Indian Institute of Information Technology Vadodara – International Campus Diu
Institute of National Importance declared by an act of Parliament
Education Hub, Kevdi, Diu (U.T) - 362520

Comparative Analysis of Table Recognition.

This is the group project of course CS391

Mentored By-

Dr. Deepika Gupta

Submitted by

Group - 01

Group Members:

Ishaan Ashish Arora - 202011012
Yashesh Bhavsar - 202011017
Dev Juneja - 202011020
Hari Gopal Nayak Jarupla - 202011024
Suraj Kumar - 202011069

November 2022

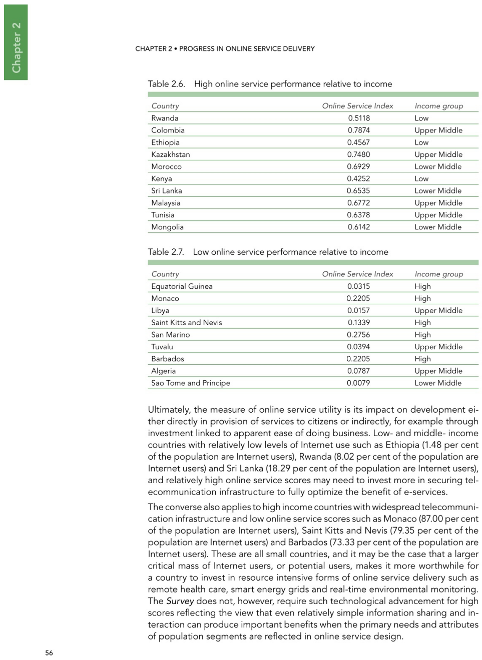
Content

Abstarct
In this particular report, we discuss the comparative analysis of the table detection models that have been developed. In our analysis, we focused on the precision, accuracy and F1 scores of the respective models. Based on the findings we try to decide on an efficient model which provides better results. In the process, we analyze the shortcomings of each of the four models. We have provided the numerical analysis of all the models. Upon training and running all the models we analysed the functioning. The motivation behind this analysis was to come up with the best possible solution to the problem of table detection(recognition)

Keywords RCNN-HRNet,Faster-CNN(Fr-CNN),YOLOv3,VGG-19

1 Introduction

The world is evolving, becoming more digitalized. Digital documents are increasingly replacing paper-based ones. These papers include various table-based data with different appearances and layouts and alignments. Table identification and table structure recognition are two subtasks of an automated table information extraction approach. While table structure recognition identifies the rows and columns to identify individual table cells, table detection identifies the area of the picture that contains the table.
We have used RCNN/HRNet as an enhanced deep learning-based end-to-end method for resolving the two sub-problems with a single model. Instance segmentation is used to address the issue of table detection. Each image undergoes table segmentation, which aims to pinpoint every instance of the table at the pixel level inside the image.To determine the structure of the table, we similarly execute table cell segmentation on each picture to forecast segmented portions of table cells inside each table. The model simultaneously predicts table and cell areas in a single inference. The model simultaneously divides tables into two categories: bordered (based on rules) and borderless (without rules).
For extracting cells from bordered tables, we use rule-based conventional text detection and line detection algorithms. We show how iterative transfer learning works to help CNN learn from less training data while also enabling it to excel on various datasets by fine-tuning it on the appropriate datasets. A novel approach of image augmentation was also inserted into the training process to enhance the accuracy of table identification and enable it learn efficiently.
In order to comprehend tabular data in document pictures, an automatic table identification approach must first solve the challenges of table detection and table structure recognition. Deep learning-based solutions are being used in more recent efforts, which also try to create an end-to-end solution. The Convolution Neural Network (CNN) model, provides an enhanced deep learning-based end-to-end method in this research for handling the issues of table identification and structure recognition. The models have tested on publicly available datasets from ICDAR 2013, ICDAR 2019, and Table-Bank. We test the accuracy results for the ICDAR 2013 and Table-Bank dataset. Additionally, using the ICDAR 2019 table structure identification dataset.


