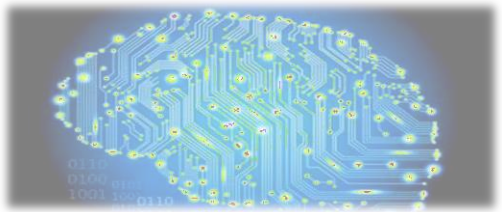
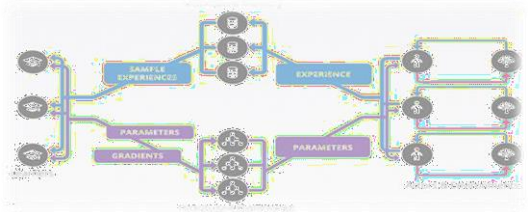
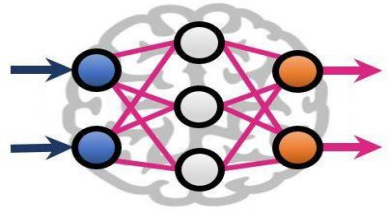
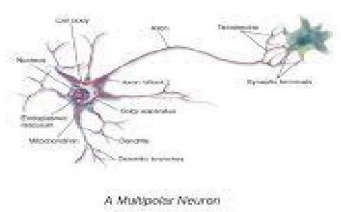
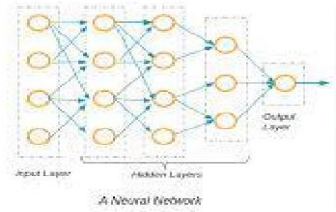





Hypernym discovery : CRIM experiments(7)

CRIM 구현 및 실험

Geonil Yun,
rjsdlf45@gmail.com
2021. 11. 09.

Overview

- **CRIM – Hypernym discovery**
  - Test with training data.
  - Check (Hypernym-Hyponym) Embedding distance.
  - Each layer output compare

