Proposing an Automatic ROI recommendation method for semiconductor package vision inspection

Keonhee Han1, Wonyong Choi2, Keejun Han1,†
1 School of Computer Engineering, Hansung University, Seoul, Korea
2 R&D Center, Genesem, Incheon, Korea
† E-mail: keejun.han@hansung.ac.kr

Research Purpose

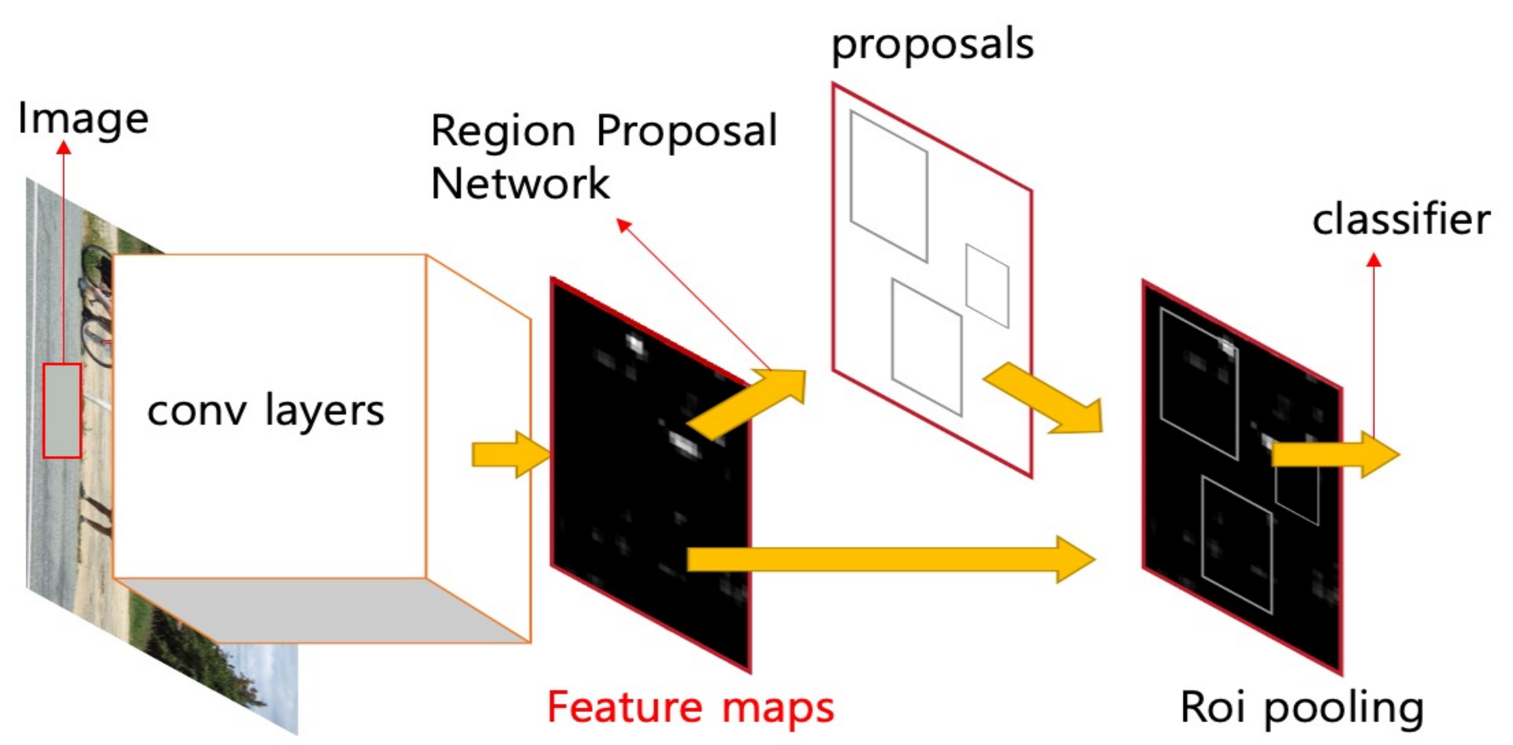
Problem

The baseline rule-based models have limitations on recommending the ROI of new packages it has never seen before.

