An Integrated Approach for Table Detection and Structure Recognition

Hai-Hong Phan, Ngo Dai Duong
Le Quy Don Technical University, Ha Noi, Viet Nam
Correspondence: hongpth@lqdtu.edu.vn; daiduong28789@hotmail.com
Communication: received 4 April, revised 29 May, accepted 31 May
Digital Object Identifier: 10.32913/mic-ict-research.v2021.n1.974

Abstract: Detecting and identifying the table structure is an important issue in document digitization. Although there have been many great strides based on current deep learning techniques, table structure identification is still a difficult and arduous problem, especially when solving the problem of digitizing text in practice. The paper proposes a solution to digitize table documents based on the Cascade R-CNN HRNet network to detect, classify tables and integrate image processing algorithms to improve table data identification results. The proposed algorithm proved effective on real data - the hydrometeorological station record book contains tables including simple and complex structures tables with over 98% accuracy.

Keywords: *Document structure analysis, table detection, detect the table structure, hydrometeorology.*

I. INTRODUCTION

Currently, digital transformation is one of the development goals of each country. The digital transformation brings great benefits to management, economic and social development. In particular, the use of digital documents instead of traditional documents is being rapidly developed. Digitizing tables in documents makes a lot of sense because tables often contain a lot of important and useful information. There are two main tasks of table digitization: table detection and table understanding. Table detection is to identify the table container, assign an identifier, and distinguish it from the text areas in the document. Table structure identification is to identify cells, rows, columns, and hierarchical relationships between cells.

This paper proposes a method to identify the table structure for the real-world problem - digitizing several types of hydrogeological records. The model is classified into tabular forms: old technical books before 2000, new technical books since 2000 and normal tables. In particular, the table usually has a simple structure, the technical book table has a complex structure.

The article's contributions include:

- Proposing to use Cascade mask R-CNN HRNetv2pW40 deep learning network for table detection and table classification, an improvement over using W=32. Achieve up to 98.74% accuracy for the table detection task with the Hydrometeorology set.

- Proposing a total solution to recognize the table structure, using techniques to extract the structure of simple tables. Use templates to handle complex structured tables encountered in practice.

The solution has achieved high accuracy for the task of detecting and classifying tables based on deep learning techniques and image processing on the Hydrometeorology dataset.

II. RELATED WORK

In 1997, P.Pyreddy and, W.B.Croft [1] first proposed a table detection method based on character alignment, holes and spaces.

Naganjaneyulu et al. [2] in 2016 proposed a model for table detection using two Hough algorithms for line detection and Harris for corner detection. The author applied it on many different types of documents and achieved an accuracy of up to 89%. This method fails with tables with uneven spatial delineation and table formatting.

Dhiran et al. [3] proposed a method to detect and extract table information from document images, using OCR for preprocessing, enhancement and contour noise removal. The author uses 1200 document images which are divided into 3 categories: type with line, type with line with column whitespace and type with only white space. This method achieves an accuracy of up to 88.7%. However, this method does not work well with documents that are divided into multiple columns.

Gilani et al. [4], propose a method with two modules of image transformation and table detection. In the article, the author proposes to use the Abbey Cloud OCR SDK and tesseract OCR to display the comparison results.

Shafait et al. [5], propose an approach to detect tables from heterogeneous documents with different table layouts. Suggest your approach to document analysis via tab-stop. Huynh-Van et al. [6], proposed a three-step combined method to detect tablespaces: 1) Zone classification, 2) Table detection based on intersection of vertical and horizontal lines and 3) define a table made up of parallel lines only. The authors use data set UW-III for the experiment to obtain the results. Achieve precision (82%) and recall (80%).

Kleber et al. [7], proposed a method to match the sample table structure using a linked graph for handwritten/printed historical documents. Table matching is performed by detecting the maximum group in the association graph, which represents the matching of the sample and document of interest line information.

Along with the development of deep learning techniques. In 2017, Schreiber et al. [8] proposed a deep learning-based approach for table detection and table structure recognition in documents or images. The author uses Faster R-CNN network for table discovery task and FCN network for table structure recognition task (specifically, table row and column recognition). This method is tested on the ICDAR 2013 dataset consisting of 67 documents with 238 pages, with the F1 measure reaching 96.77% for the table detection task and 91.44% for the table structure recognition task.

Devashish Prasad et al. [9] in 2020 proposed an end-to-end deep learning model CascadeTabNet used to detect and recognize the table structure. The article's tests are on the data sets ICDAR 2013, ICDAR 2019, Tablebank [10]. With the F1 measure achieved 90.1% for the ICDAR 2019 dataset, 94.33% for the Tablebank dataset for the table discovery task, 23.2% for the ICDAR 2019 dataset branch B2 (modern tables) for the identification table structure task.

In object detection, a certain IoU (Intersection over union) threshold is used to determine positive and negative. Usually will be trained with a low IoU threshold eg 0.5, this causes the model to always generate a lot of noise in the detection. However, if the IoU threshold is increased, the detector performance decreases. The underlying cause for this phenomenon is over-fitting in the training process, due to an exponential decline with positive samples as the IoU threshold is increased. The multi-stage object detection architecture, Cascade R-CNN proposed by Zhaowei Cai et al. in 2017 [11] aims to solve these problems. Cascade R-CNN uses several stages (usually 3), each stage resampling, thereby changing the distribution of the input hypothesis. Cascade R-CNN has relatively high accuracy with test benchmark data.

