

# Hypernym Discovery Search

## Hypernym Discovery 조사

Geonil Yun,

[rjsdlf45@gmail.com](mailto:rjsdlf45@gmail.com)

2021. 12. 14.



# Overview

---

- **Hypernym Discovery**

- Nearest Neighbor
- Term Embedding
- SVD & NN
- SGNS
- CBOW

- **Scene Text**

- PP-YOLO-tiny
- Dataset issue
- Task issue

# Hypernym Discovery

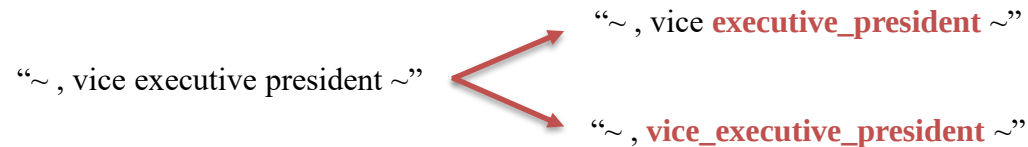
- **NLP\_HZ : Nearest Neighbor Approach**

- Preprocessing

- ① Lowercase all the words for corpus and train/gold/test data.
- ② Concatenate the phrase with underline

executive president → executive\_president

Multiple phrase in one sentence.



- Word Embeddings : using **Google word2vec** with **skip-gram model (300 dimension)**
- Trained word embeddings are used in projection learning and nearest neighbor based method

| System | Language | MAP  | MRR  | P@1  | P@3  | P@5  | P@15 |
|--------|----------|------|------|------|------|------|------|
| PL     | English  | 2.8  | 7.6  | 7.5  | 3.3  | 2.6  | 2.0  |
| NN     | English  | 13.3 | 25.1 | 18.7 | 13.9 | 13.5 | 12.5 |
| NN     | Spanish  | 16.6 | 27.2 | 19.0 | 17.2 | 16.4 | 16.1 |
| NN     | Italian  | 19.3 | 32.4 | 25   | 19.8 | 18.6 | 18.6 |

Table 1: Cross validation results of the two methods on training set(%). PL stands for the projection-learning based system. NN stands for the nearest-neighbor based method.

| Language | MAP   | MRR   | P@1  | P@3   | P@5   | P@15  |
|----------|-------|-------|------|-------|-------|-------|
| English  | 9.37  | 17.29 | 12   | 10.14 | 9.19  | 8.78  |
| Spanish  | 20.04 | 28.27 | 21.4 | 20.95 | 20.39 | 19.38 |
| Italian  | 11.37 | 19.19 | 13.1 | 12.08 | 11.23 | 10.9  |

Table 2: Results on the test data for our submissions(%).

# Hypernym Discovery

- **SJTU-NLP : Neural Hypernym Discovery with Term Embeddings**

- Embeddings

**Glove** toolkit to train the word embeddings (300 dimension)

| Embedding | Model | MAP         | MRR          | P@1         | P@3         | P@5         | P 15        |
|-----------|-------|-------------|--------------|-------------|-------------|-------------|-------------|
| Word      | TEA   | 6.10        | 11.13        | 4.00        | 6.00        | 5.40        | 5.14        |
|           | GRU   | 8.13        | 16.22        | 8.00        | <b>8.00</b> | 6.67        | 6.94        |
|           | LSTM  | 3.95        | 7.52         | 4.00        | 4.33        | 3.97        | 3.97        |
|           | CNN   | 7.32        | 13.33        | <b>8.00</b> | 9.00        | 7.80        | 6.94        |
|           | RCNN  | <b>8.74</b> | <b>12.83</b> | 6.00        | <b>9.67</b> | <b>8.87</b> | <b>9.15</b> |
| Sense     | TEA   | 4.42        | 8.71         | 0.00        | 4.04        | 4.19        | 5.31        |
|           | GRU   | 5.42        | 9.44         | 0.00        | 4.44        | 4.89        | 5.83        |
|           | LSTM  | 5.62        | 9.97         | 4.00        | 4.35        | 5.01        | 6.83        |
|           | CNN   | 6.41        | 10.92        | 2.00        | 5.01        | 5.67        | 6.29        |
|           | RCNN  | 5.79        | 9.24         | 0.00        | 4.71        | 5.29        | 6.43        |

