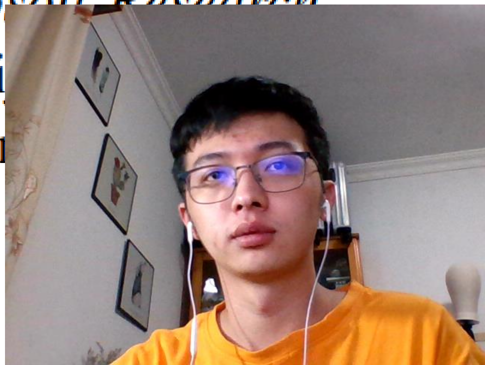
Hybrid Cascade Point Search Network for High Precision Chart Component Detection

Junyu Luo
Sichuan University
Chengdu, China
asbljy@outlook.com

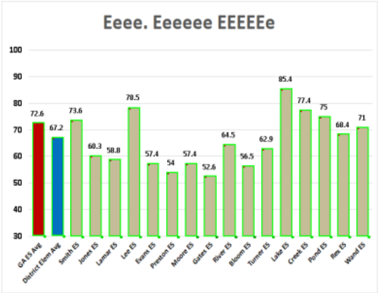
Jinpeng Wang
Microsoft Research
Beijing, china
jinpwa@microsoft.com

Chin-Yew Lin
Microsoft Research
Bei
cyl@

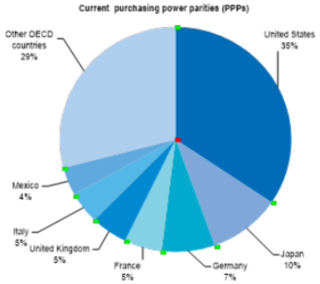


Background

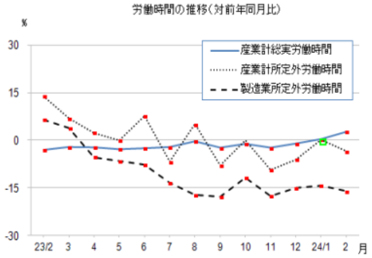
- Chart images are commonly used for data visualization.
- However, during the process they are generally stored in forms of images, bring problems to the automatic data analysis