Chapter 2

CHAPTER 2 • PROGRESS IN ONLINE SERVICE DELIVERY

Table 2.6. High online service performance relative to income

Country	*Online Service Index*	*Income group*
Rwanda	0.5118	Low
Colombia	0.7874	Upper Middle
Ethiopia	0.4567	Low
Kazakhstan	0.7480	Upper Middle
Morocco	0.6929	Lower Middle
Kenya	0.4252	Low
Sri Lanka	0.6535	Lower Middle
Malaysia	0.6772	Upper Middle
Tunisia	0.6378	Upper Middle
Mongolia	0.6142	Lower Middle

Table 2.7. Low online service performance relative to income

Country	*Online Service Index*	*Income group*
Equatorial Guinea	0.0315	High
Monaco	0.2205	High
Libya	0.0157	Upper Middle
Saint Kitts and Nevis	0.1339	High
San Marino	0.2756	High
Tuvalu	0.0394	Upper Middle
Barbados	0.2205	High
Algeria	0.0787	Upper Middle
Sao Tome and Principe	0.0079	Lower Middle

Ultimately, the measure of online service utility is its impact on development either directly in provision of services to citizens or indirectly, for example through investment linked to apparent ease of doing business. Low- and middle- income countries with relatively low levels of Internet use such as Ethiopia (1.48 per cent of the population are Internet users), Rwanda (8.02 per cent of the population are Internet users) and Sri Lanka (18.29 per cent of the population are Internet users), and relatively high online service scores may need to invest more in securing telecommunication infrastructure to fully optimize the benefit of e-services.

The converse also applies to high income countries with widespread telecommunication infrastructure and low online service scores such as Monaco (87.00 per cent of the population are Internet users), Saint Kitts and Nevis (79.35 per cent of the population are Internet users) and Barbados (73.33 per cent of the population are Internet users). These are all small countries, and it may be the case that a larger critical mass of Internet users, or potential users, makes it more worthwhile for a country to invest in resource intensive forms of online service delivery such as remote health care, smart energy grids and real-time environmental monitoring. The *Survey* does not, however, require such technological advancement for high scores reflecting the view that even relatively simple information sharing and interaction can produce important benefits when the primary needs and attributes of population segments are reflected in online service design.

56

Figure 1 Input file

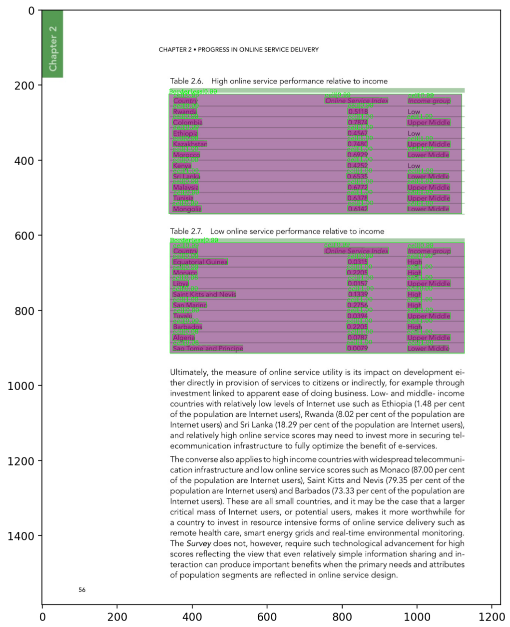


Figure 2 Output file

2 Model1:YOLOv3

In this model, a YOLO-based method is used. Considering the large difference between document objects and natural objects, this architecture introduces some adaptive adjustments to YOLOv3, including an anchor optimization strategy and two post-processing methods. For anchor optimization, the model uses k-means clustering to find anchors which are more suitable for tables rather than natural objects and make it easier for our model to find the exact positions of tables In the post-processing process, the extra whitespaces and noisy page objects (e.g. page headers, page footers and other similar objects) are removed from the predicted results so that the model can get more accurate table margins and higher IoU scores. The model is evaluated on two datasets from ICDAR 2013 and ICDAR 2017 and achieves state-of-the-art performance.

2.1 Algorithm

Algorithm 1: Erasing whitespace margin algorithm

Input: P: all pixels of the predicted region.
Output: R: result region after erasing the whitespace margins.

$orientations = \{up, down, left, right\}$;
$R = P$;
for ori **in** $orientations$ **do**
  $M = \{$all pixels belong to the margin of R in direction $ori\}$;
  while *all pixels* $\in M$ *are white color* **do**
    $R = R - M$;
    $M = \{$all pixels belong to the margin of R in direction $ori\}$;
  end
end
return R;

Figure 3 YOLOv3 algorithm

2.2 Dataset performance

This method is trained on the training set of ICDAR 2017 POD Competition. The size of input image is 416 $\times$ 416, the same as YOLOv3. Throughout the training, we use a batch size of 64, a learning rate of 0.001, a momentum of 0.9 and a decay rate of 0.0005. And we train the model for 198200 iterations in total.

2.3 Conclusion from the model

This model faces issues in detecting the width and height of tables accurately. The model also fails against strategies to deal with the particularity of unruled and less-ruled tables.