Table 2: Gold standard evaluation on general-purpose subtask.

| Embed | Model | medical      |              |      |              |              |              | music        |              |             |              |              |              |
|-------|-------|--------------|--------------|------|--------------|--------------|--------------|--------------|--------------|-------------|--------------|--------------|--------------|
|       |       | MAP          | MRR          | P@1  | P@3          | P@5          | P 15         | MAP          | MRR          | P@1         | P@3          | P@5          | P 15         |
| Word  | TEA   | 8.91         | 16.77        | 0.00 | 8.79         | 9.41         | 9.39         | 7.11         | 14.32        | 0.00        | 10.01        | 10.77        | 9.21         |
|       | GRU   | 13.27        | 21.89        | 0.00 | 13.33        | 14.89        | 14.06        | 15.20        | 20.33        | 0.00        | 17.78        | 18.67        | 15.45        |
|       | LSTM  | 11.49        | 21.11        | 0.00 | <b>17.78</b> | 12.22        | 11.83        | 14.08        | 20.77        | 0.07        | 13.33        | 16.00        | 15.00        |
|       | CNN   | <b>18.31</b> | <b>24.52</b> | 0.00 | 15.56        | <b>20.44</b> | <b>20.00</b> | <b>17.58</b> | <b>27.15</b> | 0.00        | <b>20.00</b> | <b>20.00</b> | <b>16.04</b> |
|       | RCNN  | 16.78        | 23.40        | 0.00 | 13.33        | 13.00        | 14.50        | 13.60        | 21.67        | 0.07        | 13.33        | 14.67        | 13.08        |
| Sense | TEA   | 2.01         | 4.77         | 0.00 | 2.91         | 2.77         | 3.21         | 2.59         | 5.28         | 0.00        | 2.12         | 3.01         | 2.93         |
|       | GRU   | 4.88         | 9.11         | 0.00 | 6.67         | 6.42         | 6.91         | 5.32         | 10.74        | <b>2.00</b> | 4.44         | 5.33         | 4.95         |
|       | LSTM  | 5.10         | 10.22        | 0.00 | 6.67         | 6.12         | 6.94         | 4.39         | 10.21        | 0.00        | 8.89         | 5.33         | 3.61         |
|       | CNN   | 4.15         | 7.84         | 0.00 | 4.44         | 6.09         | 6.42         | 4.75         | 9.61         | 0.00        | 6.67         | 6.67         | 4.43         |
|       | RCNN  | 4.63         | 9.84         | 0.00 | 6.67         | 6.89         | 6.43         | 4.73         | 8.56         | 0.00        | 4.44         | 6.22         | 4.94         |

Table 3: Gold standard evaluation on domain-specific subtask. “Embed” is short for “Embedding”.

# Hypernym Discovery

- The Effectiveness of Simple Hybrid System for Hypernym Discovery

- Hybrid system – Pattern based, Distributional modeling with Nearest
- Embeddings

FastText with 100 dimension

| Model Variant                 | General      |              |              | Medical      |              |              | Music        |              |              |
|-------------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
|                               | MAP          | MRR          | P@5          | MAP          | MRR          | P@5          | MAP          | MRR          | P@5          |
| Count Hearst Patterns         | 4.60         | 11.70        | 4.10         | 14.99        | 43.18        | 13.70        | 7.65         | 26.14        | 6.76         |
| SVD Hearst Patterns           | 6.19         | 15.12        | 5.65         | 15.80        | 47.01        | 14.53        | 8.89         | 29.63        | 8.05         |
| Hypernyms of Nearest Neighbor | 9.85         | 24.56        | 8.76         | 29.57        | 48.18        | 34.10        | 38.65        | 72.77        | 38.45        |
| Hybrid of Raw Count & NN      | 14.82        | 32.61        | 13.80        | 35.29        | 63.59        | 38.73        | 28.22        | 61.26        | 30.67        |
| Hybrid of SVD & NN            | <b>15.97</b> | <b>34.07</b> | <b>15.00</b> | <b>37.85</b> | <b>64.47</b> | <b>40.19</b> | <b>54.62</b> | <b>77.24</b> | <b>55.08</b> |