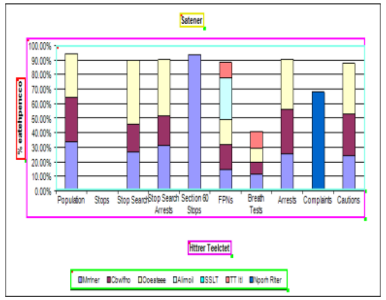
(a) Bar chart



(b) Pie chart



(c) Line chart



(d) Common chart components

Task

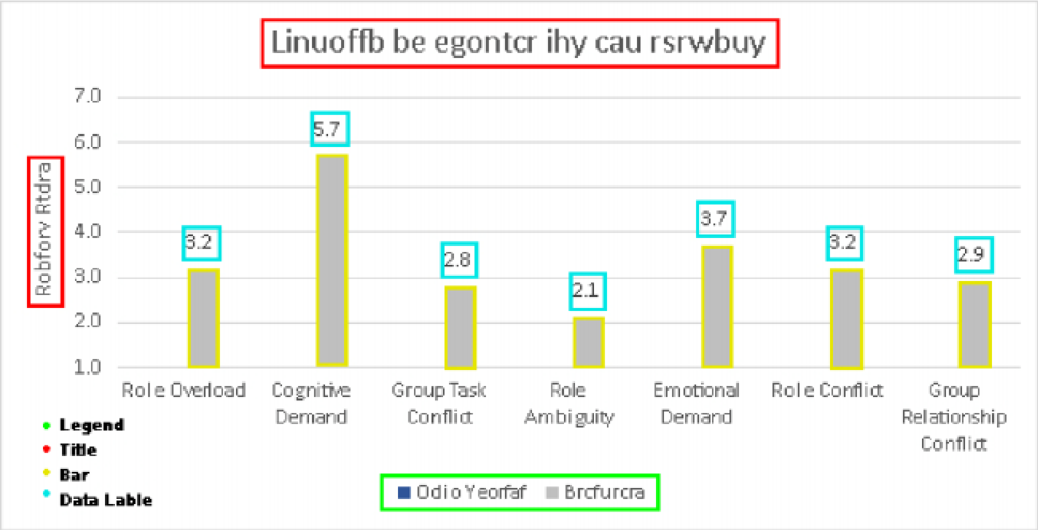
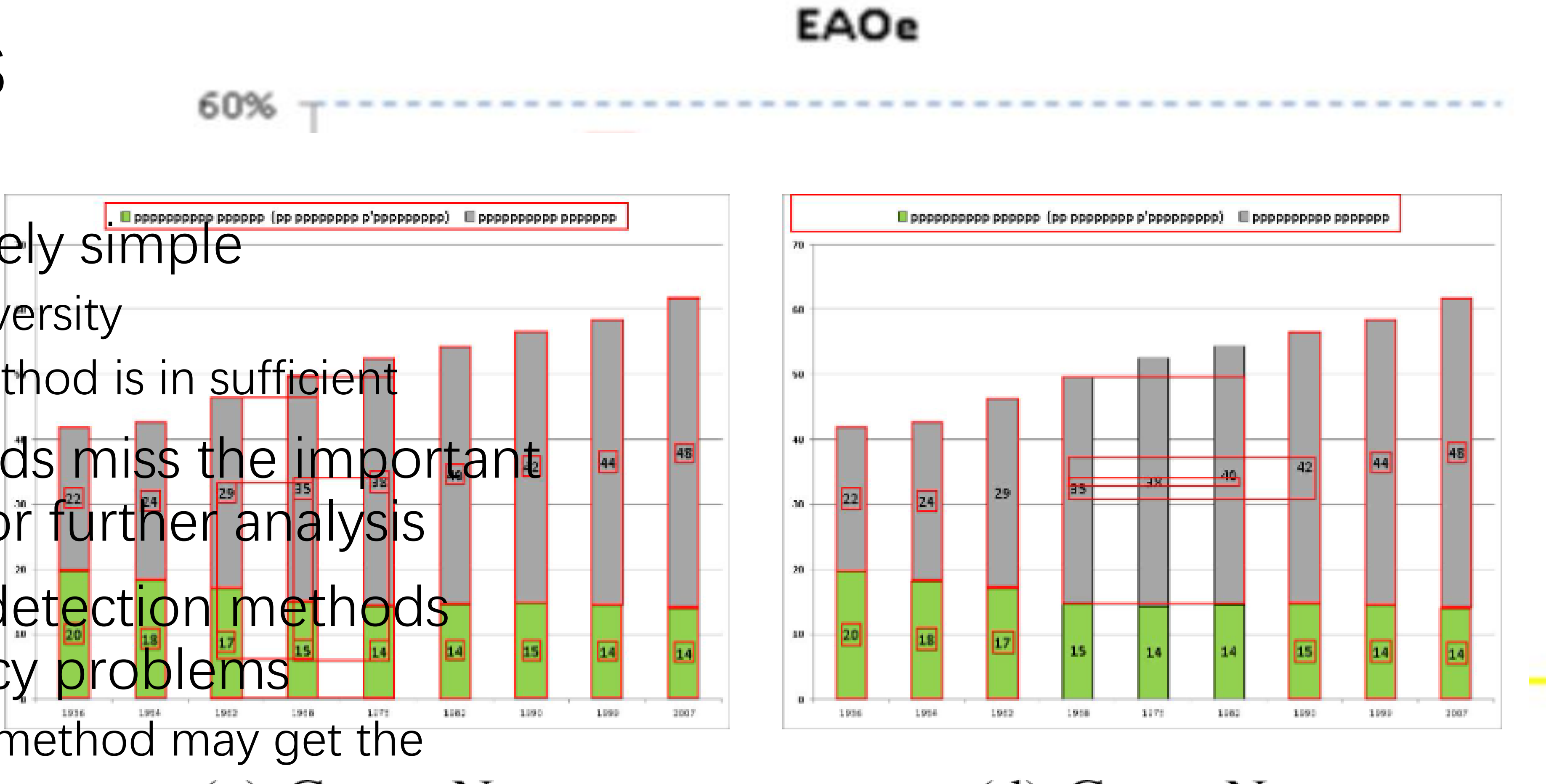


Fig. 4: An example image from ChartDet val set.

- Giving a chart image, we want to mark out each component.