3 Model2:VGG-19

In this model: a novel end-to-end deep learning model for both table detection and structure recognition is used. The model exploits the interdependence between the twin tasks of table detection and table structure recognition to segment out the table and column regions. This is followed by semantic rule-based row extraction from the identified tabular sub-regions. This model and extraction approach was evaluated on the publicly available ICDAR 2013 and Marmot Table datasets obtaining state of the art results.

3.1 Algorithm

The model uses a pre-trained VGG-19 layer as the base network. The fully connected layers (layers after pool5) of VGG-19 are replaced with two (1x1) convolution layers. Each of these convolution layers (conv6) uses the ReLU activation followed by a dropout layer having probability of 0.8 (conv6 + dropout). Following this layer, two different branches of the decoder network are appended. This is according to the intuition that the column region is a subset of the table region. Thus, the single encoding network can filter out the active regions with better accuracy using features of both table and column regions. The output of the (conv6 + dropout) layer is distributed to both decoder branches. In each branch, additional layers are appended to filter out the respective active regions. In the table branch of the decoder network, an additional (1x1) convolution layer, conv7 table is used, before using a series of fractionally strided convolution layers for upscaling the image. The output of the conv7 table layer is also up-scaled using fractionally strided convolutions, and is appended with the pool4 pooling layer of the same dimension. Similarly, the combined feature map is again up-scaled and the pool3 pooling is appended to it. Finally, the final feature map is upscaled to meet the original image dimension.

3.2 Dataset performance

VGG-19 requires both table and structure annotated datafor training.The Marmot table detection dataset and manually annotated the structure information were used. There are a total of 1016 documents containing tables including both Chinese and English documents, out of which 509 English documents are annotated and used for training.In this model we get the recall as 0.9501,precision as 0.9547 and F1-score as 0.9547.

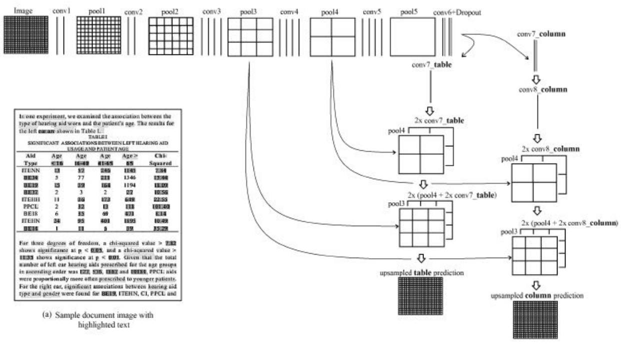


Figure 4 VGG-19

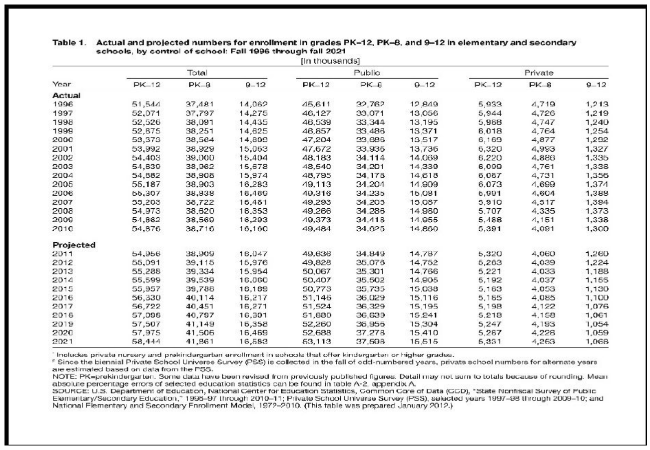


Table 1. **Actual and projected numbers for enrollment in grades PK–12, PK–8, and 9–12 in elementary and secondary schools, by control of school: Fall 1996 through fall 2021**

[In thousands]