Table 1: Comparison of model variants on all three sub-tasks of SemEval 2018 Task 9.

| Model                                      | General      |              |              | Medical      |              |              | Music        |              |              |
|--------------------------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
|                                            | MAP          | MRR          | P@5          | MAP          | MRR          | P@5          | MAP          | MRR          | P@5          |
| Hybrid of SVD & NN(Our Model)              | 15.97        | 34.07        | 15.00        | <b>37.85</b> | <b>64.47</b> | <b>40.19</b> | <b>54.62</b> | <b>77.24</b> | <b>55.08</b> |
| CRIM (Bernier-Colborne and Barriere, 2018) | <b>19.78</b> | <b>36.10</b> | <b>19.03</b> | 34.05        | 54.64        | 36.77        | 40.97        | 60.93        | 41.31        |
| vTE* (Espinosa-Anke et al., 2016)          | 10.60        | 23.83        | 9.91         | 18.84        | 41.07        | 20.71        | 12.99        | 39.36        | 12.41        |
| 300-sparsans (Berend et al., 2018)         | 8.95         | 19.44        | 8.63         | 20.75        | 40.60        | 21.43        | 29.54        | 46.43        | 28.86        |
| Most Frequent Hypernym*                    | 8.77         | 21.39        | 7.81         | 28.93        | 35.80        | 34.20        | 33.32        | 51.48        | 35.76        |

Table 2: Results for all three English sub-tasks in SemEval 2018 Task 9. Baselines are marked with \*, state-of-the-art is marked in bold.

# Hypernym Discovery

---

- **ADAPT : Skip-Gram Word Embeddings for Unsupervised Hypernym Discovery**
  - Skip-Gram with Negative Sampling algorithm - 300 dimension, 20 negative samples, 5 window size

300 dimension, 20 negative samples, 5 window size

| Domain  | Approach | MAP  | MRR   | P@1   | P@3   | P@5  | P@15 |
|---------|----------|------|-------|-------|-------|------|------|
| medical | SGNS     | 8.13 | 20.56 | 13.20 | 10.80 | 8.32 | 6.33 |
| medical | HV       | 4.40 | 13.05 | 10.60 | 5.60  | 4.27 | 3.10 |
| music   | SGNS     | 1.88 | 5.34  | 4.00  | 2.40  | 1.89 | 1.35 |
| music   | HV       | 1.79 | 5.39  | 5.00  | 2.07  | 1.62 | 1.28 |

Table 1: Automatic evaluation results for the submitted system (SGNS) and a Hypervec variant (HV).

# Hypernym Discovery

- **UMDuluth-CS8761 : Hypernym Discovery using Hearst Patterns, Co-occurrence frequencies and Word Embeddings**

- **Embeddings**

CBOW(Continuous Bag-of-Words) model.

