

# 鋼胚之印刷及手寫序號辨識

Recognize the printed and handwritten numbers on the Steel billet

第12組

葉凌瑋 / 黃博鴻 / 黃鈺凌 / 常興唯

# CONTENT

簡介  
Introduction

資料集  
Dataset

方法  
Method

結果  
Result



# 簡介

## Introduction

傳統在產線上紀錄序號，需要靠人工去辨別鋼胚上序號，這樣相當耗費人力。

Traditionally, record the numbers on the production line, requires humans to identify the numbers on the steel billet, it's too wasted human power.

因此為了達到工業自動化，減少產線上的人力成本，我們希望能夠以深度學習方法，實現鋼胚上之印刷及手寫序號辨識。

Therefore, in order to achieve industrial automation and human power wasted, we hope to use deep learning methods to realize printing and handwritten number recognition on steel billet.

這樣不僅能減少人力，更能有效的做品質控管及問題的返溯。

In this way ,not only can reduce manpower, but also do quality control and problem tracking more effective.

## 中鋼人工智慧挑戰賽-字元辨識

Training data : 12067

Testing data : 6037

H = 1028



W = 1232

Filename \_\_eN=NSp5V3VuC9dHV9664Gdp0gFf3Q

Label 3LN843054A

Top-right X 789.998

Top-right Y 452.54965

Bottom-right X 790.7092

Bottom-right Y 500.466

Top-left X 446.40326

Top-left Y 506.79227

Bottom-left X 445.69202

Bottom-left Y 458.87595

# 方法

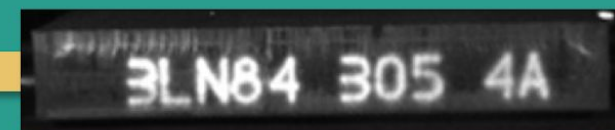
Method



## 文字偵測

Text Detections

YOLO



# 方法

Method

[class] [x center] [y center] [width] [height]



```
zidane.txt
0 0.481719 0.634028 0.690625 0.713278
0 0.741094 0.524306 0.314750 0.933389
27 0.364844 0.795833 0.078125 0.400000
```

# 方法

## Method

由於鋼胚序號皆可以用矩形框出，且背景單純，文字出現位置變化也不大。

Because the numbers on the steel billet can be framed by rectangle, and the background is simple and the position of the text does not change much.

因此我們選擇使用 YOLOv5，並選擇參數量較少的YOLOv5m6 作為 pre-trained weight。

So we choose to use YOLOv5, and choose YOLOv5m6 with fewer parameters as the pre-trained weight.

超參數部分，設定 Batch size 為 8，訓練 200 個 epochs

For the hyperparameter, we set the Batch size to 8, and train for 200 epochs.