	Total			Public			Private		
Year	PK–12	PK–8	9–12	PK–12	PK–8	9–12	PK–12	PK–8	9–12
Actual									
1996	51,544	37,481	14,062	45,611	32,762	12,849	5,933	4,719	1,213
1997	52,071	37,797	14,275	46,127	33,071	13,056	5,944	4,726	1,219
1998	52,526	38,091	14,435	46,539	33,344	13,195	5,988	4,747	1,240
1999	52,875	38,251	14,625	46,857	33,486	13,371	6,018	4,764	1,254
2000	53,373	38,564	14,809	47,204	33,686	13,517	6,169	4,877	1,292
2001	53,992	38,929	15,063	47,672	33,936	13,736	6,320	4,993	1,327
2002	54,403	39,000	15,404	48,183	34,114	14,069	6,220	4,886	1,335
2003	54,639	38,962	15,678	48,540	34,201	14,339	6,099	4,761	1,338
2004	54,882	38,908	15,974	48,795	34,178	14,618	6,087	4,731	1,356
2005	55,187	38,903	16,283	49,113	34,204	14,909	6,073	4,699	1,374
2006	55,307	38,838	16,469	49,316	34,235	15,081	5,991	4,604	1,388
2007	55,203	38,722	16,481	49,293	34,205	15,087	5,910	4,517	1,394
2008	54,973	38,620	16,353	49,266	34,286	14,980	5,707	4,335	1,373
2009	54,862	38,569	16,293	49,373	34,418	14,955	5,488	4,151	1,338
2010	54,876	38,716	16,160	49,484	34,625	14,860	5,391	4,091	1,300
Projected									
2011	54,956	38,909	16,047	49,636	34,849	14,787	5,320	4,060	1,260
2012	55,091	39,115	15,976	49,828	35,076	14,752	5,263	4,039	1,224
2013	55,288	39,334	15,954	50,067	35,301	14,766	5,221	4,033	1,188
2014	55,599	39,539	16,060	50,407	35,502	14,905	5,192	4,037	1,155
2015	55,957	39,788	16,169	50,773	35,735	15,038	5,163	4,053	1,130
2016	56,330	40,114	16,217	51,146	36,029	15,116	5,185	4,085	1,100
2017	56,722	40,451	16,271	51,524	36,329	15,195	5,198	4,122	1,076
2018	57,098	40,797	16,301	51,880	36,639	15,241	5,218	4,158	1,061
2019	57,507	41,149	16,358	52,260	36,956	15,304	5,247	4,193	1,054
2020	57,975	41,506	16,469	52,688	37,278	15,410	5,287	4,228	1,059
2021	58,444	41,861	16,583	53,113	37,598	15,515	5,331	4,263	1,068

[1] Includes private nursery and prekindergarten enrollment in schools that offer kindergarten or higher grades.
[2] Since the biennial Private School Universe Survey (PSS) is collected in the fall of odd-numbered years, private school numbers for alternate years are estimated based on data from the PSS.
NOTE: PK=prekindergarten. Some data have been revised from previously published figures. Detail may not sum to totals because of rounding. Mean absolute percentage errors of selected education statistics can be found in table A-2, appendix A.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "State Nonfiscal Survey of Public Elementary/Secondary Education," 1996–97 through 2010–11; Private School Universe Survey (PSS), selected years 1997–98 through 2009–10; and National Elementary and Secondary Enrollment Model, 1972–2010. (This table was prepared January 2012.)

Figure 5 VGG-19: raw input image

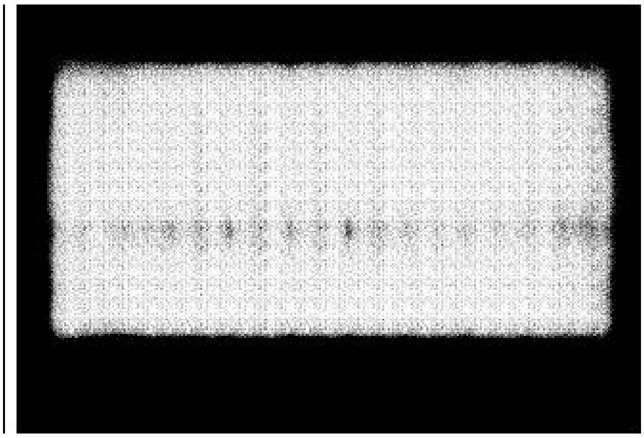


Figure 6 VGG-19 table mask

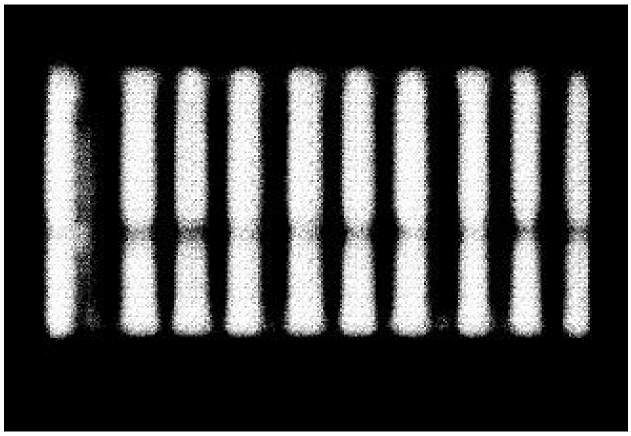


Figure 7 VGG-19 column mask

3.3 Conclusion from the model

The VGG-19 model,is a novel deep learning model trained on dual tasks of table detection and structure recognition in an end-to-end fashion. Existing approaches to information extraction treat detection and structure recognition as two distinct problems to be solved independently. VGG-19 is the first model to jointly address both tasks simultaneously, by exploiting the inherent interdependence between table detection and table structure identification. VGG-19 utilizes the knowledge from previously learned tasks and can transfer that knowledge to newer, related ones demonstrating transfer learning.