10 window size

300 embedding dimension size

|      | Cooc | Hearst | Phi  | Is-A | Merged |
|------|------|--------|------|------|--------|
| MRR  | .103 | .020   | .025 | .165 | .188   |
| MAP  | .050 | .008   | .012 | .071 | .080   |
| P@1  | .061 | .013   | .012 | .140 | .152   |
| P@3  | .055 | .008   | .012 | .076 | .087   |
| P@5  | .048 | .008   | .012 | .066 | .075   |
| P@15 | .047 | .007   | .011 | .062 | .070   |

Table 1: Test Data 1A English - Concept Scores

|      | Cooc | Hearst | Phi  | Is-A | Merged |
|------|------|--------|------|------|--------|
| MRR  | .000 | .000   | .008 | .090 | .099   |
| MAP  | .000 | .000   | .003 | .036 | .037   |
| P@1  | .000 | .000   | .004 | .069 | .081   |
| P@3  | .000 | .000   | .003 | .041 | .045   |
| P@5  | .000 | .000   | .003 | .035 | .036   |
| P@15 | .000 | .000   | .003 | .031 | .030   |

Table 2: Test Data 1A English - Entity Scores

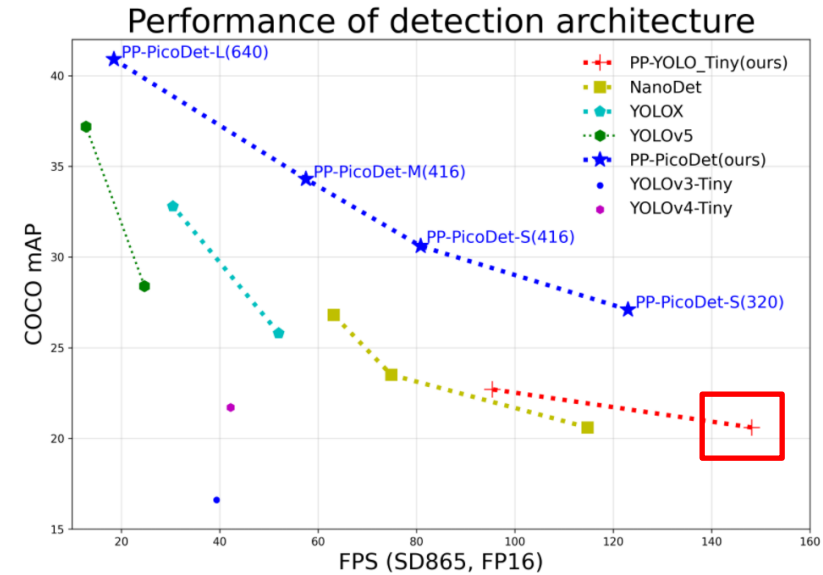
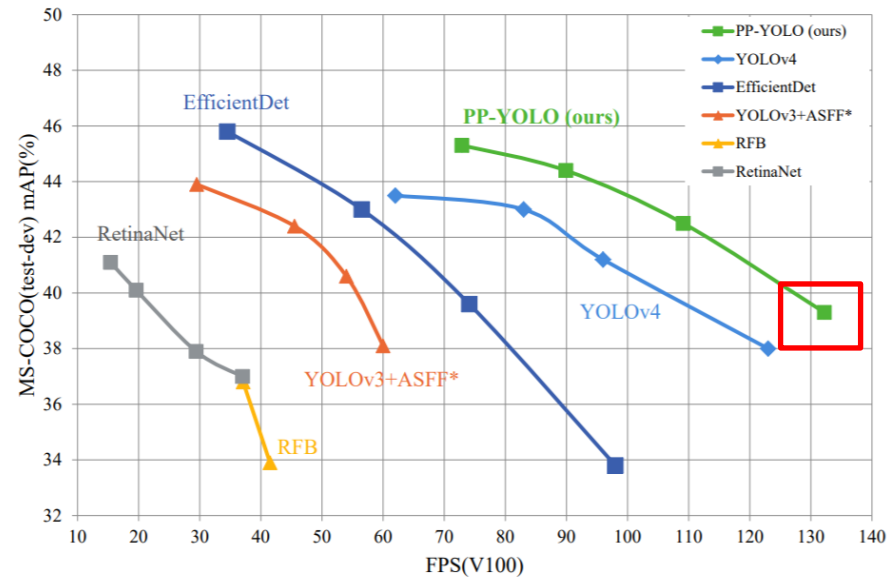
# Hypernym Discovery