Including the previously discussed cold start problem, we propose a new deep learning approach for setting Auto ROI during the Teaching process in the context of semiconductor vision inspection.

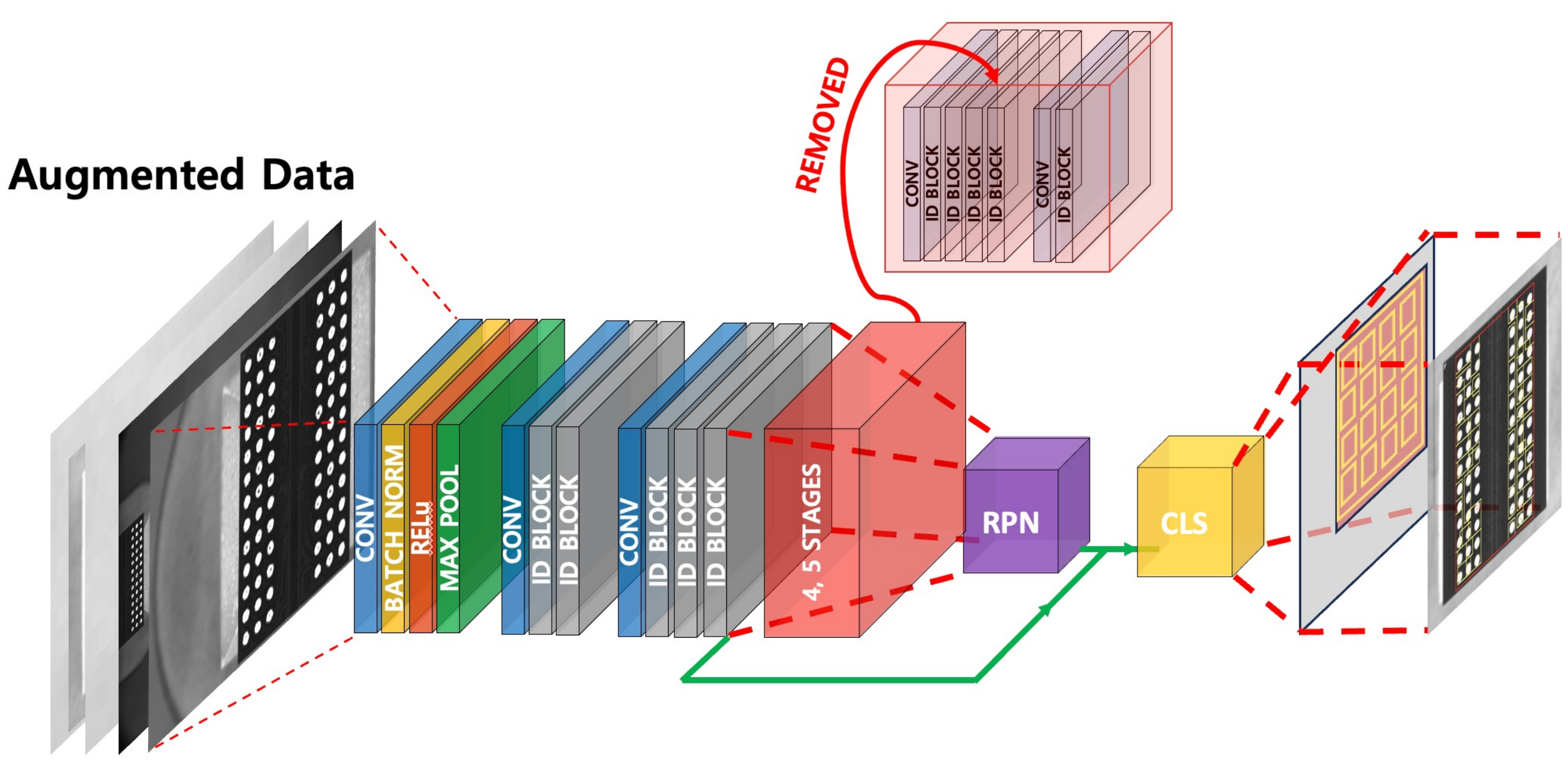
Related Work

Faster Region-based Convolutional Neural Network (Faster R-CNN)



Faster R-CNN is a single, unified network for object detection composed of two modules. The first module is a deep fully convolutional network that proposes regions, and the second module is the Fast R-CNN detector that uses the proposed Regions. (ShaoqingRen et al., 2015).

Method



Step 01 - Preprocessing

Generate data from images and Ground Truth boxes for each anchor.