4 Model3:Faster-CNN

This model is an end-to-end system for table detection and structure recognition in document images called FCNN. This method is data driven, based on deep learning, and hence does not require any heuristics or rules to detect tables and to recognize their structure. This approach makes FCNN applicable to both, images as well as born-digital documents (e.g. PDFs, Word documents, and web pages, as they can be converted to images).

4.1 Algorithm

The first step in table understanding is detecting the locations of tables within a document. Conceptually, the problem is similar to the detection of objects in natural scene images. Therefore, in this approach domain adaptation is used and transfer learning by utilizing deep learning-based object detection frameworks originally created for natural scene images and tested their ability to cope with tabular structures in scanned document images.
The FRCNN model consists of two distinct parts: First they generate region proposals based on the input image by a so-called region proposal network (RPN). Afterwards, these proposals are classified using a Fast-RCNN network. Both modules share parameters and can be trained end-to-end.
After a table has successfully been detected and its location is known to the system, the next step is in understanding its contents is to recognize and locate the rows and columns which make up the physical structure of the table. This step is inherently different from the preceding table detection. The key difference is not only that there are significantly more rows and columns present in a table image than there are tables in a document but these tabular structures are generally located in very close proximity. These two factors make this task so difficult for FRCNN and ask for a different approach.

4.2 Dataset performance

As FCNN is based on a data-driven approach, there was the need for a sufficiently large dataset. The largest the publicly available dataset is the Marmot dataset for table recognition1 published by the Institute of Computer Science and Technology of Peking University.
The FCN-based segmentation models of FCNN were trained for 60,000 iterations employing a standard SGD optimizer with a fixed learning rate of 10-10 and classical the momentum of 0.99.

Figure 8 FR-CNN output(row detecting)

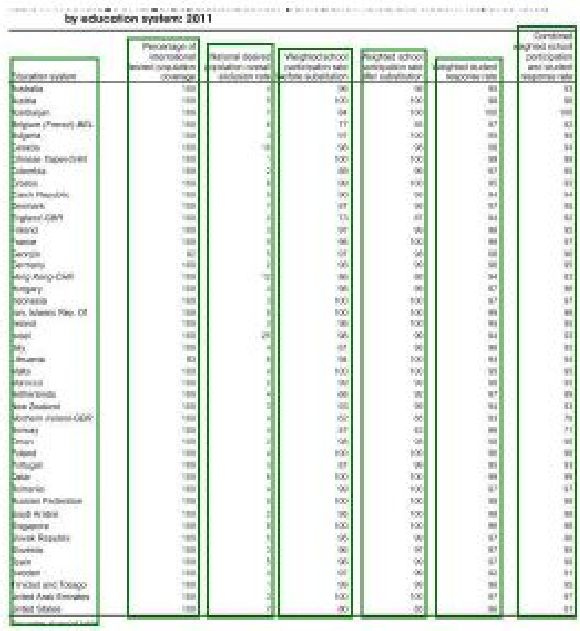
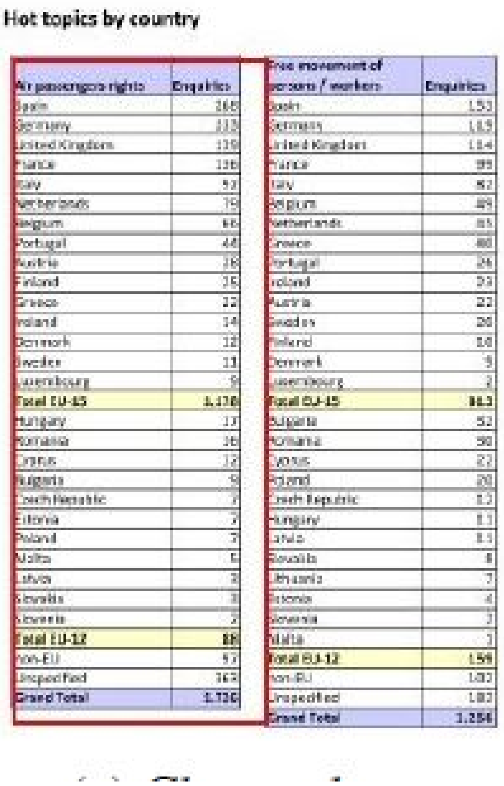


