meeting 8/25

應名宥

Table of contents

01

Discriminator

SOO test result

discriminator

a city street line with brick buildings and trees.

random

maximize

```
population : 100
offspring  : 100
generation : 1000
```

discriminator

a city street line with brick buildings and trees.

random

minimize

```
population : 100
offspring  : 100
generation : 1000
```

02

Problem

rethinking of word-label structure

1. **context**

a city street line with brick buildings and trees.

/ \

brick buildings

2. **top k important word**

a dog and a cat are sitting on an orange sofa

Afghan Hound
Golden Retriever
Airedale Terrier
...

Persian Cat
Donskoy
Oriental Bicolour
...

03

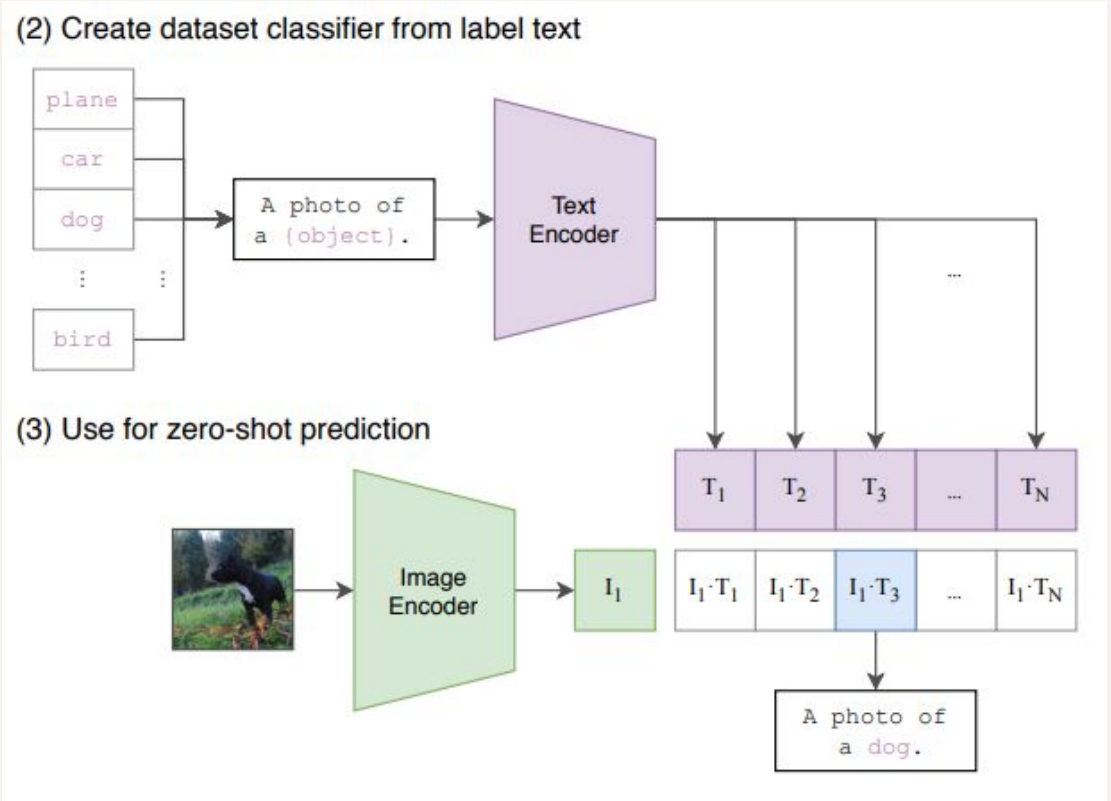
text-label similarity

test text-label similarity effectiveness

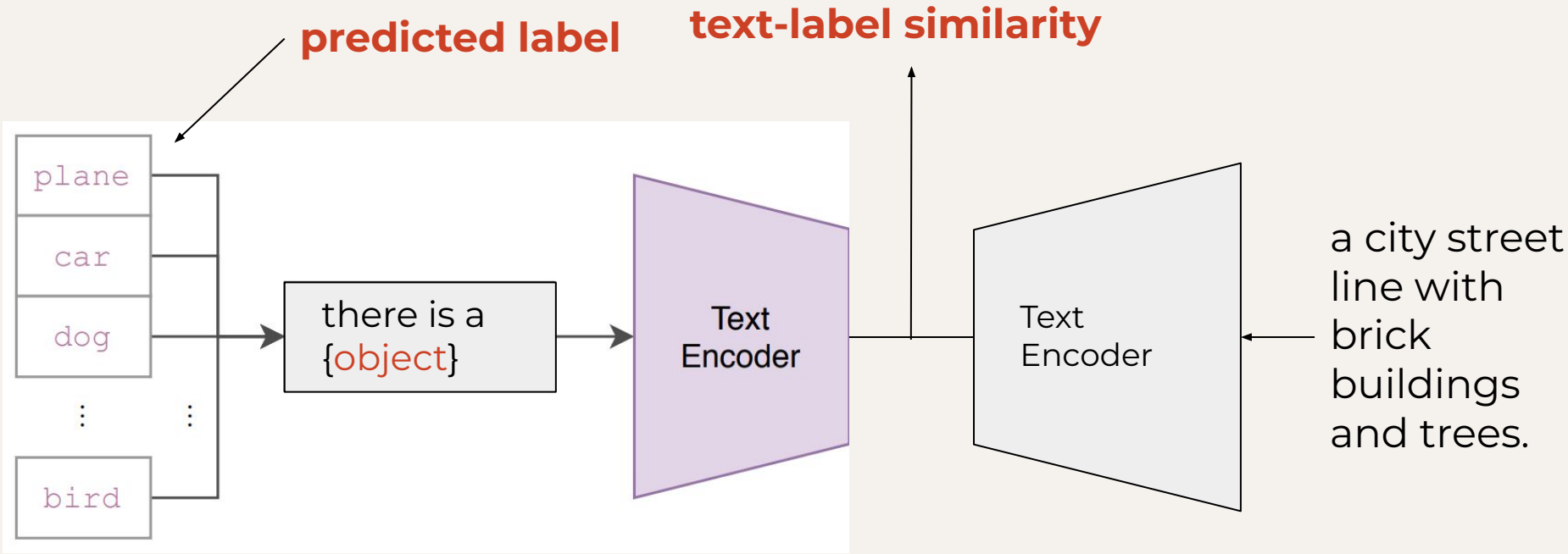
clip training setting

train from scratch on a dataset of 400 million (image, text) pairs collected from the internet.

a photo of a {object}



text-label similarity



UKPLab/sentence-transformers: Multilingual Sentence & Image Embeddings with BERT (github.com)

text-label similarity

$$u_i = f_{txt}^{\mathrm{S}}(l_i)$$
$$v = f_{txt}^{\mathrm{S}}(h)$$
$$sim = cos(u_i, v)$$

u_i : the i-th label belongs to Image-Net
h : the caption being used generating image

text-label similarity

$$index, sim' = pick(sim, \alpha)$$
$$sim'' = softmax(sim')$$
$$w = f^{\mathrm{V}}_{img}(x)$$
$$score = \sum_{i \in index} (w_i \times sim''_i)$$

pick : return index and value for whose value is large then threshold α
w : label distribution predicted by pretrained VGG network
score : sum of the element wise product between w and *sim*