Most of the networks developed used for classification such as Alexnet, VGGNet, GoogleNet, ResNet are usually followed by Lenet-5's rule of reducing the spatial size of feature maps, connecting convolutions from degrees high resolution to low resolution in a sequence and using these low resolution representations for classification. Meanwhile, high-resolution representations are needed for tasks such as semantic segmentation, human action computation, and object detection. Jingdong Wang et al. proposed that HRNet (High-Resolution Net) [12] is capable of maintaining high resolution representations throughout the processing. The HRNetV2p instance is made up of a feature tower from HRNetV2 representatives used for the object detection task and evaluated against the COCO dataset.

These methods, which mostly deal with photos of construction documents for competitions, have not yet processed and proven the method's effectiveness with real-world data. Second, the data is usually of a historical or modern type only. While the data to be digitized in practice can include both. The article focuses on proposing algorithms to recognize table structure on actual data from hydrometeorological stations. This data has a high difficulty because it contains many different table structures, besides there are both old and new documents, so the image quality is much different.

III. PROPOSED METHODS

In this paper, we propose an algorithm intergrating deep learning and traditional to solve the real problem - digitizing tables of hydrometeorological documents. Towards this goal, the main strategy of the paper includes:

- Based on state-of-the-art CNN architecture for table segmentation and detection.

- Tranfer learning for Hydrometeorological data to improve performance.

- Use the template table to better identify the table structure in the document.

With the collected initial data, we divide the Hydrometeorology data into the following 3 types of tables: table_skt1_old, table_skt1_new, table_normal. The table_skt1_old is the old table of the meteorological observation book from 1969 to 1985. The table_skt1_new is the table of the basic meteorological observation book from 2000s. Table_normal is the remaining table, both old and new, including specific tables for each type of documents.

We propose a solution to recognize the table structure as shown in Figure 1. Accordingly, the original document image is included in the Cascade mask RCNN HRNetv2pW40 model to detect the tables in the document. In addition, the model also classifies the table into 3 types, corresponding to 3 output branches: Table_normal, Table_skt1_old and

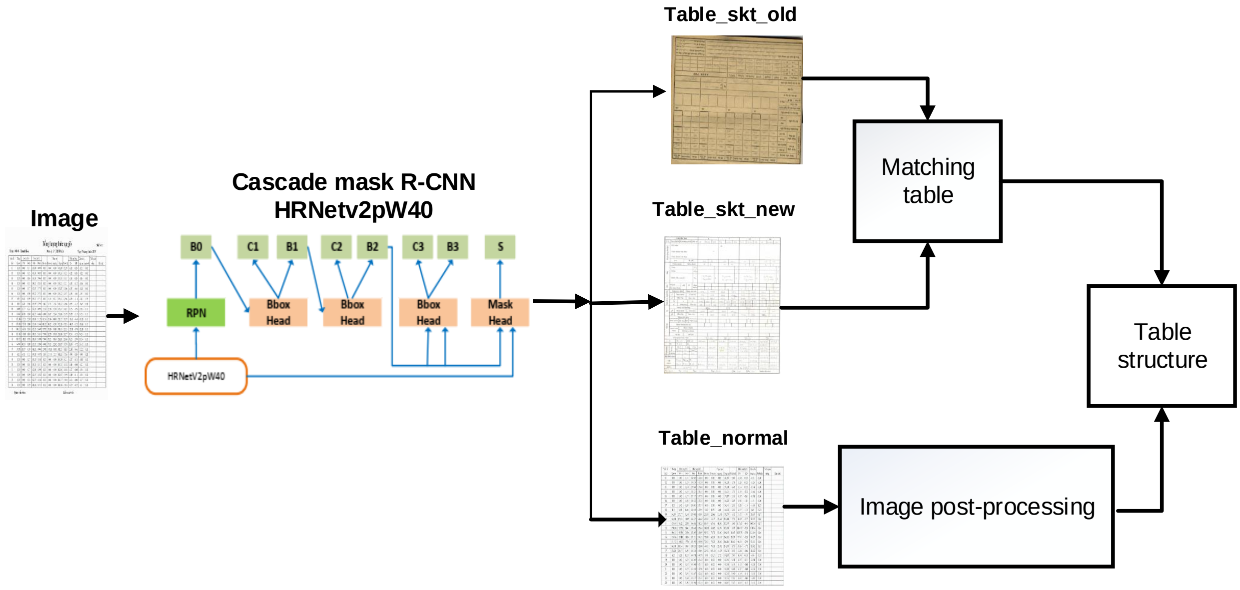


Figure 1: The proposed pipeline for identifying the table structure in the Hydrometeorology documents.

Table_skt1_new. Tables in the table_normal branch would be processed by traditional image processing methods to recognize the cell areas in the table and the relationship between them. Besides, the proposed method applies alignment techniques and template tables to identify the structure of the table for table_skt1_new and table_skt1_old.

1. The proposed table detection model

In this paper, the prososed method intergrates Cascade mask R-CNN and HRNetV2pW40. Cascade mask R-CNN is very similar to Cascade R-CNN, this model extends the ability to segment objects by adding a segmentation branch as done with Mask RCNN.