- Comparison

| Model            | Embedding   | dimension  | Window size | MAP          | MRR          | P@1          |
|------------------|-------------|------------|-------------|--------------|--------------|--------------|
| NLP_HZ           | SGNS        | 300        | N/A         | 13.3         | 25.1         | 18.70        |
| SJTU-NLP         | Glove       | 300        | N/A         | 8.74         | 12.83        | 6.00         |
| SVD&NN           | FastText    | 100        | N/A         | 15.97        | 34.07        | 15.00        |
| ADAPT            | SGNS        | 300        | 5           | N/A          | N/A          | N/A          |
| UMDuluth-CS8761  | CBOW        | 300        | 10          | 8.00         | 18.80        | 15.2         |
| CRIM             | SGNS        | N/A        | N/A         | 19.78        | 36.10        | <b>29.67</b> |
| RMM              | SGNS        | 200        | N/A         | <b>27.12</b> | <b>39.07</b> | 23.41        |
| <b>Best CRIM</b> | <b>SGNS</b> | <b>300</b> | <b>2</b>    | <b>8.12</b>  | <b>18.26</b> | <b>15.06</b> |

# Scene Text

- PP-YOLO tiny



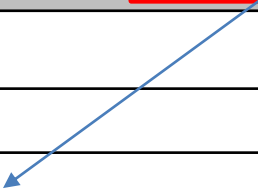
- PP-YOLO tiny for word region detection
- Compared to YOLOv4-Tiny, FPS is much faster

# Scene Text

- PP-YOLO tiny at COCO-2017

| Model   | backbone   | input shape | Box AP val | FPS on V100 |
|---------|------------|-------------|------------|-------------|
| PP-YOLO | ResNet50vd | 608         | 44.8       | 72.9        |
|         |            | 512         | 43.9       | 89.9        |
|         |            | 416         | 42.1       | 109.1       |
|         |            | 320         | 38.9       | 132.2       |

| Model        | Model Size | input shape | Box AP val | FPS on Kirin 990 4xCores (Cortex-A55@1.86 GHz) |
|--------------|------------|-------------|------------|------------------------------------------------|
| PP-YOLO tiny | 4.2MB      | 320         | 20.6       | 92.3                                           |
|              | 4.2MB      | 416         | 22.7       | 65.4                                           |



|       |                                                                                                                                             |
|-------|---------------------------------------------------------------------------------------------------------------------------------------------|
| CPU   | <ul style="list-style-type: none"><li>• 2x Cortex-A76@2.86 GHz</li><li>• 2x Cortex-A76@2.09 GHz</li><li>• 4x Cortex-A55@1.86 GHz</li></ul>  |
| GPU   | <ul style="list-style-type: none"><li>• 16-core Mali-G76</li></ul>                                                                          |
| AI    | <ul style="list-style-type: none"><li>• HUAWEI Da Vinci Architecture</li><li>• 1x Ascend Lite + 1x Ascend Tiny</li><li>• HiAI 2.0</li></ul> |
| Modem | <ul style="list-style-type: none"><li>• 2G/3G/4G</li></ul>                                                                                  |

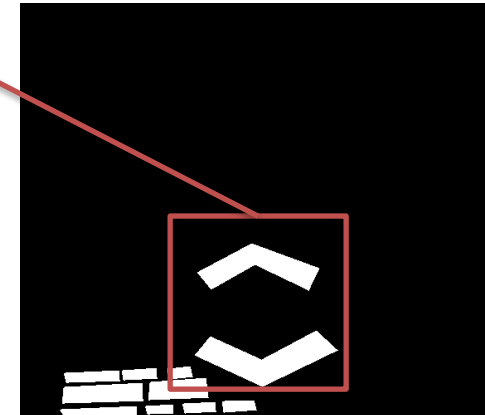
# Scene Text

- Dataset Issue
  - Support expression

| Expression           | Explanation                                                                               |
|----------------------|-------------------------------------------------------------------------------------------|
| $x_1, y_1, x_2, y_2$ | $(x_1, y_1)$ is the top left coordinate, $(x_2, y_2)$ is the bottom right coordinate.     |
| $x_1, y_1, w, h$     | $(x_1, y_1)$ is the top left coordinate, $w$ is width of object, $h$ is height of object. |
| $x_c, y_c, w, h$     | $(x_c, y_c)$ is center of object, $w$ is width of object, $h$ is height of object.        |

