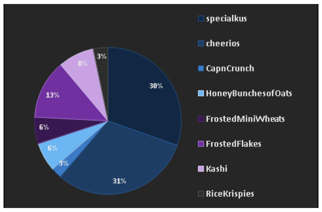
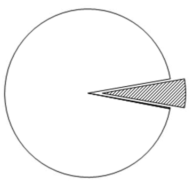


Ground Truth:	0.30	0.31	0.03	0.06	0.06	0.13	0.08	0.03
ChartOCR:	0.30	0.31	0.02	0.06	0.06	0.13	0.08	0.03
Revision:	<u>0.02</u>	<u>0.03</u>	<u>0.38</u>	<u>0.03</u>	<u>0.05</u>	<u>0.41</u>	<u>0.06</u>	<u>0.02</u>
RotationRNN:	0.30	0.31	0.02	0.06	0.06	0.13	0.08	0.03

Figure 6: Both ChartOCR and RotationRNN perform reasonably well for this image, but the Revision method does not work well due to the black background. (Underlines indicate un-matched results from ground-truth)



Ground Truth:	0.06	0.94		
ChartOCR:	0.06	0.94		
Revision:	<u>0.07</u>	<u>0.00</u>	<u>0.90</u>	<u>0.89</u>
RotationRNN:	<u>0.03</u>	0.94	<u>0.02</u>	<u>0.01</u>

Figure 7: For pie charts that contain detached sectors, only ChartOCR has satisfying performance. (Underlines indicate un-matched results from ground-truth)

charts. Rule-based method is not good at extracting stacked elements as shown in Figure 5 (a)(b) and it can be disturbed by the text content. Compared with ResNet+Faster-RCNN in (d), ChartOCR is better at detecting the borders. Figure 5 (b) and (f) show that the commercial product ThinkCell also suffer from the poor generalization problem. In (f) it fails to detect the actual bar components and mistakenly treat the background ruler line as targets.

Pie Chart Cases: Since not every method provides the result for sector extraction[4], here we compare the final numerical output in Figure 6 and 7. In the case of rare background color or detached sector, ChartOCR can still make precise prediction while other methods are severely disrupted.

Line Chart Cases: In this section we only show our results due to the lack of comparable automatic extraction methods. For simple cases such as the first row of Figure 8, ChartOCR gives pretty good result. For hard examples in the second row, the performance is not very satisfying. The reason is that the QUERY network can not deal with complicated situations where multiple line segments are entangled.

7.3.1 Efficiency Analysis

In terms of the time efficiency of each methods, we show the running time in Table 4 which includes the comparison

[4]RotationRNN directly output the percentage of each sector.

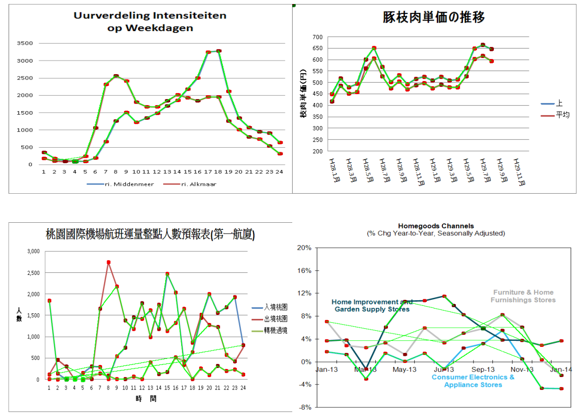


Figure 8: Detection result for line charts. First row: easy samples. Second row: hard samples. Green lines are the predicted lines and red dots are the extracted key points.

Table 4: Average Running Time

Methods	Bar	Pie	Line
ChartOCR	0.206s	0.193s	0.507s
ResNet+Faster-RCNN	0.120s	-	-
Revision	20.032s	5.423s	-
ResNet+RotationRNN	-	0.421s	-

with both the deep learning methods and rule-based methods. As we can see deep methods show a great advantage in time efficiency. For line type ChartOCR takes twice the time of other types, because we added an additional QUERY network as mentioned in Section 3.3.3. The QUERY network does not share parameters with the keypoint detection network. In situations where time efficiency is highly demanded, we can merge these two networks into one single common backbone to reduce the processing time.

8. Conclusion

In this paper we proposed ChartOCR network for precise data value extraction on chart images by combining the rule-based methods and deep neural network based methods. We also introduced a novel benchmark dataset ExcelChart400K that comes with detailed annotations for the chart components. This dataset lays a stepping stone for further research on chart understanding. Our experiments on multiple datasets show that ChartOCR has better performance than both pure rule-based and traditional deep end-to-end methods. Compared with deep models, our chart extractor can be easily generalized for different type of charts such as bar, line and pie charts. Compared with rule-based methods, our approach has much higher precision. For future work, we will expand this work for more chart types.