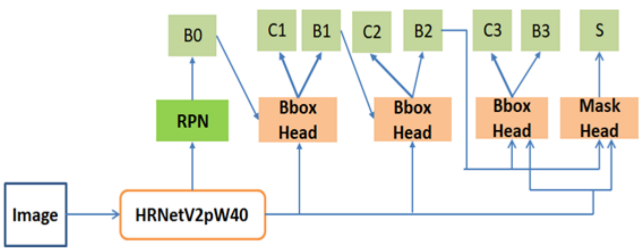


Figure 2: Illustration of using Cascade Mask model R-CNN HRNetV2pW40 for table detection in Hydrometeorological documents

As shown in Figure 2, HR_NetV2p_W40 model extracts feature map of input image I. The Region Proposal Network (RPN) predicts the original proposed features from the feature maps. "Bbox Head" takes RoI features as input and makes predictions for each RoI. Each "Bbox Head" makes two predictions: predicting the bounding box and classifying. We indicate "B" as the bounding boxes predicted by the "head". "C" is the classification prediction of the object. "Mask Head" predicts the mask of the object and "S" denotes the output segment. At the inference step, the object detected by "Bbox Head" is supplemented with the segment mask performed by "Mask Head".

The model bases on the MMdetection library [13]. We use the default cascade mask configuration of rcnn hrnetv2p w32 20e for our model during the experiment and analysis.

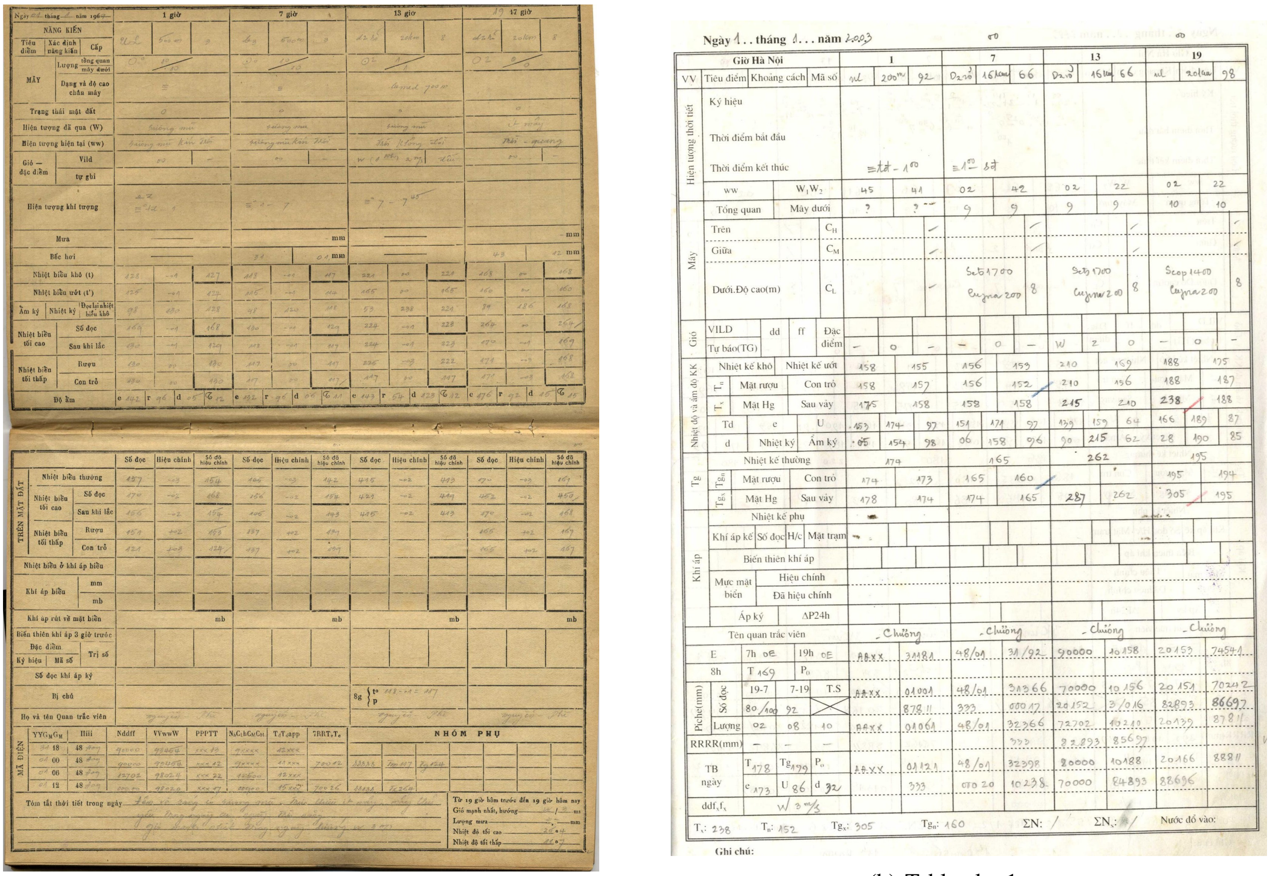
2. Transfer learning

For table detection task, we use iterative transformation learning technique. First, we use the general data containing many different tables, like table with border, table without border, table lacking border. After training, we re-adjust with the hydrographic materials of the table, in order to create a model that can infer well with this special the data set.

Figure 4 illustrates the transformation process for problem release table accountants.

First, we use a model that has been trained with generic data. We only perform the task of detecting tables, all tables in the document will be annotated as a class (table class).

Second, we refine the model trained in the first step. With data obtained from documents of hydrometeorological



(a) Table skt_1_old

(b) Table skt_1_new

Figure 3: Table data labeled table_skt_1_new and table_skt_1_old

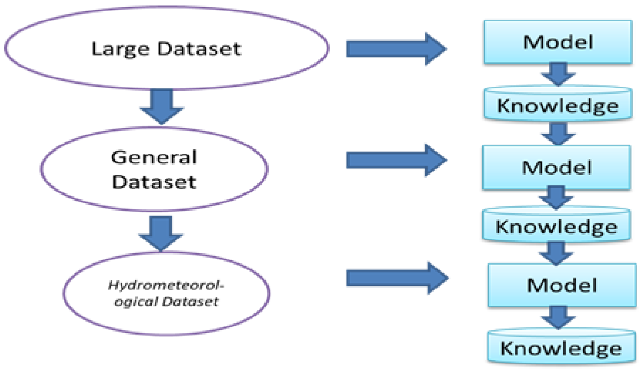


Figure 4: Two stage transfer learning for the Hydrometeorological table detection

centers of Vietnam. In this step, we classify the table into 3 types: table_skt_1_old, table_skt_1_new, table_normal. Table_skt_1_old are old technical books, historical documents (since 1969) and complex structure.