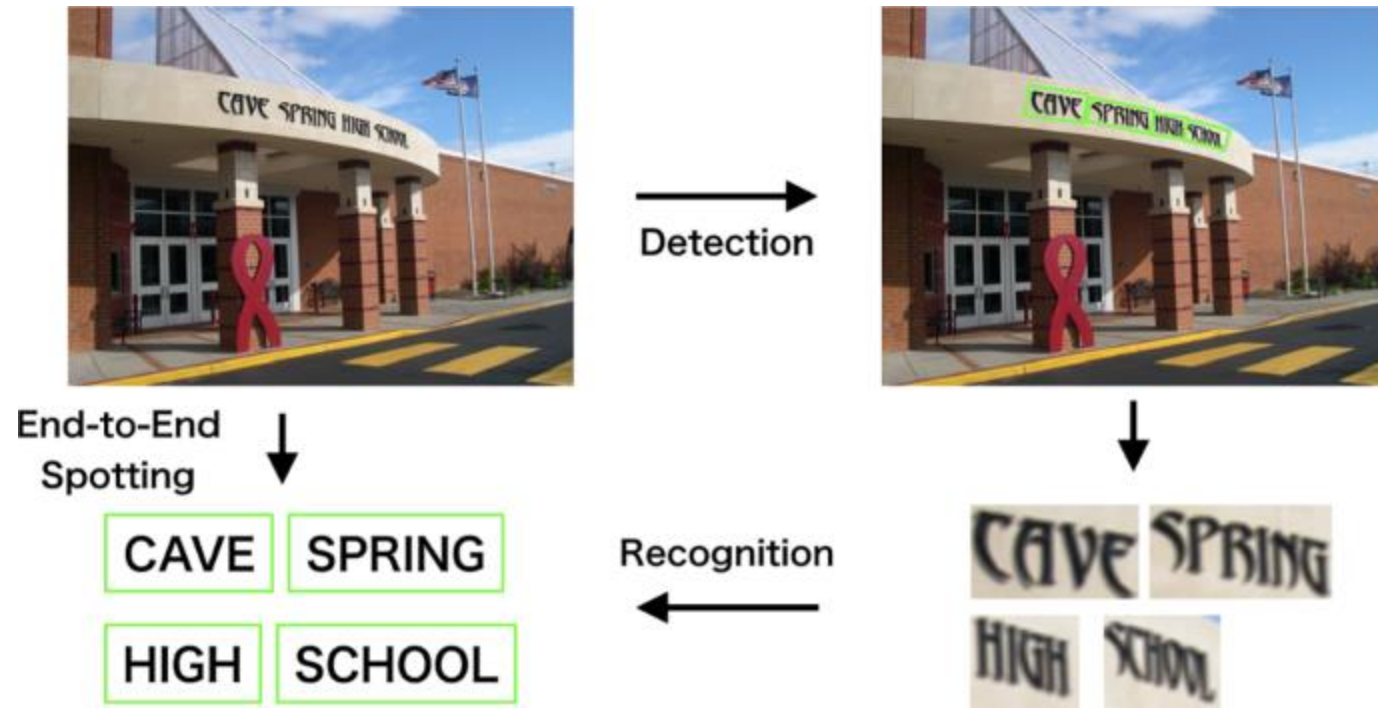
x: [[214 280 362 349 284 231]], y: [[325 290 320 347 316 346]],  
x: [[230 292 358 384 293 211]], y: [[402 430 395 419 463 425]],  
x: [[ 53 120 121 56]], y: [[446 443 456 458]], ornt: [u'h'], transcr  
x: [[123 165 166 125]], y: [[443 440 453 455]], ornt: [u'h'], transc  
x: [[168 179 179 167]], y: [[439 439 452 453]], ornt: [u'#'], transc  
x: [[178 206 209 180]], y: [[441 438 451 454]], ornt: [u'h'], transc  
x: [[ 53 148 149 51]], y: [[463 458 480 483]], ornt: [u'm'], transc  
x: [[157 225 228 155]], y: [[457 452 476 479]], ornt: [u'm'], trans  
x: [[ 48 141 142 53]], y: [[490 486 499 498]], ornt: [u'm'], transc  
x: [[151 185 185 154]], y: [[486 484 495 496]], ornt: [u'h'], transc  
x: [[196 236 237 199]], y: [[483 481 493 495]], ornt: [u'h'], transc  
x: [[244 283 286 246]], y: [[480 479 491 494]], ornt: [u'm'], trans