- It contains positional difference information between ground truth boxes and anchor boxes.

$$\hat{G}_x = P_w d_x(P) + P_x \quad \hat{G}_y = P_h d_y(P) + P_y \quad \hat{G}_w = P_w \exp(d_w(P)) \quad \hat{G}_h = P_h \exp(d_h(P))$$

$$t_x = (G_x - P_x)/P_w \quad t_y = (G_y - P_y)/P_h \quad t_w = \log(G_w/P_w) \quad t_h = \log(G_h/P_h)$$

The value 't' which indicates the amount by which P(anchor box) should be moved to regress towards the G(ground truth box), is subsequently employed in the generation of RPN target data. (Ross Girshick et al., 2014)

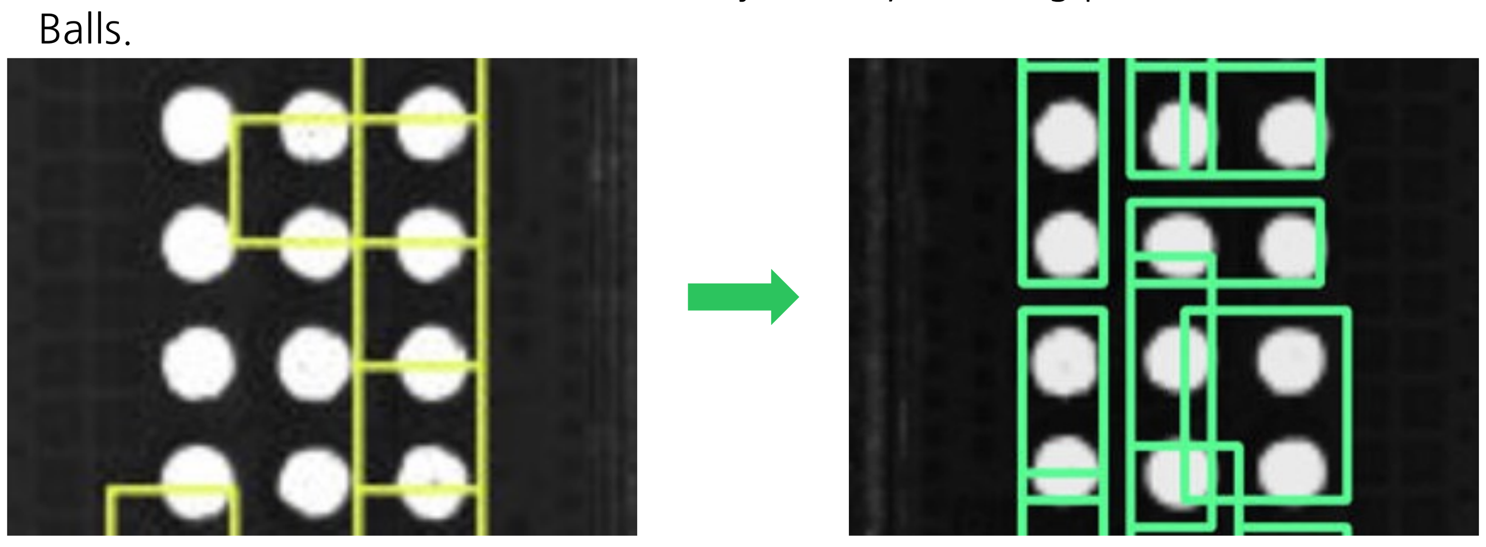
Generate object presence labels.