- **Scene Text**
  - Dataset – Total-Text, ICDAR-2015

CRIM – Hypernym Discovery

- **Test with training data**

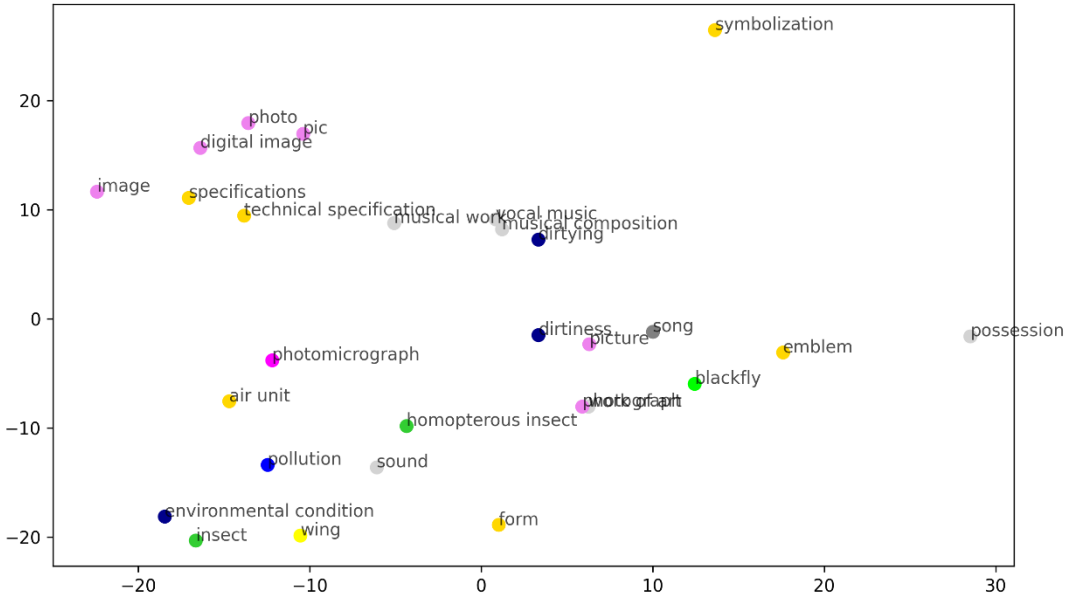
  - **Evaluation**

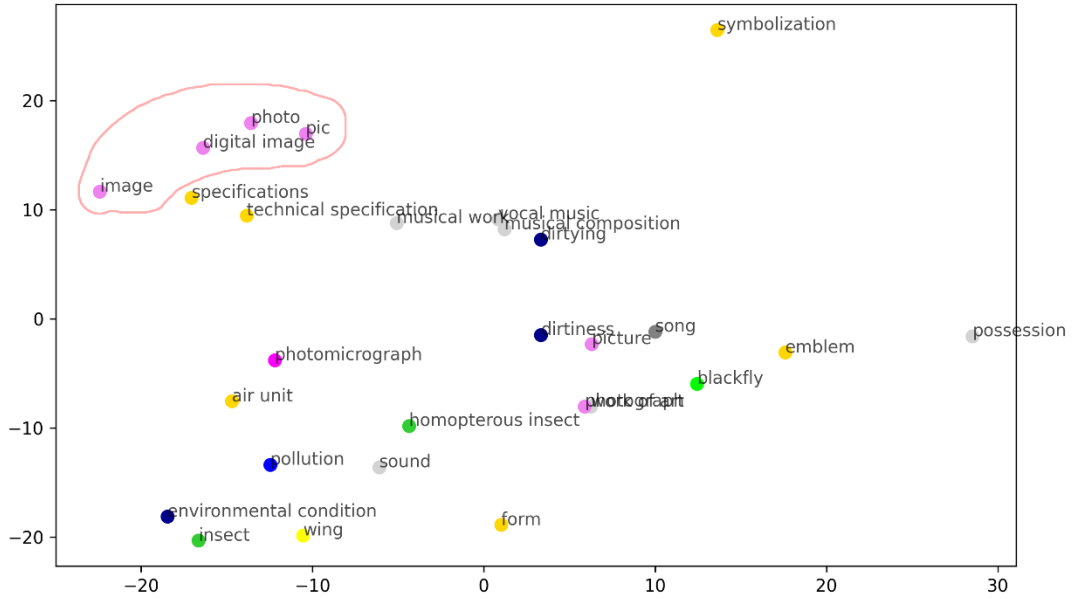
Model	MAP	MRR	P@1
CRIM	***19.54***	***35.94***	***29.67***
Hybrid-before	4.94	11.62	9.26
Hybrid-win2-50epoch	8.12	18.26	15.06
Supervised-win2-50epoch	7.95	19.04	16.06
Unsupervised – win2 - 60epoch	4.50	9.92	7.33
Train data	**40.52**	**57.74**	**45.73**

    - Checking the results using the training data shows some performance.