Table_skt_1_new are new, complex tables. Table_normal are simple structured tables that can be processed by traditional or deep learning methods.

3. Analyze table structure based on sample table

For some types of tables with complex structure, historical documents, it is very difficult to automatically recognize the table structure. For special data with known information about some tables, several techniques can be used to identify the structure of the table as shown in Figure 5. First, after the table is detected in the document with the model in the previous figure, the tables table_skt_1_old and table_skt_1_new are truncated from the table container. Second, we use ORB (Binary Robust Independent Elementary Features) technique to extract features of the image, then use RANSAC algorithm to match key points in two images, find the identity matrix and transform the input image according to the identity matrix. We get the output image with a high match to the sample image. Third, we define the table template as xml. Enclose box coordinates around table cells and show their relationship. The cells are then OCR to detect the filled text in the defined table structure.

IV. Results and Analysis

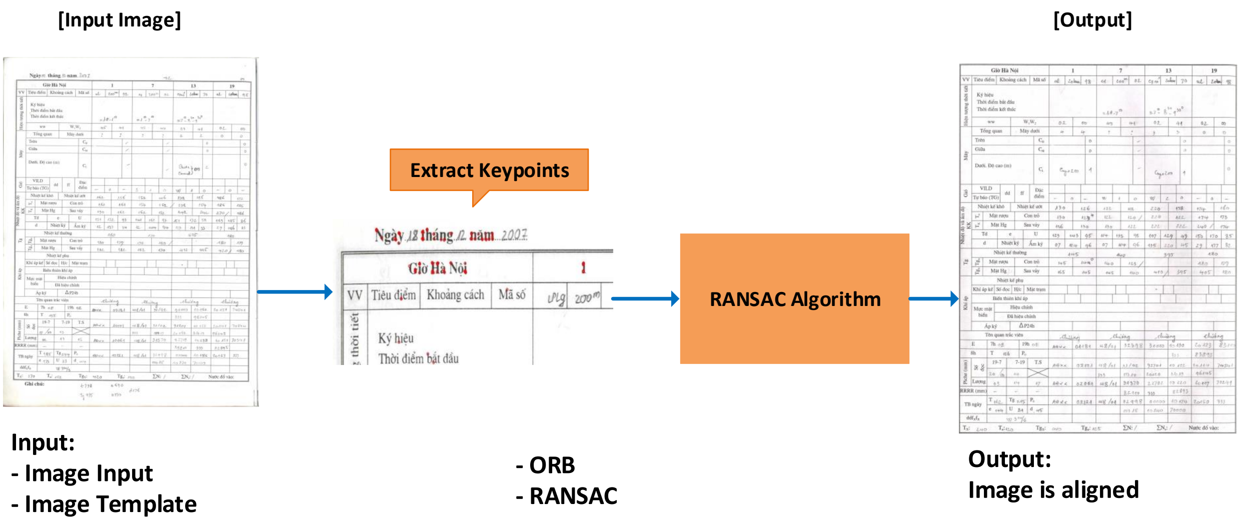
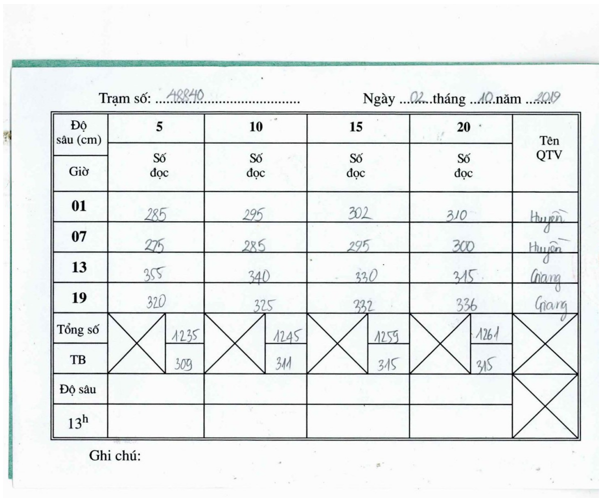