Figure 9 FR-CNN output(column detecting)

4.3 Conclusion from the model

This model further utilized recently published insights from deep learning-based semantic segmentation research for recognizing structures within tables. However, the model fails to resolve its persisting issues with recognizing structures which are in very close proximity to other elements of interest in an image.
Failure cases



Hot topics by country

Air passengers rights	Enquiries	Free movement of persons / workers	Enquiries
Spain	265	Spain	153
Germany	223	Germany	115
United Kingdom	139	United Kingdom	114
France	136	France	95
Italy	93	Italy	92
Netherlands	79	Belgium	49
Belgium	66	Netherlands	45
Portugal	44	Greece	40
Austria	38	Portugal	26
Finland	35	Ireland	23
Greece	22	Austria	22
Ireland	14	Sweden	20
Denmark	12	Finland	10
Sweden	11	Denmark	5
Luxembourg	9	Luxembourg	2
Total EU-15	1,176	Total EU-15	813
Hungary	17	Bulgaria	52
Romania	16	Romania	50
Cyprus	12	Cyprus	27
Bulgaria	9	Poland	20
Czech Republic	7	Czech Republic	12
Estonia	7	Hungary	11
Poland	7	Latvia	11
Malta	5	Slovakia	8
Latvia	3	Lithuania	7
Slovakia	3	Estonia	4
Slovenia	2	Slovenia	2
Total EU-12	88	Malta	1
non-EU	97	Total EU-12	158
Unspecified	365	non-EU	103
Grand Total	1,726	Unspecified	183
		Grand Total	1,284

Figure 10 Close Columns

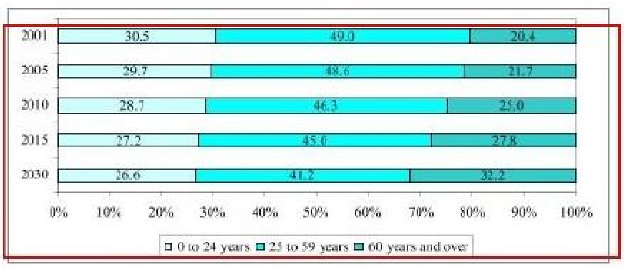


Figure 11 Bar chart confusion

Sample Group	Some Year 1 Head Start Participation	No Year 1 Head Start Participation	Total
All Randomly Assigned (N=4,667):			
3-Year-Old Cohort			
Head Start Group	85.1%	14.9%	100%
Control Group	17.3%	82.7%	100%
4-Year-Old Cohort			
Head Start Group	79.8%	20.2%	100%
Control Group	13.9%	86.1%	100%

Figure 12 Nested row hierarchy

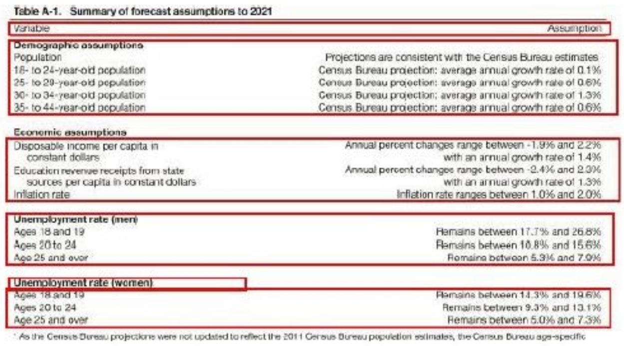


Table A-1. Summary of forecast assumptions to 2021

Variable	Assumption
Demographic assumptions	
Population	Projections are consistent with the Census Bureau estimates
18- to 24-year-old population	Census Bureau projection: average annual growth rate of 0.1%
25- to 29-year-old population	Census Bureau projection: average annual growth rate of 0.6%
30- to 34-year-old population	Census Bureau projection: average annual growth rate of 1.3%
35- to 44-year-old population	Census Bureau projection: average annual growth rate of 0.6%
Economic assumptions	
Disposable income per capita in constant dollars	Annual percent changes range between -1.9% and 2.2% with an annual growth rate of 1.4%
Education revenue receipts from state sources per capita in constant dollars	Annual percent changes range between -2.4% and 2.3% with an annual growth rate of 1.3%
Inflation rate	Inflation rate ranges between 1.0% and 2.0%
Unemployment rate (men)	
Ages 18 and 19	Remains between 17.7% and 26.8%
Ages 20 to 24	Remains between 10.8% and 15.6%
Age 25 and over	Remains between 5.3% and 7.9%
Unemployment rate (women)	
Ages 18 and 19	Remains between 14.3% and 19.6%
Ages 20 to 24	Remains between 9.3% and 13.1%
Age 25 and over	Remains between 5.0% and 7.3%