CRIM – Hypernym Discovery

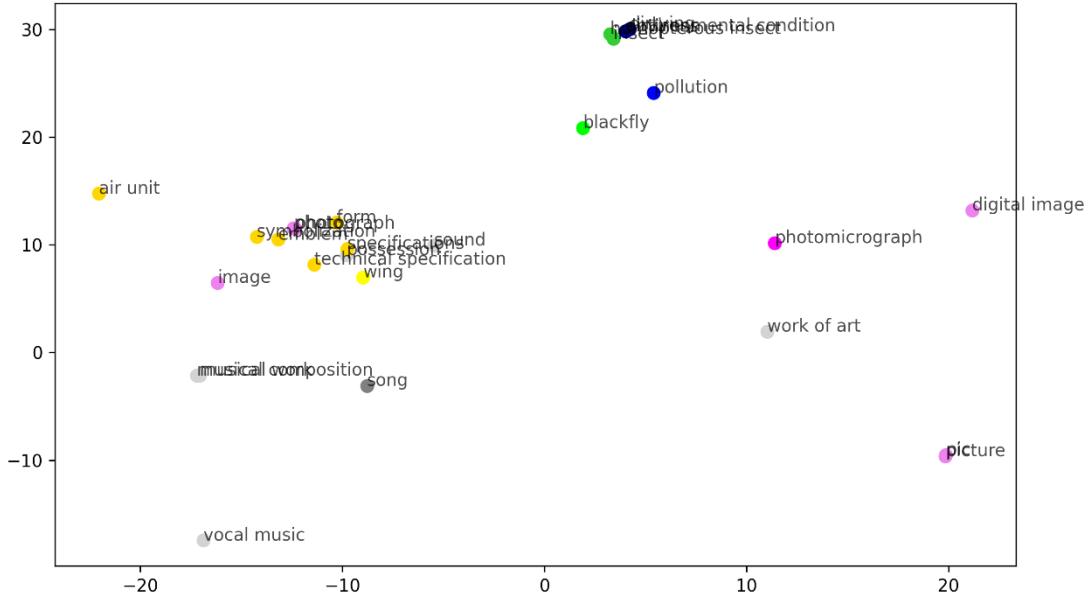
- **Embedding visualization**
  - ➢ **Embedding unsupervised for training data**

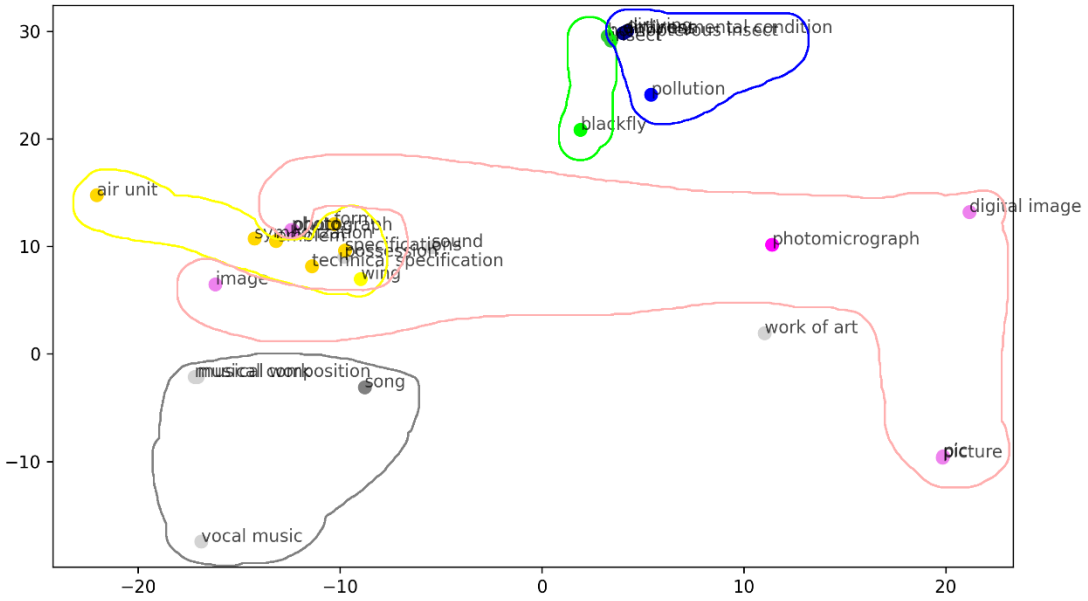




CRIM – Hypernym Discovery

- **Embedding visualization**
  - ➢ **Embedding supervised for training data**





CRIM – Hypernym Discovery

- **Compare output of each layer**
  - **Query(“blackfly”), hypernym(“insect”) : Training data(correct)**

$$P_i = \left(\phi_i \cdot e_q\right)^T, i \in \{1, \dots, 24\}$$

Projection Matrix(ϕ)	Embedding (e_q)	Projected query (P)
[[[0.1026, …, -0.0044, 0.0469], …],…]	[[0.2139, …, 0.1675, 0.8245]]	[[[-0.7345], …, [0.5673]], …]
shape : (24, 100, 100) (kernel , dim, dim)	shape : (1, 100) (batch, dim)	shape : (24, 100, 1) (kernel, dim, batch)