Challenges

- Chart are relatively simple
  - Yet contains diversity
  - Rule-based method is in sufficient
- End2End Methods miss the important middle results for further analysis
- General object detection methods contains accuracy problems
  - Region-based method may get the wrong border
  - Key-point-based method may miss group the key points



(c) CornerNet

(d) Cen

Model

- Key Point Proposal
- Dynamically search the object
- Point Pairing Module

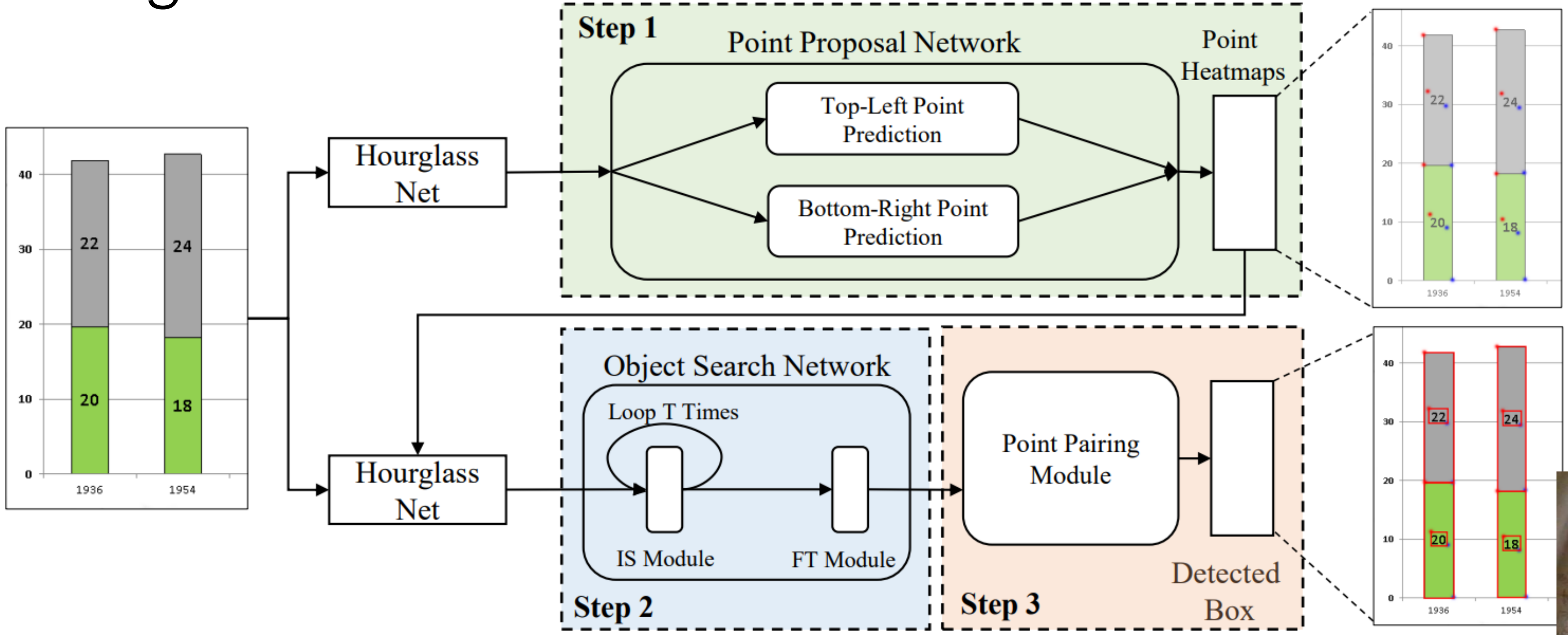
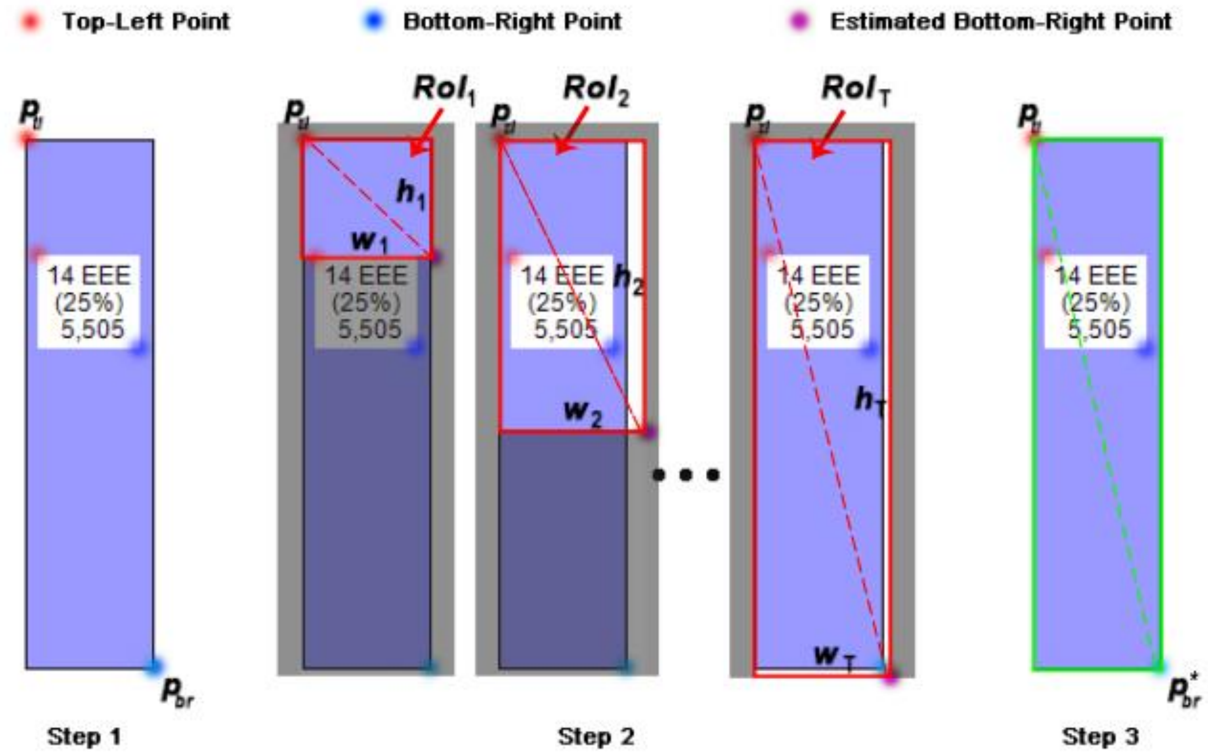
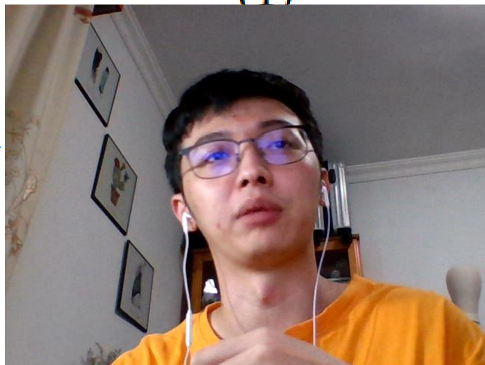