[1] As the Census Bureau projections were not updated to reflect the 2011 Census Bureau population estimates, the Census Bureau age-specific

Figure 13 Very close rows

5 Model4:RCNN-HRNet

5.1 Pipeline architecture

It predicts the segmentation masks for tables of two types as bordered and borderless. Next in the pipeline, there are separate branches for bordered and borderless tables. Depending on the type of the detected table it is further processed by its respective branch post-processing module. Post-processing modules perform trivial tasks of arranging and cleaning the outputs of the model.
In the borderless branch, the predicted cells are detected inside the table are into rows and columns based on their positions. Then the model estimates the missing table lines using the positions of identified rows and columns. Based on these lines, for undetected cells, detect the cells using a contour-based text detection algorithm. And finally, Row-span and Col-span cells are also identified after estimating the lines.
In the bordered branch, a conventional algorithm of line detection is used to detect lines of bordered tables. The cells are identified using the line intersection points. And within each cell, the text regions are detected by using the contour-based text detection algorithm.

5.2 Algorithm

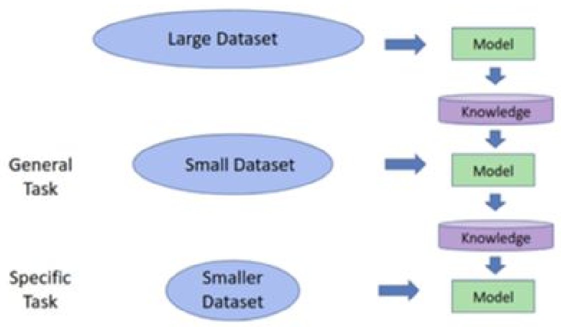


Figure 14 HRNet Alogrithm

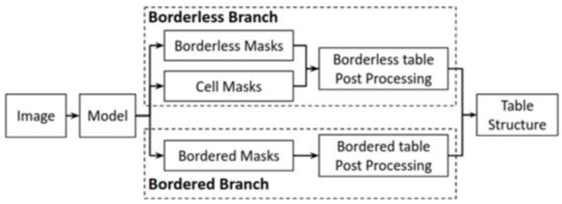


Figure 15 HRNet Pipeline

5.3 Conclusion from the model

The model starts learning for a general task and iteratively it learns to perform well on specific tasks. The model recognizes structures within tables by predicting table cell masks while using the line information as well. Improving the post-processing modules can further enhance the accuracy of the end-to-end model.The model performs better on various public datasets.

6 Experimental Analysis

Accuracy is a metric that generally describes how the model performs across all classes. It is useful when all classes are of equal importance. It is calculated as the ratio between the number of correct predictions to the total number of predictions.

Accuracy=

$$\frac{True_{positive} + True_{negative}}{True_{positive} + True_{negative} + False_{positive} + False_{negative}} \tag{1}$$

Precision is calculated as the ratio between the number of Positive samples correctly classified to the total number of samples classified as Positive (either correctly or incorrectly). The precision measures the model's accuracy in classifying a sample as positive.

$$Precision = \frac{True_{positive}}{True_{positive} + False_{positive}} \tag{2}$$

Recall is calculated as the ratio between the number of Positive samples correctly classified as Positive to the total number of Positive samples. The recall measures the model's ability to detect Positive samples. The higher the recall, the more positive samples detected.

$$Recall = \frac{True_{positive}}{True_{positive} + False_{negative}} \tag{3}$$

F1-score is a metric which takes into account both precision and recall

$$F1_{score} = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

Model	Precision	Recall	F1 Score
R-CNN HRNet [1]	0.9892	0.9921	0.9906
Faster R-CNN [2]	0.9615	0.9740	0.9677
VGG-19 [3]	0.9628	0.9697	0.9662
YOLOv3 [4]	0.9712	0.9821	0.9766

After thorough analysis of the metrics that affect the CNN models,we have come to the conclusion that the RCNN/HR-Net model has the best results for all the metrics.This model then proves to be the best among the other models presented in this report.