- Check dot product with **projection matrix[0][0]** and **embeddings** : **-0.7345**

CRIM – Hypernym Discovery

- **Compare output of each layer**
  - **Query(“blackfly”), hypernym(“insect”) : Training data(correct)**

$$s = P \cdot e_h$$

Projected query (P)

[[[-0.7345], …, [0.5673]], …]

shape : (24, 100, 1) (kernel, dim, batch)

↓ **reshape**

shape : (1, 24, 100) (batch, kernel, dim)

Embedding (e_h)

[[-0.0273, …, 0.4245, -0.1077]]

shape : (1, 100) (batch, dim)

↓ **reshape**

shape : (1, 100, 1) (batch, dim, 1)

Score (P)

[[[-0.5638], …, [-0.6596]]]

shape : (1, 24, 1) (batch, kernel, 1)

- Check dot product with **projected query[0][0]** and **embeddings** : **-0.5638**

CRIM – Hypernym Discovery

- **Compare output of each layer**

  - **Query(“blackfly”), hypernym(“insect”) : Training data(correct)**

$$y = \sigma(W \cdot s + b)$$

Weight	Score (s)	Bias
[[0.0046, 0.1145, 0.0756, -0.0050, -0.1738, -0.1368, 0.1907, 0.0827, 0.0561, 0.0326, 0.1108, 0.0940, -0.0572, -0.2038, -0.0188, -0.0582, 0.1875, -0.0253, 0.1804, 0.0033, 0.1115, 0.1209, -0.1907, -0.0608]]	[-0.5638, 5.5314, 7.2603, 1.2169, -1.7639, -6.0869, 7.8552, 15.9347, 17.3474, 10.5739, 8.7186, 14.8626, -10.0919, -4.8536, -14.9141, -14.6340, 5.1422, -0.2579, 6.0947, -19.0973, 4.9274, 5.3083, -2.8579, -0.6596]	[0.1852]

$$\sigma(15.4787) = 1$$

- Check for correct hypernym
- likelihood 1

CRIM – Hypernym Discovery

- **Compare output of each layer**

  - **Query(“blackfly”), hypernym(“hate”) : Training data(incorrect)**

$$y = \sigma(W \cdot s + b)$$

Weight	Score (s)	Bias
[[0.0046, 0.1145, 0.0756, -0.0050, -0.1738, -0.1368, 0.1907, 0.0827, 0.0561, 0.0326, 0.1108, 0.0940, -0.0572, -0.2038, -0.0188, -0.0582, 0.1875, -0.0253, 0.1804, 0.0033, 0.1115, 0.1209, -0.1907, -0.0608]]	**[-4.1955, -7.1086, -4.1250, 13.8438, 1.4566, 4.5636, -9.0193, -2.1757, -3.2347, 0.1974, -2.4694, -0.9075, 3.7260, 3.6365, 21.5789, -0.9208, -2.8388, 5.6795, -1.3012, -48.2848, -2.1753, -5.0976, 2.4201, 3.9550]**	[0.1852]

$$\sigma(-8.2769) = 0.0003$$

  - Check for incorrect hypernym
  - likelihood 0.0003

Issue & Future

- **last week**
  - RMM(Recurrent Mapping Model)
  - Use pre-trained embeddings.
  - More experiments.
  - **Test with training data.**
  - **Check embedding distance.**
  - **Compare output of each layer.**

- **Future works**
  - RMM(Recurrent Mapping Model)
  - Use pre-trained embeddings.
  - More experiments.