# 方法

Method



文字辨識  
Text Recognition

3LN843054A



# 方法

Method

Original



Perspective



Rotate



Shrink



Blur



Stretch



Noise



Contrast  
Brightness



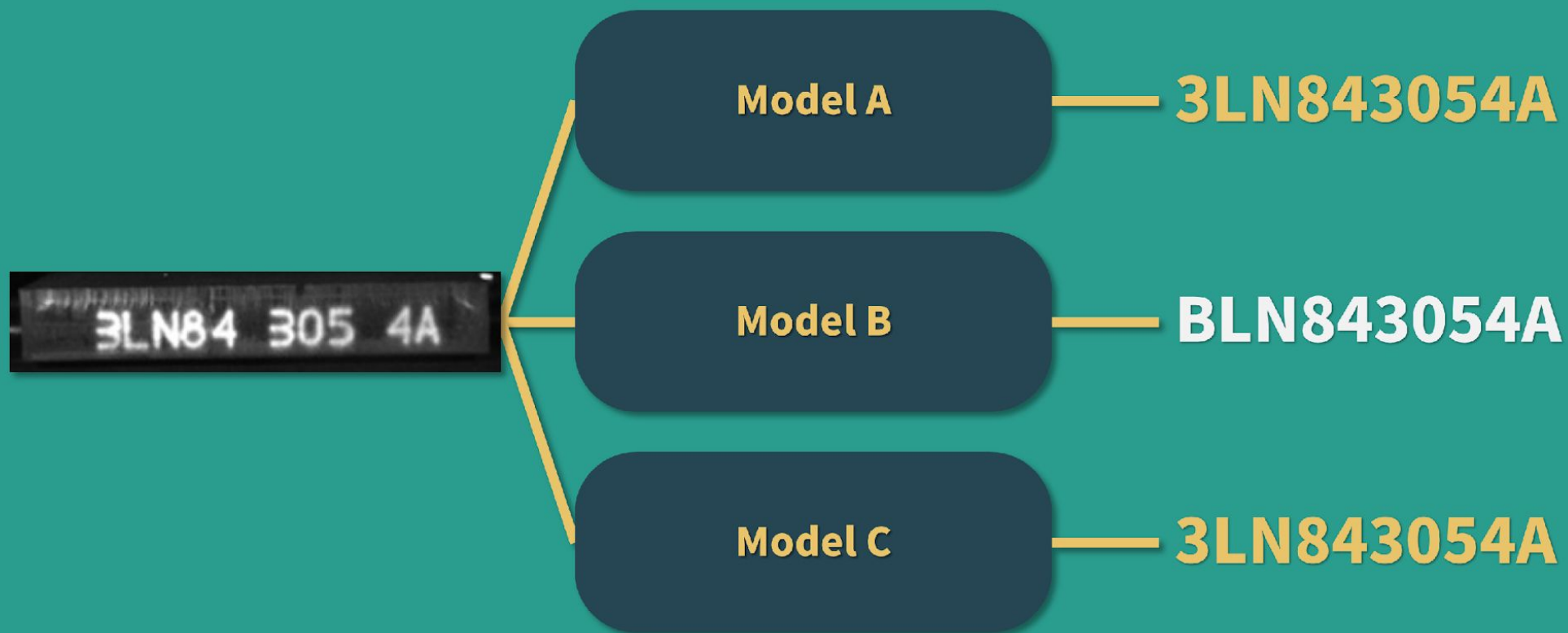
# 方法

## Method

<b>model</b>	<b>Train data</b>		<b>FT</b>
TRBA	Original images	1x	N
TRBA	Original images ∙ Blur ∙ Noise	9x	N
TRBA	Original images ∙ Blur ∙ Noise ∙ Contrast / Brightness	10x	N
CRNN+CTC	All Data Augmentation	20x	N
TRBA	All Data Augmentation	20x	N
TRBA	Original images	1x	Y
TRBA	Original images ∙ Blur ∙ Noise	9x	Y
TRBA	Original images ∙ Blur ∙ Noise ∙ Contrast / Brightness	10x	Y
CRNN+CTC	All Data Augmentation	20x	Y
TRBA	All Data Augmentation	20x	Y

# 方法

Method



# 方法

Method

model	Train data		FT
TRBA	Original images	1x	N
TRBA	Original images 、 Blur 、 Noise	9x	N
CRNN+CTC	All Data Augmentation	20x	N
TRBA	All Data Augmentation	20x	N
TRBA	Original images 、 Blur 、 Noise 、 Contrast / Brightness	10x	Y
TRBA	All Data Augmentation	20x	Y

# 方法

Method

## Score

$$\sum_{i=1}^N err(pred_i, label_i) + \frac{1}{N} lev(pred_i, label_i) / \max(|pred_i|, |label_i|)$$

錯誤字串個數

Numbers of error string

每個字串Levenshtein Distance的平均值

Average of every string's Levenshtein Distance

# 結果

Result

model	Train data		FT	Score
TRBA	Original images	1x	N	284
TRBA	Original images ◊ Blur ◊ Noise	9x	N	238
TRBA	Original images ◊ Blur ◊ Noise ◊ Contrast / Brightness	10x	N	220
CRNN+CTC	All Data Augmentation	20x	N	338
TRBA	All Data Augmentation	20x	N	198
TRBA	Original images	1x	Y	537
TRBA	Original images ◊ Blur ◊ Noise	9x	Y	355
TRBA	Original images ◊ Blur ◊ Noise ◊ Contrast / Brightness	10x	Y	345
CRNN+CTC	All Data Augmentation	20x	Y	418
TRBA	All Data Augmentation	20x	Y	283

# 結果

Result

Ensemble model	Validation score	Public test score
2	178	74
3	169	67
4	162	62
5	161	65
6	157	60
7	157	64
8	160	63

# 結果

## Result

這次實驗的Validation score，從一開始直接拿原始資料硬train一發得到的284。

The Validation score of this experiment was 284, which was obtained by training the original data directly in the beginning.

經過資料增強後提升到198，最後再透過ensemble的技巧提升到157。

After data augmentation, it increased to 198, and finally increased to 157 through ensemble.

使用pretrain model fine-tune的model，雖然結果都不太好，但在ensemble時可以彌補train from scratch的不足。

Using the pretrain model fine-tune model, although the results are not very good, but in the ensemble, it can make up for the lack of train from scratch.



## 結果 Result

最後使用6個model ensemble，在private data結果為118.0028，第7名。  
Finally, we use 6 model ensembles, and the result on private data was 118.0028, 7th.

辨識失敗的字串大多為手寫字串及重疊的文字。這兩類的圖片資料比較少，也無法透過資料增強方式產生。

Most of the failed to be recognized were handwritten strings and overlapping texts. These two types of image data are too few and cannot be generated through data augmentation.



**THANK YOU  
FOR  
LISTENING**

**THE END**