- We generate labels for anchor boxes with an IOU exceeding a certain threshold, indicating the presence or absence of objects in relation to the ground truth.

Step 03 - Region Proposal

Specify the base layer shared between RPN and the classification layer.

- Replace VGG with ResNet-50.
- Exclude the last two stages(4, 5 Stage) of ResNet-50. It addresses the constraint related to the minimum object size, enabling precise detection of Balls.



Region Proposal Layer (RPN)

- The model assigns each anchor to all coordinates in the input image and learns the difference between the ground truth box for object presence and bounding box regression.

Step 04 - Classification

Classification Layer

- We train for the object's location and its corresponding class information.

Step 05 - Region Aggregation

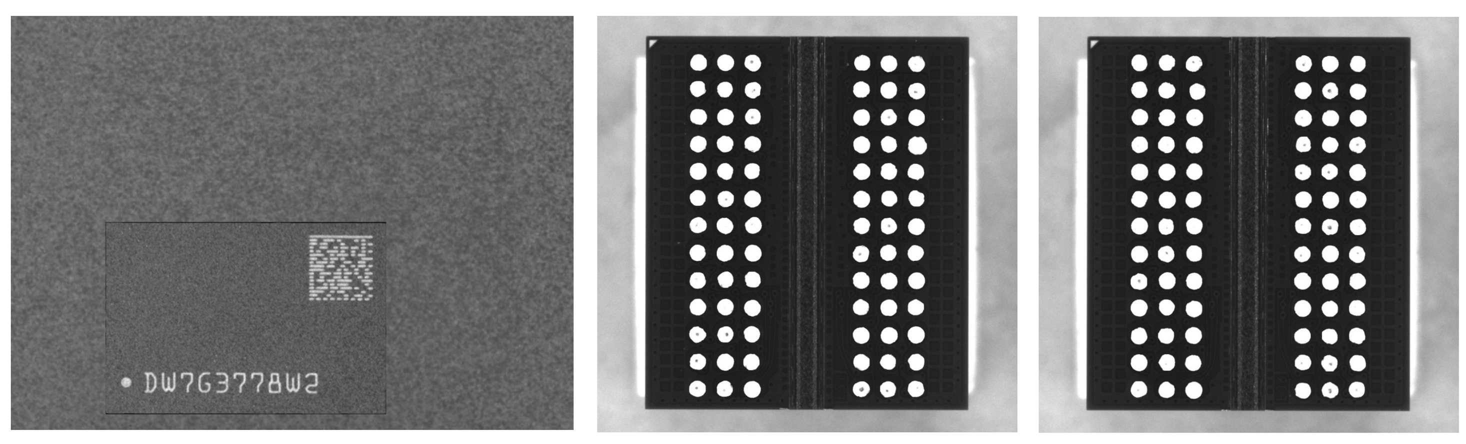
Perform Region Aggregation for all detected Ball areas.

- We employed a straightforward approach, aggregating the regions of balls using the coordinates of the smallest and largest values among the detected regions as the final ROI.