Figure 5: Identifying table structure based on sample table

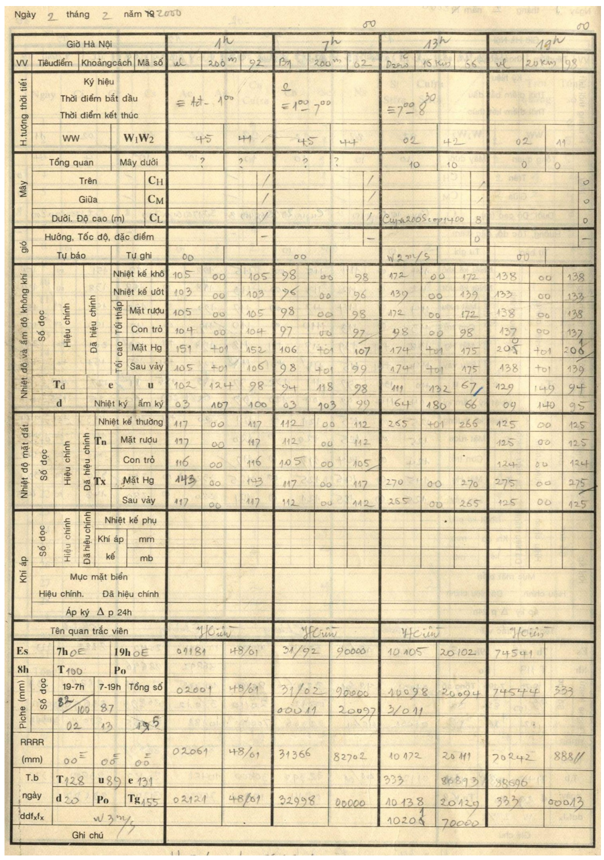


Trạm số: ...48840... Ngày ...02...tháng ...10...năm ...2019

Độ sâu (cm)	5		10		15		20		Tên QTV
Giờ	Số đọc		Số đọc		Số đọc		Số đọc		
01	285		295		302		310		Huyền
07	275		285		295		300		Huyền
13	355		340		330		315		Giang
19	320		325		332		336		Giang
Tổng số		1235		1245		1259		1261	
TB		309		311		315		315	
Độ sâu									
13h									

Ghi chú:

(a) Table has a simple structure



(b) The table has a complex structure like table_skt1_new

Figure 6: Some tables are labeled table_normal

1. Dataset Preparation

We experiment on 3 data sets: - The general dataset: The dataset is aggregated from three databases ICDAR19 [14], Marmot [15] and Github [16]. The dataset has been labeled and provided by Devashish Prasad [9]. The dataset is ultilized for table detection task, table classification and

TABLE I: COMPARISON OF TABLE FINDING RESULTS ON GENERAL DATA.
(data compiled from three databases ICDAR19 [15], Marmot [15] and Github [16])

Model	**IoU**				**Wags**
	0.6	0.7	0.8	0.9	
Cascade mask RCNN HRNetv2pW32 [9]	0.9604	0.9606	0.9658	**0.9696**	0.9641
Cascade mask RCNN HRNetv2pW40	**0.9667**	**0.9669**	**0.9674**	0.9693	**0.9675**

TABLE II: COMPARISON TABLE FINDING RESULTS ON DATA SET TABLEBANK.

Model	**Word**			**Latex**			**Word+Latex**		
	P	R	F1	P	R	F1	P	R	F1
X101(Word) [17]	0.9352	0.9398	0.9375	0.9905	0.5851	0.7356	0.9579	0.7474	0.8397
X152(Word) [17]	0.9418	0.9415	**0.9416**	0.9912	0.6882	0.8124	0.9641	0.8041	0.8769
X101(Latex) [17]	0.8453	0.9335	0.8872	0.9819	0.9799	0.9809	0.9159	0.9587	0.9368
X152(Latex) [17]	0.8476	0.9264	0.8853	0.9816	0.9814	**0.9815**	0.9173	0.9562	0.9364
X101(Word+Latex) [17]	0.9178	0.9363	0.9270	0.9827	0.9784	0.9806	0.9526	0.9592	**0.9559**
X152(Word+Latex) [17]	0.9229	0.9266	0.9247	0.9837	0.9752	0.9795	0.9557	0.9530	0.9543
The proposed method (Word -3000 sample)	0.9563	0.9514	**0.9538**	0.9276	0.9826	0.9543	0.9365	0.9687	0.9523
The proposed method (Latex - 2000 sample)	0.9112	0.8698	0.8900	0.9845	0.9842	**0.9844**	0.9646	0.9466	0.9556
The proposed method (All 4000 sample)	0.9286	0.9211	0.9248	0.9749	0.9797	0.9773	0.9620	0.9588	**0.9604**

TABLE III: TABLE FINDING RESULTS AFTER FINETURNING MODEL AND DATA ON KTTV DATA.

Model	**IoU**				**Wags**
	0.6	0.7	0.8	0.9	
Cascade mask RCNN HRNetv2pW40 + small data is dilation	0.9784	0.9855	0.9855	0.9855	0.9837
Cascade mask RCNN HRNetv2pW40 fineturning general dataset	0.9723	0.9774	0.9794	0.9803	0.9773
Cascade mask RCNN HRNetv2pW40 fineturning general dataset + dilation	**0.9865**	**0.9865**	**0.9883**	**0.9883**	**0.9874**

TABLE IV: RESULTS OF CLASSIFICATION OF TYPES OF HYDROLOGICAL TABLES.

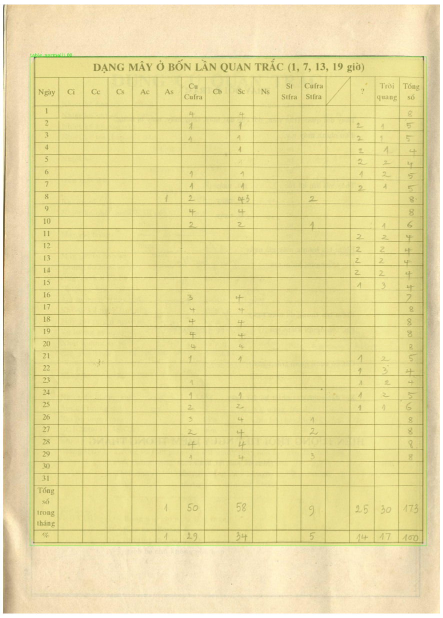
Table type name	**Correct identification number**	**Amount to be identified**	**Probability**
Table_normal	115	115	1.0
Table_sk1_old	233	233	1.0
Table_sk1_new	190	190	1.0

no-table regions. This general dataset includes a total of 1934 photos, the documents in the photos are new, have good image quality, and are available in 2 languages: English and Chinese.

