October 1, 2022

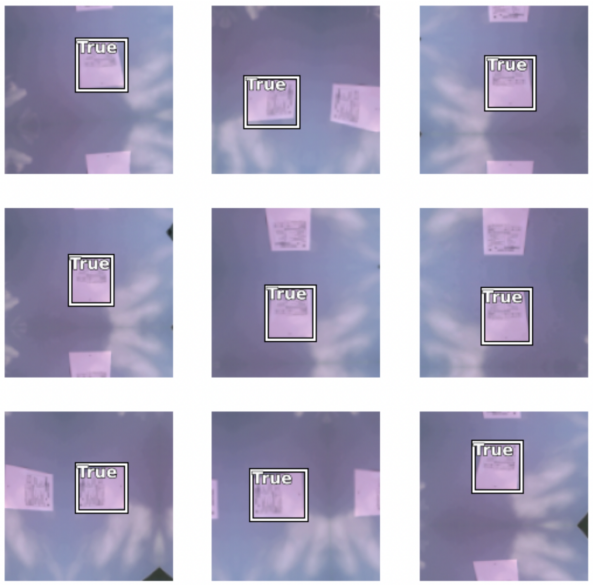
Hi [],

We hope you are doing well! Our team has been making some progress and we want to update you with our current findings.

Updates:

- Cropped images without documents using the PIL library, modified the groundtruth files to accommodate for images without documents.
- Finished making a balanced binary classification model (with images with documents + images w/o documents)
- Finished following the Object detection tutorial (https://walkwithfastai.com/Object_Detection) with our datasheet dataset. Our current

Results look like this for bounding boxes:



epoch	train_loss	valid_loss	time
0	3.680068	3.457086	01:43
1	3.499872	3.165926	01:40
2	2.977562	2.248234	01:39
3	2.687737	2.838647	01:37
4	2.472650	2.793158	01:38
5	2.319677	2.391312	01:38
6	2.149541	1.903899	01:40
7	2.044715	1.705064	01:40
8	1.962769	1.647363	01:39
9	1.885270	1.575182	01:38

Results for the binary classification:

epoch	train_loss	valid_loss	error_rate	time
0	0.573977	0.064204	0.023529	00:12

epoch	train_loss	valid_loss	error_rate	time
0	0.000360	0.000079	0.000000	00:11

[For Rohan] Do you happen to have the bounding box resources that you mentioned in your previous email? As Jessie mentioned, we are using the RetinaNet model from fastai

[For Chris and Lakshmi] We have a question regarding the anchor boxes and aspect ratios for bounding boxes. Could you possibly point us towards some resources to better get some theoretical understanding of it, and how this might be implemented in our bounding box model?

For our next steps, we would also like to know how we would go about the text detection process and do you have any resources that you think would be beneficial to us.

Here is the github link for your reference:
https://github.com/alixintong/btt-ai-know-your-customer

Thank you,
Han Pham, Alice Li, Jiaxin (Jessie) Yang, Blessing Nwogu