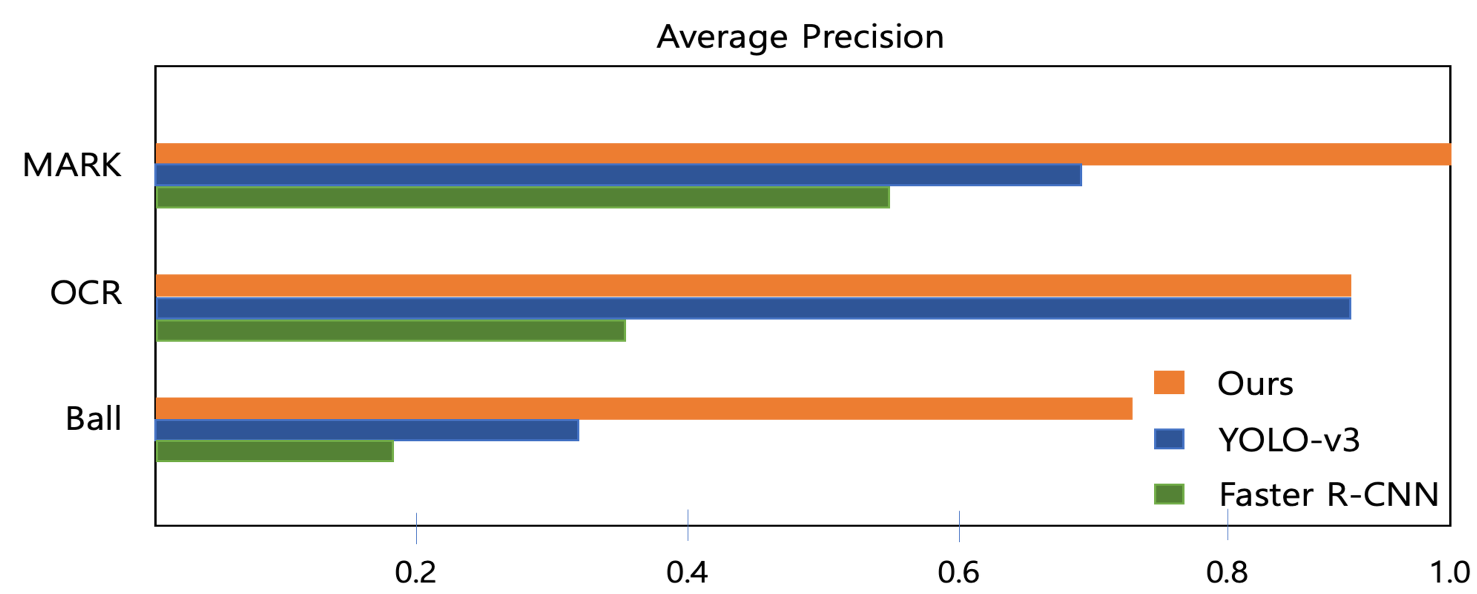
Experiment

01 Datasets

- Dataset : Using 1500 images from BGA Bottom Vision collected in the actual industrial field, we conducted data augmentation, along with 1000 images from Mark and OCR Top Vision.