Polygon coordinates.



# Scene Text

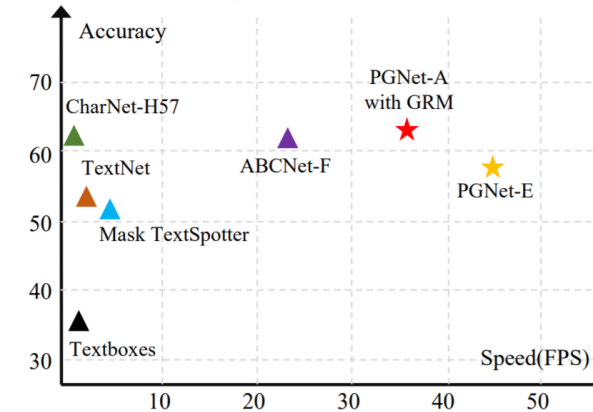
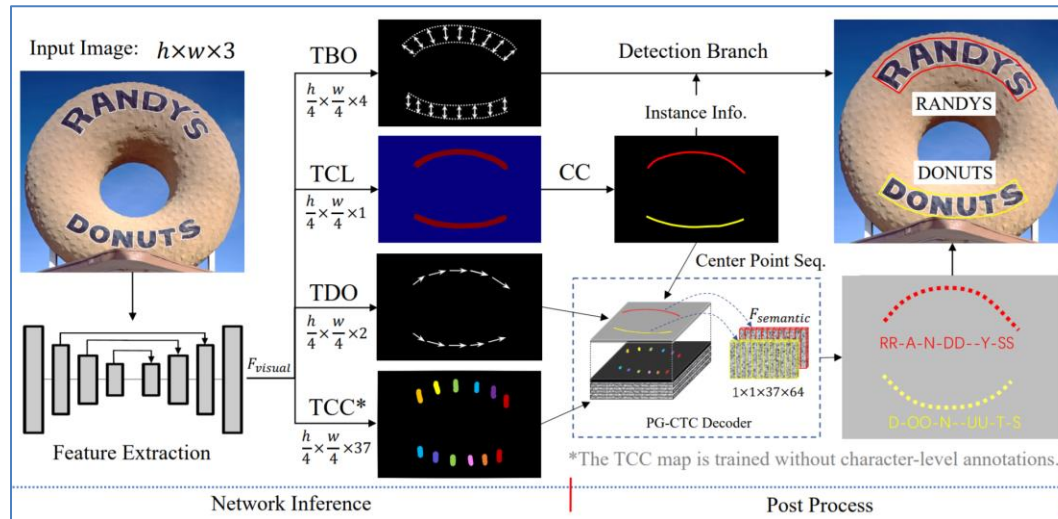
- Task issue



- TextFuseNet, DB-ResNet are scene text detection models.

# Scene Text

- Task issue
  - Text Spotting (PGNet-E)



- End-to-end text spotting sota model : PGNet



---

Thank you