation function (Sigmoid)
- Absolute difference provides the symmetry property

$$\phi_k(B)_{i,j} = \begin{cases} 0 & \text{if } B_{i,j} = 0\\ \sigma\left(\mathsf{MLP}_{\tilde{\theta}}\left(\left|x_i^{(k)} - x_j^{(k)}\right|\right)\right) & \text{otherwise} \end{cases}$$



Table Detection Framework

GNN Architecture

Experimental Validation

Conclusion





Table Detection Framework

GNN Architecture Graph Residual Block

Experimental Validation

Conclusion

► Idea of ResNet [1]

- GNN layers with a skip connection
- Edge weights are learned at the beginning of the block



[1] K. He et. al., Deep residual learning for image recognition, in CVPR, 2016.



Experimental Validation

Conclusion

GNN Architecture Objective functions

- ▶ Node classifier: Linear classifier with Softmax operation
- Edge classifier: Binary Cross entropy
 - 0 Edge connects two different regions
 - ▶ 1 Edge connects elements in the same region





Table Detection Framework $\circ \circ \circ \circ \circ \circ \bullet$

Experimental Validation

Conclusion

Table Detection

- Discard 0'ed edges
- Subgraphs with nodes classified as Table are considered
- The confidence score of these subgraphs are thresholded for the final decision







Experimental Validation



Table Detection Framework0000000

Experimental Validation

Conclusion

Datasets

CON-ANONYM

- 960 documents
- ▶ 8 region annotation
- Common car invoices
- Not publicly available

RVL-CDIP

- Overall 25,000 images
- ▶ 5 region annotation
- Selected 518 invoice class
- Publicly available ¹

^[3] A. W. Harley et. al., Evaluation of deep convolutional nets for document image classification and retrieval, in ICDAR, 2015.



¹ https://zenodo.org/record/3257319

_

_

Table Detection Framework

Experimental Validation $0 \bullet 0 \circ 0$

Conclusion

 21		C	<u>0</u> 1	гc
a	ιa	5		ιJ

CON-ANONYM	RVL-CDIP
950 (665, 95, 195)	518 (362, 52, 104)
1252	518
1202	485
8	6
245.50	124.03
1354.81	619.55
	CON-ANONYM 950 (665, 95, 195) 1252 1202 8 245.50 1354.81



Table Detection Framework

Experimental Validation

Conclusion

Node/Edge Classification

Task	CON-ANONYM			RVL-CDIP		
rasik	All	Table	Edge	All	Table	Edge
Pow 2 + Edge	82.8 84.2	96.4 97.0		57.8 58.2	80.9 79.1	_ 84.1
Pow 5 + Edge	82.7 84.5	96.2 97.2	_ 93.4	56.5 62.3	82.3 83.9	_ 84.0



Table Detection Framework

Experimental Validation

Conclusion

Table Detection

Intersection over Union

Task	CON-ANONYM			RVL-CDIP			
	F1-Score	Precision	Recall	F1-Score	Precision	Recall	
Pow 2	69.4	65.8	73.4	28.6	23.9	35.4	
+ Edge	70.8	65.2	77.6	30.8	26.7	36.5	
Pow 5	68.4	65.3	71.8	22.6	20.0	26.0	
+ Edge	73.7	78.4	69.5	30.8	25.2	39.6	



Table Detection Framework

Experimental Validation

Conclusion

Table Detection

Proper detection





Table Detection Framework

Experimental Validation

Conclusion

Table Detection

Proper detection





Table Detection Framework

Experimental Validation

Conclusion

Table Detection

- Proper detection
- Preprocessing problems





Experimental Validation

Conclusion

Table Detection

- Proper detection
- Preprocessing problems
- Tabular layout





Conclusions and Future Work



Experimental Validation

Conclusion

Conclusions and Future Work

- ► First *Table Detection* based on structural information.
- ► A Graph models the underlying structure of the document.
- Publicly available RVL-CDIP invoice dataset.
- Deal with anonymized data.

Generalize to unconstrained tabular layout.



Thank you for your attention!

Pau Riba Computer Vision Center priba@cvc.uab.cat http://www.cvc.uab.cat/people/priba/