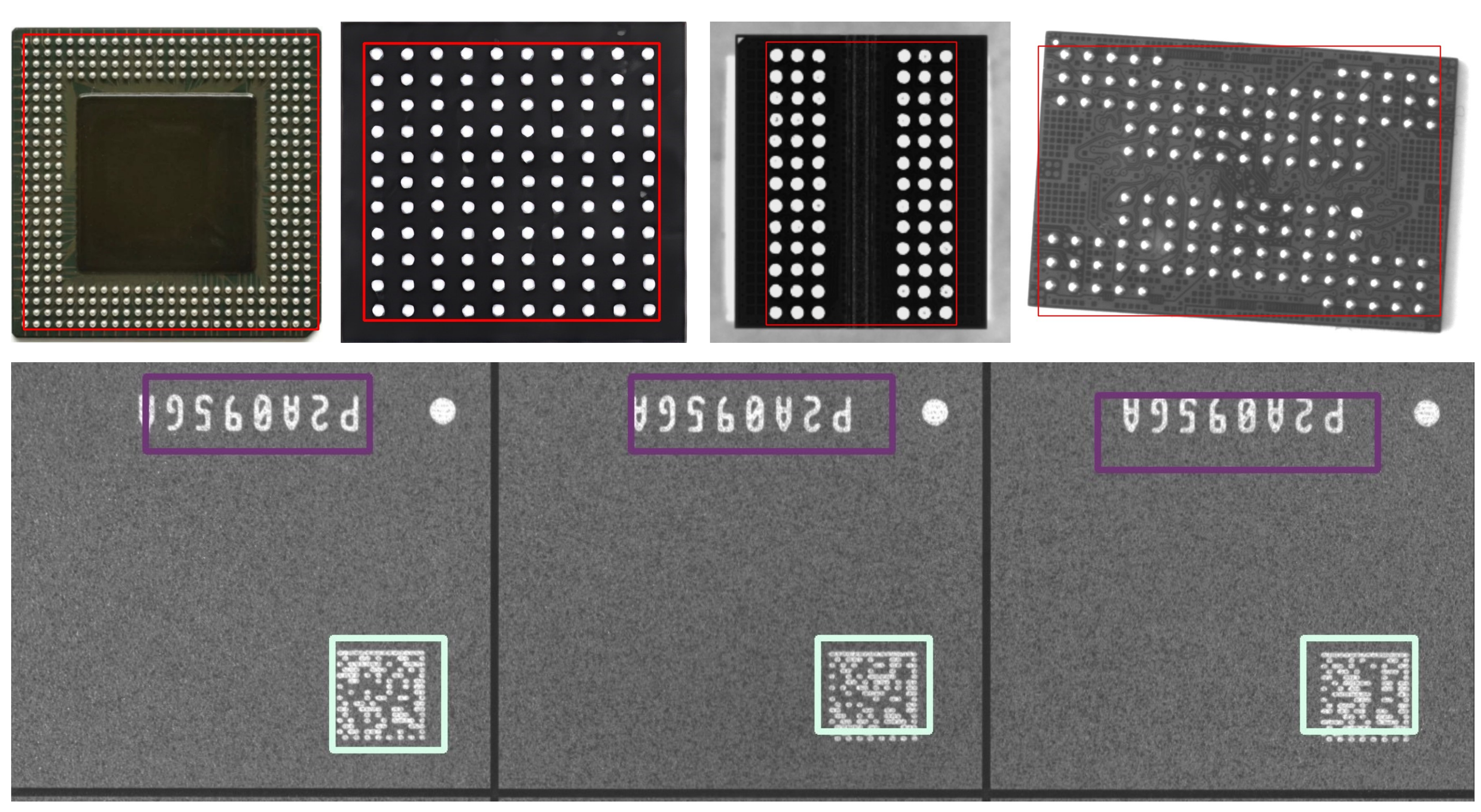
02 Performance of our framework



Model	Ball AP	OCR AP	MARK AP	mAP
Faster R-CNN	**19.0 %**	**36.0 %**	**56.0 %**	**36.89 %**
YOLO-v3	**33.0 %**	**92.0 %**	**74.0 %**	**66.3 %**
Ours	**72.0 %**	**92.0 %**	**100.0 %**	**87.96 %**

- Despite our model having a shallower architecture compared to existing models in the semiconductor packaging domain, it achieves the highest mAP.

04 Result



- The results demonstrate that our model adeptly detects the regions of various package balls and accurately identifies Marks and OCR. In particular, it showed its capability to detect the majority of balls in one package compared to previous comparative models that had undergone training.

05 Limitation

The model dynamically detect the areas related to the balls, but it is not feasible within this model to specifically detect individual balls at precise locations during the teaching process.

Conclusion and future work

We proposed a novel approach to Auto ROI using deep learning and demonstrated successful ROI detection across various packages. Furthermore, by enhancing the base layer, we effectively detected even smaller objects with high performance.

Future work

The research will implement an ensemble model that combines the attention-based object detection model with other modalities.