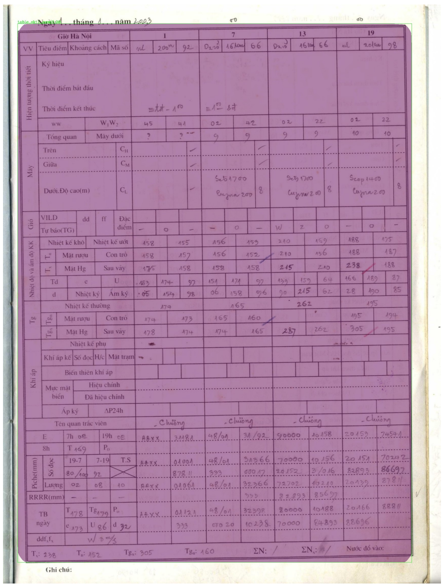
- Tablebank dataset [10]: Tablebank . . . to evaluate the quality of the table detector. Tablebank is a relatively large data set of 278582 images divided into 2 main types, Latex and Word.

- Hydro-meteorological data: The first collected hydro-meteorological data set includes 780 high-resolution scanned images (360dbi), equivalent to the image width of 6 inches. Data is taken from 13 different types of technical books related to meteorology, for example: basic meteorological observation book, soil temperature monitoring book. The data is divided and labeled into 3 types of tables table_skt1_old, table_skt_1_new , table_normal.

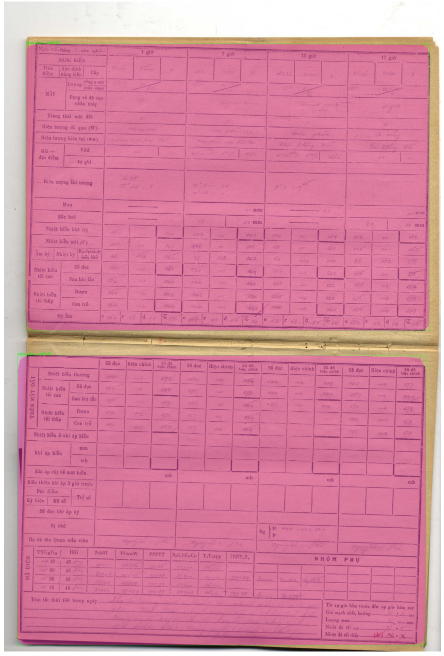
In Figure 3 (a) the skt1_old tables are the old tables of the meteorological observation books dating from 1969 to 1985. The pictures are often of poor quality, due to old books, faded ink, and bad paper. Figure 3 (b) skt_1_new are the tables of the basic meteorological observation books from 2000 onwards. These document images are relatively good, but the table structure is very complicated, the vertical lines are short, not aligned from the end to the end, so it is difficult to identify the links from the table cells automatically. These two types of tables have complex structures.



(a) Detect table table_nomal



(b) Detect table table_stk1_new



(c) Detect table table_stk1_old

Figure 7: Some results of the hydrometeorological data set table detection.

Figure 6 is a description of the table_normal table type. Figure 6-a The table_nomal table has a simple structure, these tables can be identified by the traditional methods of row and column identification, or using learning deep[14], Figure 6-b is a type of table with a complex structure that is highly similar to table_skt1_new. However, because the amount of data we currently have is too small to classify this table as a separate type, we have included it with the table_normal table type. In addition, this is also used to test the robustness of the model's classifier. In total, our data includes 780 images, in which table_normal has 243 images, table_skt1_old has 300 images, table_skt1_new has 237 images. We will rely on this data set to evaluate the quality of the model.

2. Results and analysis

We conduct model testing on the data sets presented in section IV.2.a. Compare some CNN network models, and the capabilities of our proposed model. We also tested some post-processing techniques related to output table alignment. Our experiments are performed on Google Colaboratory platform with P100 PCIE GPU with 16GB GPU memory, Intel(R) Xeon(R) CPU @2.30GHz and 12.72 GB Ram.

For each dataset, we use 60% for training set, 40% for testing set. Specifically with the general dataset, we use 1160 images for the training set, 774 images for the test set; With the KTTV dataset, we use 486 images for the training set, 294 images for the test set. First, we use common data to evaluate the panel detection of two models Cascade mask R-CNN HRNetv2pW32 (CascadeTabNet [9]) and Cascade mask R-CNN HRNetv2pW40 (W represents feature width).