7 Conclusion

The report provides a detailed comparative analysis of table recognition. In the world of machine learning and deep learning, this topic is of great significance. Research in this domain has accelerated in recent years due to the growth of various deep-learning models, the complete method for identifying structures and detecting tables. It is demonstrated that current CNN architectures using instance segmentation, which were taught to recognise objects in photos of real scenes, are also highly good at detecting tables. Additionally, to learn well from a short quantity of data, picture augmentation methods and iterative transfer learning can be applied. The model begins by learning for a generic task, then over time, it progressively gains the ability to excel at more precise tasks. By predicting table cell masks while also employing line information, models -RCNN/HRNet, VGG-19, YOLOv3 and Faster-CNN can identify structures within tables. The accuracy of RCNN/HRNet is more than all the other models mentioned in the report. The accuracy of this model can be further be improved by using post-processing modules. Our future work would be to further improve the HRNet model by training and testing it on other possible datasets and to overcome the present failure cases.

References

[1] Prasad, D. et al. (2020) “Cascadetabnet: An approach for end to end table detection and structure recognition from image-based documents,” 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) [Preprint].Available at: https://doi.org/10.1109/cvprw50498.2020.00294.

[2] Schreiber, S. et al. (2017) “DeepDeSRT: Deep Learning for detection and structure recognition of tables in document images,” 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) [Preprint]. Available at: https://doi.org/10.1109/icdar.2017.192.

[3] Paliwal, S.S. et al. (2019) “TableNet: Deep Learning Model for end-to-end table detection and tabular data extraction from scanned document images,” 2019 International Conference on Document Analysis and Recognition (ICDAR) [Preprint]. Available at: https://doi.org/10.1109/icdar.2019.00029.

[4] Huang, Y. et al. (2019) “A YOLO-based table detection method,” 2019 International Conference on Document Analysis and Recognition (ICDAR) [Preprint]. Available at: https://doi.org/10.1109/icdar.2019.00135.

[5]S. Arif and F. Shafait. Table detection in document images using foreground and background features. In 2018 Digital Image Computing: Techniques and Applications (DICTA), pages 1–8, 2018.

[6]Zhaowei Cai and Nuno Vasconcelos. Cascade r-cnn: Delving into high quality object detection. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.

[7]Zhaowei Cai and Nuno Vasconcelos. Cascade R-CNN: high quality object detection and instance segmentation. CoRR, abs/1906.09756, 2019.

[8]Kai Chen, Jiaqi Wang, Jiangmiao Pang, Yuhang Cao, Yu Xiong, Xiaoxiao Li, Shuyang Sun, Wansen Feng, Ziwei Liu, Jiarui Xu, Zheng Zhang, Dazhi Cheng, Chenchen Zhu, Tianheng Cheng, Qijie Zhao, Buyu Li, Xin Lu, Rui Zhu, Yue Wu, Jifeng Dai, Jingdong Wang, Jianping Shi, Wanli Ouyang, Chen Change Loy, and Dahua Lin. Mmdetection: Open mmlab detection toolbox and benchmark. CoRR, abs/1906.07155, 2019.

[9]Zewen Chi, Heyan Huang, Heng-Da Xu, Houjin Yu, Wanxuan Yin, and Xian-Ling Mao. Complicated table structure recognition. ArXiv, abs/1908.04729, 2019.

[10]J. Fang, X. Tao, Z. Tang, R. Qiu, and Y. Liu. Dataset, ground truth and performance metrics for table detection evaluation.

[11]W. Farrukh, A. Foncubierta-Rodriguez, A. Ciubotaru, G.Jaume, C. Bejas, O. Goksel, and M. Gabrani. Interpreting data from scanned tables. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), volume 02, pages 5–6, 2017.

[12]L. Gao, Y. Huang, H. D´ejean, J. 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), pages 1510–1515, 2019.

[13]A. Gilani, S. R. Qasim, I. Malik, and F. Shafait. Table detection using deep learning. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), volume 01, pages 771–776, 2017.

[14]Max G¨obel, Tamir Hassan, Ermelinda Oro, and Giorgio Orsi.A methodology for evaluating algorithms for table understanding in pdf documents. In Proceedings of the 2012 ACM Symposium on Document Engineering, DocEng ’12, page 45–48, New York, NY, USA, 2012. Association for Computing Machinery.

[15]M. G¨obel, T. Hassan, E. Oro, and G. Orsi. Icdar 2013 table competition. In 2013 12th International Conference on Document Analysis and Recognition, pages 1449–1453, 2013.

[16]D. He, S. Cohen, B. Price, D. Kifer, and C. L. Giles. Multiscale multi-task fcn for semantic page segmentation and table detection. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), volume 01, pages 254–261, 2017.