SOO test

a city street line with brick buildings and trees.

random

SOO test

a zebra is walking in the snow.

random

optimizated

```
population : 100
offspring  : 100
generation : 250
```

score

a zebra is walking in the snow.

```
zebra        : 0.6903 (softmax : 1)

final score : 99.9994
```

analyze

a zebra is walking in the snow.

text - image score : **30.41**

zebra , probability = **1.00**
tiger , probability = 0.00
gazelle , probability = 0.00
prairie chicken , probability = 0.00
tiger cat , probability = 0.00

SOO test

a green train is coming down the tracks.

random

optimizated

```
population : 100
offspring  : 100
generation : 250
```

score

a green train is coming down the tracks.

```
bullet train          : 0.5303 (softmax : 0.333)
electric locomotive   : 0.5053 (softmax : 0.324)
steam locomotive      : 0.5607 (softmax : 0.343)


final score           : 34.242
```

analyze

a green train is coming down the tracks.

text - image score : **25.64**

steam locomotive , probability = **1.00**
electric locomotive , probability = 0.00
passenger car , probability = 0.00
freight car , probability = 0.00
streetcar , probability = 0.00

04

MOO

combination all features

MOO test

a city street line with brick buildings and trees.

random

optimizated

```
population : 100
offspring  : 100
generation : 250
```

analyze

a city street line with brick buildings and trees.

text - image score : **32.97**

streetcar , probability = **0.28**
traffic light , probability = 0.09
viaduct , probability = 0.08
bell cote , probability = 0.06
palace , probability = 0.04