a) Experimental results on the general data set

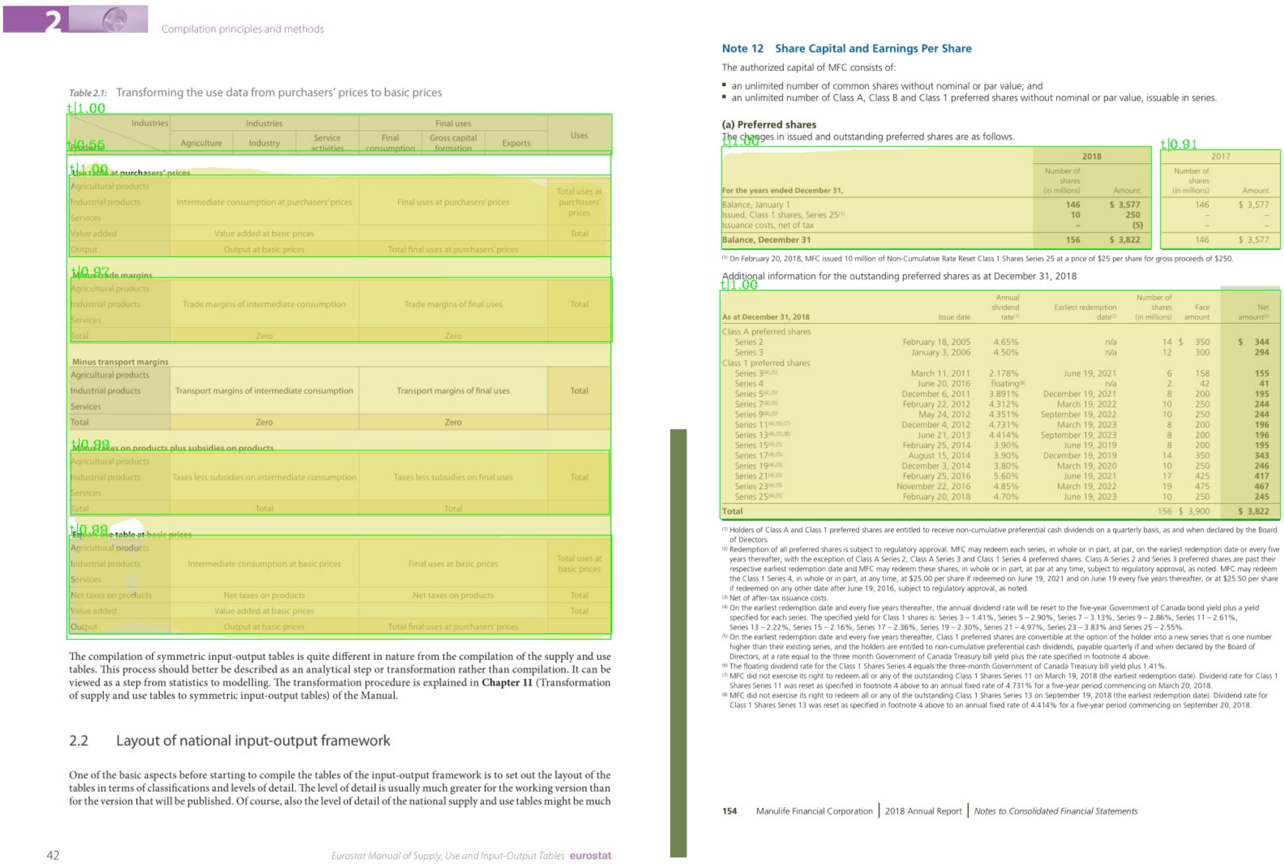
From the results of Table I, it is found that the Cascade mask R-CNN model HRNetv2pW40 has higher accuracy. This can be explained by the fact that we use more feature widths. Therefore, the Cascade mask R-CNN HRNetv2pW40 model will be selected to be used to conduct further evaluation tests. Table I compares the results between the models, based on the F1 score. The proposed model achieves higher results than the Cascade mask RCNN HRNetv2pW32 model in most IoU indexes with an average accuracy of 96.75%.

b) Experimental results on the TableBank data set

Second, we evaluate the table detection quality of the model on the Tablebank dataset. The results show that the model achieves the best results of 96.44% with the F1 score, with both Latex and Word data. While, we only use 4000 samples for training instead of using 260582 training samples of original setting. The test set uses the split ratio of the Tablebank dataset (5719 for latex, 2281 words, 8000 samples for both)

c) Experimental results on hydrometeorological data set

Third, we test the Cascade mask R-CNN HRNetv2pW40 model with hydrometeorological data, combined with the



(b)

(a)

Figure 8: Some error table detection results.

use of dilation enhancement technique. Here, we did not use the smudge technique to enhance the image as old hydrometeorological images are of poorer quality than the original modern images used in [9], the use of smudge will cause the image to lose the features. required to detect the table. We use Cascade mask model RCNN HRNetv2pW40 with data divided into 3 types of tables. From the experimental results in Table III, we see that the model gives relatively good panel detection results. After fine-tuning the model and data, the table detection results increased by about 2% compared to applying Cascade mask RCNN HRNetv2pW40 only, averaging 98.74%.

Table IV presents the results of table classification of phytosanitary datasets performed with threshold IoU = 0.8. The probability of table classification is 100%. Some results of the panel detection are shown in Figure 7. Figure 7-a, b, c illustrates the blue highlighted detection panel area and the label of the detected panel type of the corresponding table_sk1_new, table_normal and table_skt1_old.

Some of the error cases in table detection are mentioned in Figure 8. Figure 8-a is a case of missing table part recognition, a table is split into 5 sub-tables caused by the rows in this table being far apart. The model is interpreted as separate tables. Figure 8-b is an error identification case, table 1 is divided into 2 tables vertically, because the columns in the table are far apart and the horizontal line is broken between the two columns.

On the KTTV dataset, the results suggest the post-processing step of image alignment, using the sample table and ORB (Binary Robust Independent Elementary Features) and RANSAC algorithms to identify the structure of the table, conducting experiments on the skt_new table. The red and blue lines in Figure 9-a represent the area of interest as a result of the template table alignment and definition. The dark red and purple lines in Figure 9-a represent a match between the table of interest and the sample table. Thereby we see the cells in the table are clearly identified. Next with the cells that have been previously marked based on the sample table. Will proceed to split the cells and perform OCR and save the results to

(a) (b)

Figure 9: Illustration of using sample table to identify cells of interest.

the excel file. Figure 9-b is the result of the table extracted to the excel file, we see that the table structure recognition is very good. However, the handwriting recognition part is not good, achieving an accuracy of 65%, it is necessary to have another study to handle this handwriting.

