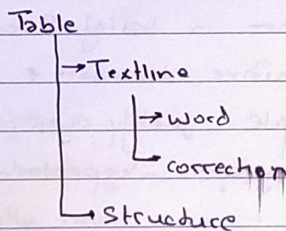
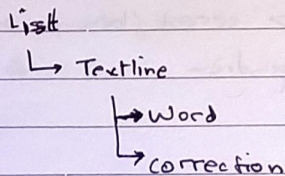
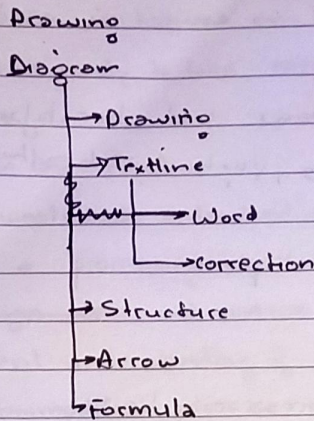
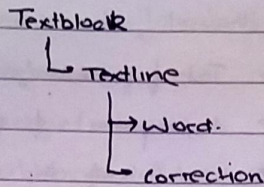


Hierarchical Object detection problem

→



Formula

Approaches

- ① flatten the hierarchy and detect all levels of hierarchy together
- ② Detect the outermost level and then zoom in according to classification of bounding box regression.

③ Scene graph generation

Disadvantage of approach ①

- Hundreds of predictions for the same image
- Model slows down drastically

$$2 \times 7 + 12$$

$$12 \times 7 + 2$$

$$21 + ③$$

OCR needs an image patch ~~from~~ from a textline.

create the xml in the aforementioned hierarchy and implement a dataloader to modify the hierarchy based on model to be predicted.

No need to predict word and correction positions in textline textline.