MOO test

a zebra is walking in the snow.

random

optimizated

```
population : 100
offspring  : 100
generation : 250
```

analyze

a zebra is walking in the snow.

text - image score : **36.28**

zebra , probability = **1.00**
tiger , probability = 0.00
dalmatian , probability = 0.00
impala , probability = 0.00
tiger cat , probability = 0.00

MOO test

a green train is coming down the tracks.

random

optimizated

```
population : 100
offspring  : 100
generation : 250
```

analyze

a green train is coming down the tracks.

text - image score : **30.08**

electric locomotive , probability = **0.43**
steam locomotive , probability = **0.42**
passenger car , probability = 0.07
freight car , probability = 0.07
streetcar , probability = 0.01

test

a green train running on the rails.

a green train is coming down the tracks.

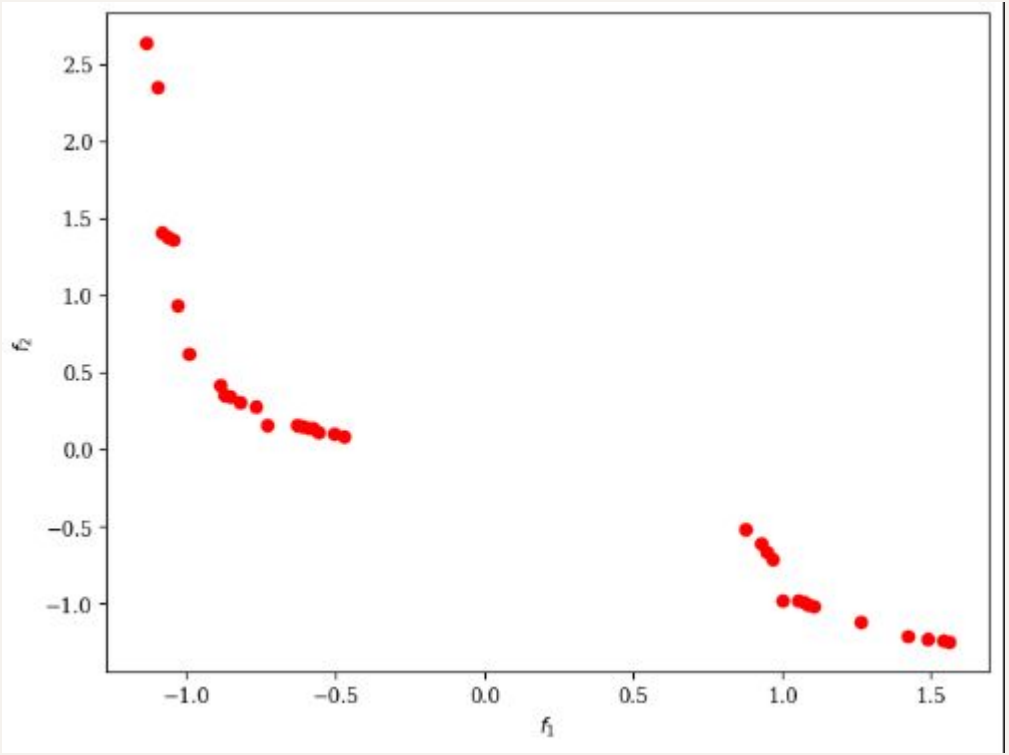
distributation

a green train is coming down the tracks.