V. CONCLUSION

This paper proposes an overall method based on the Cascade mask R-CNN HRNetw40 model to detect table objects, combining ORB and RANSAC techniques to digitize table documents. The results show that the proposed algorithm is capable of detecting and classifying document tables very well. The table detection results reached 98.74% on the F1 score and recognized different types of tables with a very high accuracy rate on the three test data sets. The results obtained in the study show the feasibility of applying deep learning techniques in the process of digitizing documents, especially tabular documents.

REFERENCES

[1] P. Pyreddy and W. B. Croft, "Tintin: A system for retrieval in text tables," in *Proceedings of the Second ACM International Conference on Digital Libraries*, ser. DL '97. New York, NY, USA: Association for Computing Machinery, 1997, p. 193–200. [Online]. Available: https://doi.org/10.1145/263690.263816

[2] G. V. S. S. K. R. Naganjaneyulu, N. V. Sathwik, and A. Narasimhadhan, "A multi clue heuristic based algorithm for table detection," in *2016 IEEE Region 10 Conference (TENCON)*, 2016, pp. 1246–1249.

[3] T. Dhiran and R. Sharma, "Table detection and extraction from image document," *International Journal of Computer & Organization Trends*, vol. 3, no. 7, pp. 275–278, 2013.

[4] A. Gilani, S. R. Qasim, I. Malik, and F. Shafait, "Table detection using deep learning," 09 2017.

[5] F. Shafait and R. Smith, "Table detection in heterogeneous documents," in *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems*, ser. DAS '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 65–72. [Online]. Available: https://doi.org/10.1145/1815330.1815339

[6] T. Huynh-Van, K. Nguyen-An, T. L. B. Khanh, H.-J. Yang, T. A. Tran, and S.-H. Kim, "Learning to detect tables in document images using line and text

information," in *Proceedings of the 2nd International Conference on Machine Learning and Soft Computing*, ser. ICMLSC '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 151–155. [Online]. Available: https://doi.org/10.1145/3184066.3184091
[7] F. Kleber, M. Diem, H. Dejean, J.-L. Meunier, and E. Lang, "Matching table structures of historical register books using association graphs," in *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, 2018, pp. 217–222.
[8] S. Schreiber, S. Agne, I. Wolf, A. Dengel, and S. Ahmed, "Deepdesrt: Deep learning for detection and structure recognition of tables in document images," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 01, 2017, pp. 1162–1167.
[9] D. Prasad, A. Gadpal, K. Kapadni, M. Visave, and K. Sultanpure, "Cascadetabnet: An approach for end to end table detection and structure recognition from image-based documents," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 2439–2447.
[10] M. Li, L. Cui, S. Huang, F. Wei, M. Zhou, and Z. Li, "Tablebank: Table benchmark for image-based table detection and recognition," *CoRR*, vol. abs/1903.01949, 2019. [Online]. Available: http://arxiv.org/abs/1903.01949
[11] Z. Cai and N. Vasconcelos, "Cascade r-cnn: High quality object detection and instance segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 5, pp. 1483–1498, 2021.
[12] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 5686–5696.
[13] K. Chen, J. Wang, J. Pang, Y. Cao, Y. Xiong, X. Li, S. Sun, W. Feng, Z. Liu, J. Xu, Z. Zhang, D. Cheng, C. Zhu, T. Cheng, Q. Zhao, B. Li, X. Lu, R. Zhu, Y. Wu, J. Dai, J. Wang, J. Shi, W. Ouyang, C. C. Loy, and D. Lin, "Mmdetection: Open mmlab detection toolbox and benchmark," *CoRR*, vol. abs/1906.07155, 2019. [Online]. Available: http://arxiv.org/abs/1906.07155
[14] L. Gao, Y. Huang, H. Déjean, J.-L. Meunier, Q. Yan, Y. Fang, F. Kleber, and E. Lang, "Icdar 2019 competition on table detection and recognition (ctdar)," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, 2019, pp. 1510–1515.
[15] J. Fang, X. Tao, Z. Tang, R. Qiu, and Y. Liu, "Dataset, ground-truth and performance metrics for table detection evaluation," in *2012 10th IAPR International Workshop on Document Analysis Systems*, 2012, pp. 445–449.
[16] sgrpanchal31, "sgrpanchal31/table-detection-dataset." [Online]. Available: https://github.com/sgrpanchal31/table-detection-dataset
[17] Doc-Analysis, "doc-analysis/tablebank." [Online]. Available: https://github.com/doc-analysis/TableBank
[18] S. A. Siddiqui, I. A. Fateh, S. T. R. Rizvi, A. Dengel, and S. Ahmed, "Deeptabstr: Deep learning based table structure recognition," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, 2019, pp. 1403–1409.

Hai-Hong Phan is currently a lecturer of Faculty of Information Technology in Le Quy Don Technical University She received a Ph.D. degree in computer science from the University of CY Cergy Paris, France, in 2019. She received the B.S. and M.S. degrees in computer science from Le Quy Don Technical University, Vietnam, in 2009 and 2013, respectively. Her researches focus on computer vision, image processing, document analysis, action recognition, and machine learning.
Email: hongpth@lqdtu.edu.vn

Ngo Dai Duong is currently studying for master degree at Le Quy Don Technical University. He received the B.S. degree in Information Technology from Le Quy Don Technical University in 2014. His researches focus on computer vision, document analysis and machine learning.
Email: daiduong28789@hotmail.com