Scene Text

- **Dataset – Total-text**

- **1255 training images, 500 test images**
- **Annotated with polygons in word-level**

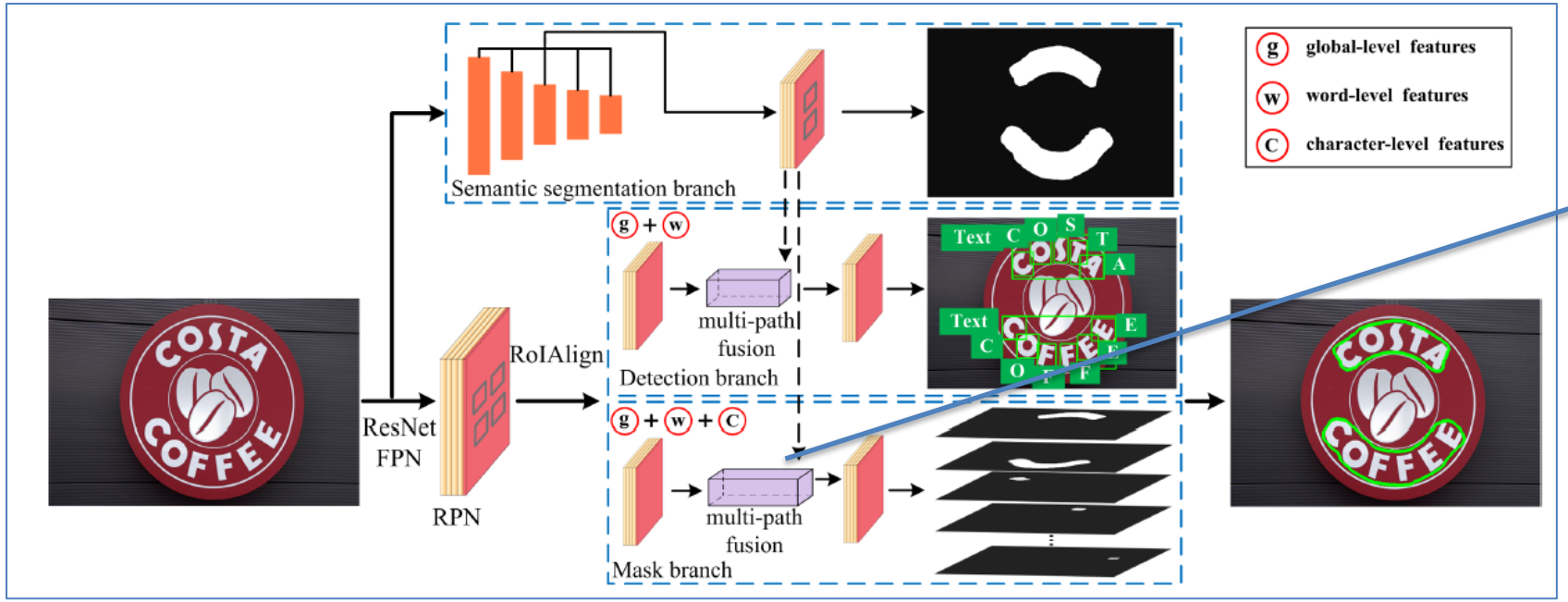
Scene Text

- **Dataset – ICDAR2015**

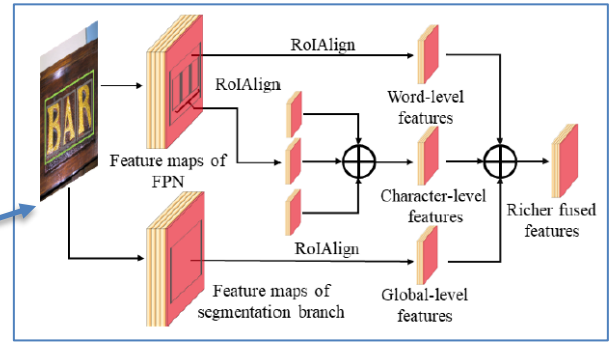
- **1000 training images, 500 test images**
- **Annotated with quadrangles in word-level**