f1 : image-text similarity

f2 : text-label similarity

#nds : **34**



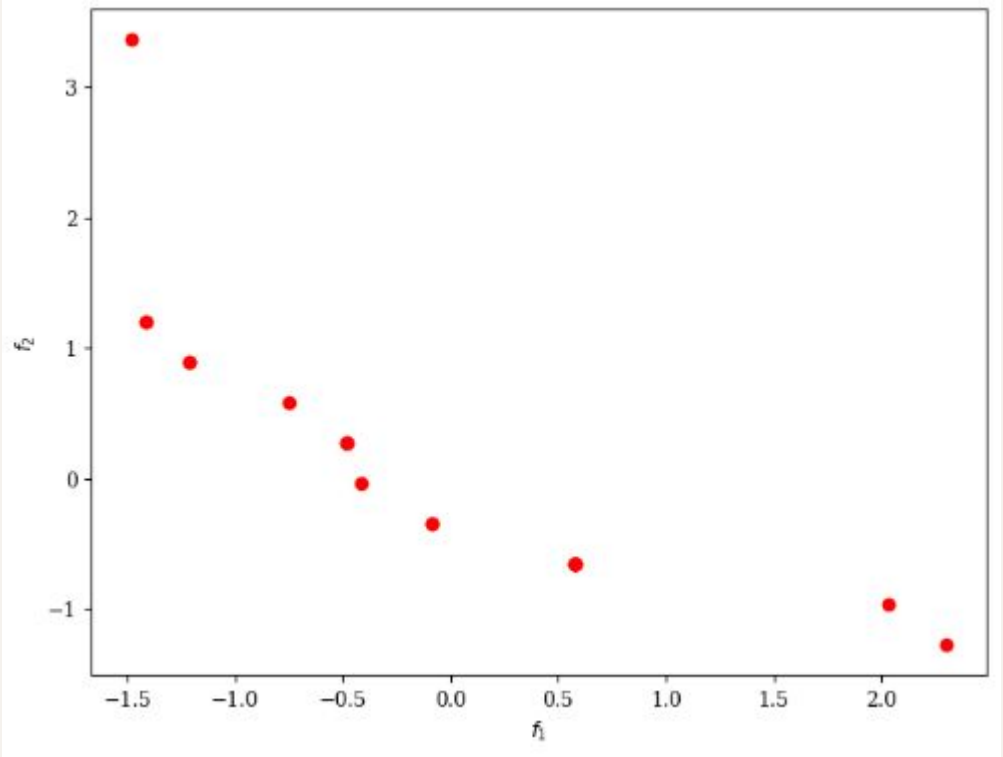
distributation

a zebra is walking in the snow.

f1 : image-text similarity

f2 : text-label similarity

#nds : **23**



05

T test

check different optimization setting

experiments setting

- #nds : 12
- objective :
  - image-text similarity
  - text-label similarity
- #generation
  - [50, 100, 250, 500, 1000]
- tail : single tail
- type : paired t-test

50 v.s. 100

30.22	27.16
30.10	28.48
30.05	32.48
29.83	32.93
29.66	33.49
30.18	28.46
30.20	27.17
28.05	34.13
28.00	34.13
29.57	33.72
30.08	30.12
29.39	34.03

30.64	33.98
29.76	34.24
30.27	34.20
30.44	34.19
30.18	34.22
30.57	34.13
29.83	34.24
30.13	34.22
30.08	34.24
30.42	34.20
30.59	34.12
30.20	34.21

50 vs 100
image-text sim
0.007001
text-label sim
0.002629

100 v.s. 250

30.64	33.98
29.76	34.24
30.27	34.20
30.44	34.19
30.18	34.22
30.57	34.13
29.83	34.24
30.13	34.22
30.08	34.24
30.42	34.20
30.59	34.12
30.20	34.21

30.27	31.11
33.01	26.88
32.93	27.91
30.18	31.44
29.42	32.67
32.35	29.63
30.00	32.15
30.10	32.11
32.71	29.19
29.35	32.69
32.96	26.98
29.52	32.62

100 vs 250
image-text sim
0.053529
text-label sim
0.000061

250 v.s. 500

30.27	31.11
33.01	26.88
32.93	27.91
30.18	31.44
29.42	32.67
32.35	29.63
30.00	32.15
30.10	32.11
32.71	29.19
29.35	32.69
32.96	26.98
29.52	32.62

29.57	34.24
30.64	33.60
30.69	33.57
30.44	34.09
30.57	33.70
29.59	34.24
30.20	34.18
30.52	33.70
30.35	34.14
30.18	34.19
29.54	34.24
30.71	33.56

250 vs 500
image-text sim
0.061783
text-label sim
0.000109

500 v.s. 1000

29.57	34.24	30.20	33.82
30.64	33.60	30.40	32.42
30.69	33.57	28.56	34.09
30.44	34.09	30.32	33.30
30.57	33.70	28.52	34.12
29.59	34.24	28.59	34.07
30.20	34.18	28.30	34.16
30.52	33.70	30.47	32.34
30.35	34.14	30.44	32.38
30.18	34.19	30.15	33.82
29.54	34.24	30.10	33.84
30.71	33.56	26.54	34.17

500 vs 1000
image-text sim
0.029997
text-label sim
0.043135

conclusion

50 vs 100	100 vs 250	250 vs 500	500 vs 1000
image-text sim	image-text sim	image-text sim	image-text sim
0.007001	0.053529	0.061783	0.029997
text-label sim	text-label sim	text-label sim	text-label sim
0.002629	0.000061	0.000109	0.043135