ECE 364 Project Option #1: Identification of Chinese Characters in an Ensemble Canvas
Spring 2025
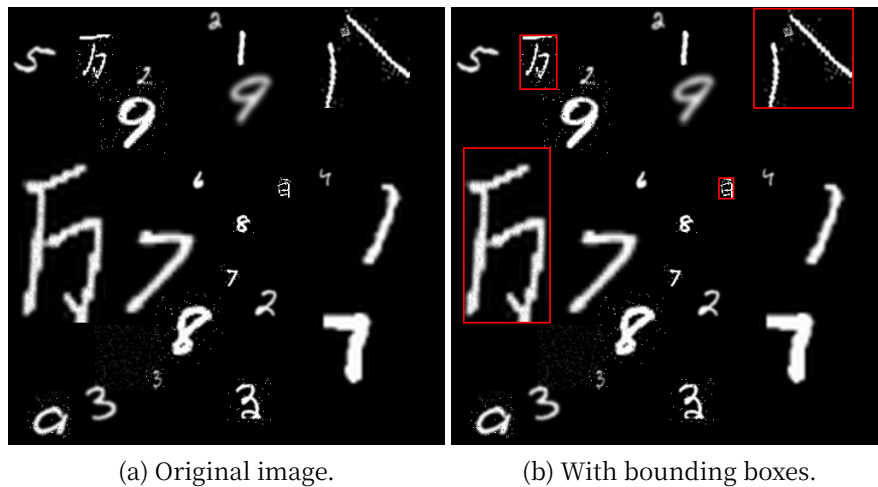

(a) Original image.   (b) With bounding boxes.

Fig. 1: Example from the CCD-ND dataset with is bouding boxes.

Fig. 2: Annotation for Fig. 1.

1 Project Background

We have created a dataset of images that are collages containing both: the standard MNIST handwritten numerical digits, which we have used in this course 100 times before, and handwritten Chinese characters taken from the Chinese MNIST dataset. We'll call this dataset Chinese Character Recognition with Numerical Distractors (CCD-ND) dataset.

For this project, we have provided the dataset that contains 1000 annotated images for training, 256 annotated images for validation, and 512 unannotated images for testing. All images are 512×512 pixels in size, and there are a bunch of random scalings to vary the sizes of all the characters. There are 15 possible Chinese characters and 10 decimal digits.

The unannotated images will be used to evaluate your models on Kaggle. This is a bounding box and classification problem. We want to find the bounding box of the Chinese characters and each test example in one of two classes. For evaluation, please generate a `prediction.csv` file in the following format. The accuracy score calculated by Kaggle will be used as an evaluation metric. You can use the `calculate_AP` method from `topic_-1/utils/metrics.py` to evaluate your models.

Note: To ensure effective evaluation, the header and content of the `prediction.csv` file must exactly match the format, and the case of the text matters.

You can use the "submit prediction" button at the upper right corner in the Kaggle page to submit. Remember to write the names of all group members in the description. After submitting, you need to go to the "Submissions" tag and press the "select" box after your best submission, so your submission will be shown on the leaderboard.

Id	Predictions
0	...
1	...
2	...
...	...

14	{
	"image_filename" : "00014.png" ,
	"boxes" : [{ "char" : "八" , "bbox" : { "x_min" : 0.69140625, "y_min" : 0.001953125, "x_max" : 0.919921875, "y_max" : 0.23046875 }},
	{ "char" : "百" , "bbox" : { "x_min" : 0.611328125, "y_min" : 0.388671875, "x_max" : 0.646484375, "y_max" : 0.4375 }},
	{ "char" : "万" , "bbox" : { "x_min" : 0.15625, "y_min" : 0.0625, "x_max" : 0.2421875, "y_max" : 0.1875 }},
	{ "char" : "万" , "bbox" : { "x_min" : 0.02734375, "y_min" : 0.3203125, "x_max" : 0.2265625, "y_max" : 0.720703125 }}]
	}

...	...

Table 1: Format of `prediction.csv`

2 Deliverables

The following are the deliverables of this project:

- A model that outputs the bounding box and character annotation of all the Chinese characters in a set of images and outputs this to a CSV.
- The number of parameters in the model must not exceed 30 million.
- You are free to use any publicly available model (pre-trained or otherwise) with or without augmentation, but it is not a requirement. You can also augment the data as you see fit.

As a starting point, check out some of the object detection models in torchvision and the classic object detection tutorial.

3 Submission

1. Submit all your code, including training and evaluation, as a `.zip` file.
2. Submit `prediction.csv` file to Kaggle for scoring and evaluation
3. Submit a 2-page report (1-inch margin, 12-point font) and include

– Your approach, model, and any other design choice.
– Hyperparameters that you used for training.
– Training and test results.
– Any other interesting details about the approach or model.