Fig. 3: A diagram of the hybrid cascade pairing network. Point Proposal Network (PPN), Object Search Network (OSN
Point Pairing Module (PPM) work in a cascade order.

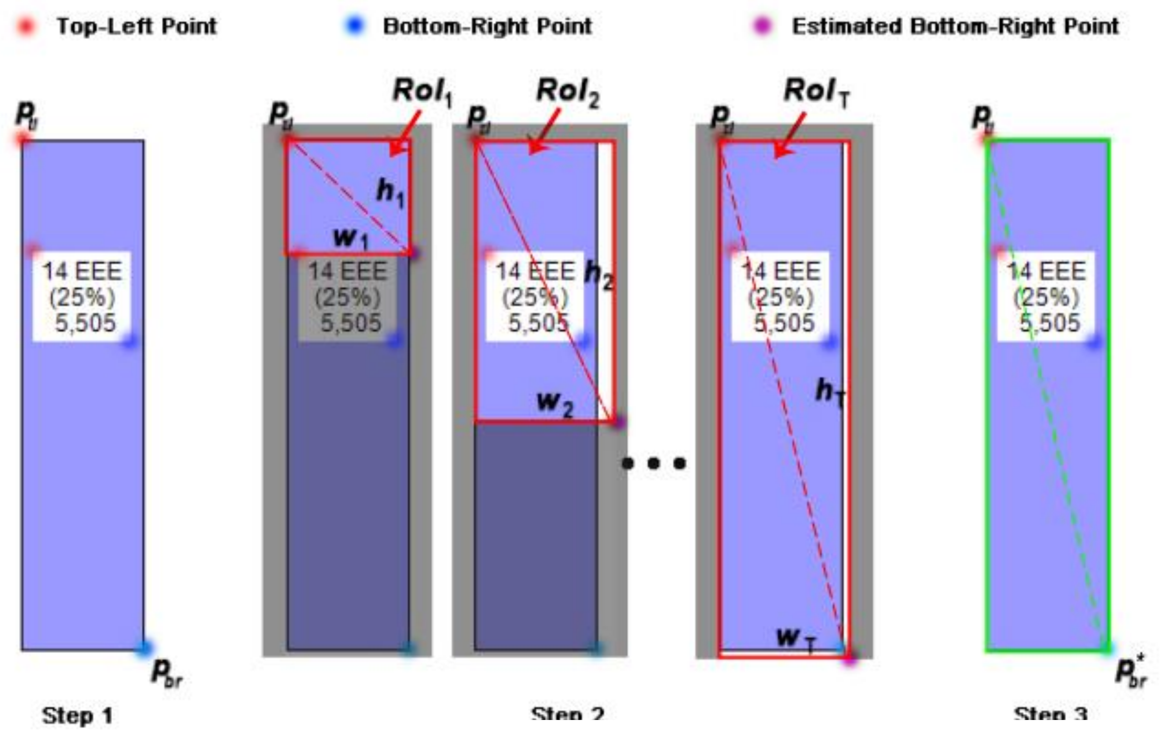
Key Point Proposal



$$L_{heat} = \frac{-1}{N}\sum_{c=1}^{C}\sum_{i=1}^{H}\sum_{j=1}^{W}\begin{cases}(1-\hat{y}_{cij})^{\alpha}log(\hat{y}_{cij}), \\ (1-y_{cij})^{\beta}(\hat{y}_{cij})^{\alpha}log(1-\end{cases} \quad (1)$$



Dynamically search the object - ISM



Iterative Search Module (ISM): ISM searches the main region of the current object in a few iterations.

$$(f_w, f_h) = \mathrm{ISM}\left(RoI_t^{\mathrm{ISM}}, p_{tl}.label, Scale_t\right) \quad (3)$$

$$p_{tl}.w_{t+1} = p_{tl}.w_t \times f_w, \qquad p_{tl}.h_{t+1} = p_{tl}.h_t \times f_h \quad (4)$$

Dynamically search the object - FTM

$$(d_w, d_h) = \mathbf{FTM}(RoI^{\mathrm{FTM}}, p_{tl}.label, Scale_T) \quad (6)$$

$$p_{tl}.w = p_{tl}.w_T \times (1 + d_w) \quad (7)$$

$$p_{tl}.h = p_{tl}.h_T \times (1 + d_h) \quad (8)$$

2) Fine Truing Module (FTM): We add the FTM to further refine the predictions using linear di
ISM. It is different from ISM.

Point Pairing Module

Algorithm 1 Point Pairing Algorithm

Input: top-left point p_{tl}, top-k bottom-right points $\{p_{br}\}$, size of predicted region $(p_{tl}.w, p_{tl}.h)$, threshold of IoU T_{IoU}, threshold of the score T_{score}, ratio of candidate region γ