Scene Text

- **State-of-the-art (TextFuseNet)**



TextFuseNet Framework



multi-path fusion architecture

Scene Text

- **State-of-the-art (TextFuseNet)**

Method	ICDAR2013				ICDAR2015				Total-Text				CTW-1500			
	R	P	F	FPS	R	P	F	FPS	R	P	F	FPS	R	P	F	FPS
CTPN [Tian *et al.*, 2016]	83.0	93.0	88.0	7.14	52.0	74.0	61.0	7.1	-	-	-	-	-	-	-	-
SegLink [Shi *et al.*, 2017]	83.0	87.7	85.3	**20.6**	76.8	73.1	75.0	-	-	-	-	-	-	-	-	-
TextSnake [Long *et al.*, 2018]	-	-	-	-	80.4	84.9	82.6	1.1	74.5	82.7	78.4	-	85.3	67.9	75.6	-
TextBoxes++* [Liao *et al.*, 2018a]	86.0	92.0	89.0	-	78.5	87.8	82.9	2.3	-	-	-	-	-	-	-	-
RRD* [Liao *et al.*, 2018b]	86.0	92.0	89.0	-	80.0	88.0	83.8	-	-	-	-	-	-	-	-	-
PixelLink* [Deng *et al.*, 2018]	87.5	88.6	88.1	-	82.0	85.5	83.7	7.3	-	-	-	-	-	-	-	-
Mask TextSpotter [Lyu *et al.*, 2018]	88.6	95.0	91.7	4.6	81.0	91.6	86.0	4.8	55.0	69.0	61.3	-	-	-	-	-
TextField [Xu *et al.*, 2019]	-	-	-	-	80.5	84.3	82.4	6.0	79.9	81.2	80.6	-	79.8	83.0	81.4	-
CTD [Liu *et al.*, 2019a]	-	-	-	-	-	-	-	-	71.0	74.0	73.0	-	69.8	74.3	73.4	13.3
CSE [Liu *et al.*, 2019b]	-	-	-	-	-	-	-	-	79.1	81.4	80.2	2.4	76.0	81.1	78.4	2.6
MSR* [Xue *et al.*, 2019]	-	-	-	-	78.4	86.6	82.3	4.3	74.8	83.8	79.0	-	78.3	85.0	81.5	-
PSENet [Wang *et al.*, 2019a]	-	-	-	-	84.5	86.9	85.7	1.6	78.0	84.0	80.9	3.9	79.7	84.8	82.2	3.9
PAN [Wang *et al.*, 2019b]	-	-	-	-	84.0	81.9	82.9	**26.1**	**89.3**	81.0	85.0	**39.6**	81.2	86.4	83.7	**39.8**
ATRR [Wang *et al.*, 2019c]	89.7	93.7	91.7	-	86.0	89.2	87.6	-	76.2	80.9	78.5	-	80.2	80.1	80.1	-
SAE [Tian *et al.*, 2019]	-	-	-	-	85.0	88.3	86.6	-	-	-	-	-	77.8	82.7	80.1	-
SPCNet [Xie *et al.*, 2019]	90.5	93.8	92.1	-	85.8	88.7	87.2	-	82.8	83.0	82.9	-	-	-	-	-
LOMO* [Zhang *et al.*, 2019]	-	-	-	-	87.6	87.8	87.7	-	79.3	87.6	83.3	-	76.5	85.7	80.8	-
DB (ResNet-50) [Liao *et al.*, 2020]	-	-	-	-	83.2	91.8	87.3	12.0	82.5	87.1	84.7	32.0	80.2	86.9	83.4	22.0
Our (ResNet-50)	89.5	95.1	92.2	7.7	88.9	91.3	90.1	8.3	83.2	87.5	85.3	7.1	85.0	85.8	85.4	7.3
Our (ResNet-101)	**92.3**	**96.5**	**94.3**	4.0	**89.7**	**94.7**	**92.1**	4.1	85.3	**89.0**	**87.1**	3.3	**85.4**	**87.8**	**86.6**	3.7

Thank you