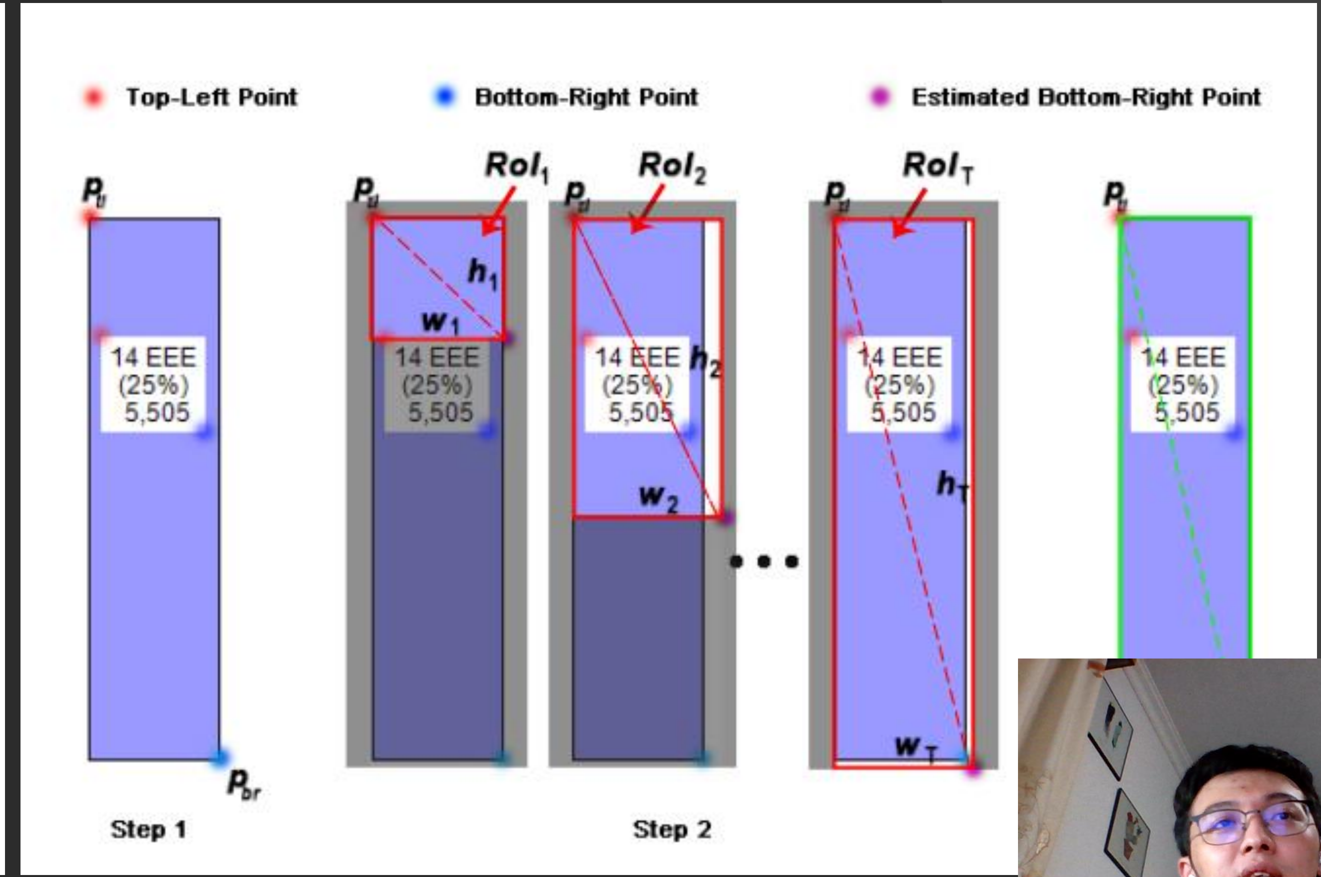
Output: the paired bottom-right point p^*_{br}

initial the max score $S_{max} = 0$

$p_{obj} = (p_{tl}.i + p_{tl}.w,\ p_{tl}.j + p_{tl}.h)$

$\{p_{br}\}'$ = **select** p_{br} **from** $\{p_{br}\}$
    where p_{br} **in** $(p_{obj}.i \pm \gamma \cdot p_{tl}.w,\ p_{tl}.j \pm \gamma \cdot p_{tl}.h)$
    and $p_{br}.label == p_{tl}.label$

for $p_{br} \in \{p_{br}\}'$ **do**
  $S_{IoU} = \text{IoU}(\text{bbox}(p_{tl}, p_{obj}), \text{bbox}(p_{tl}, p_{br}))$
  if $S_{IoU} > T_{IoU}$ **then**
    $S_{cur} = S_{IoU} \times p_{br}.score$
    if $S_{cur} > S_{max}$ and $S_{cur} > T_{score}$ **then**
      $S_{max} = S_{cur}$
      $p^*_{br} = p_{br}$
    end if
  end if
end for



Experiments

Method	AP	$AP_{0.5}$	$AP_{0.75}$	$AP_{0.8}$	$AP_{0.85}$	$AP_{0.9}$	$AP_{0.95}$
Retinanet 101	0.459	0.729	0.497	0.389	0.253	0.110	0.012
Faster-RCNN 101	0.580	0.805	0.664	0.578	0.434	0.233	0.056
Cascade-RCNN 101	0.647	0.831	0.723	0.660	0.552	0.375	0.149
CornerNet	0.646	0.783	0.717	0.674	0.587	0.429	0.225
CenterNet	0.666	0.820	0.742	0.685	0.592	0.429	0.205
HCPN w/o PPM	0.697	**0.870**	0.775	0.714	0.610	0.429	0.222
HCPN	**0.706**	0.868	**0.778**	**0.723**	**0.623**	**0.457**	**0.261**
HCPN (Bar Only)	0.810	0.934	0.873	0.837	0.757	0.635	[illegible]
Revision (Bar Only)	0.330	0.598	0.316	0.217	0.112	0.032	[illegible]

Conclusions

- we presented HCPN, a new framework for precise object detection.
- The experiments proved that our method effectively combining the strengthens of region and point based methods on chart component detection task.

Thanks