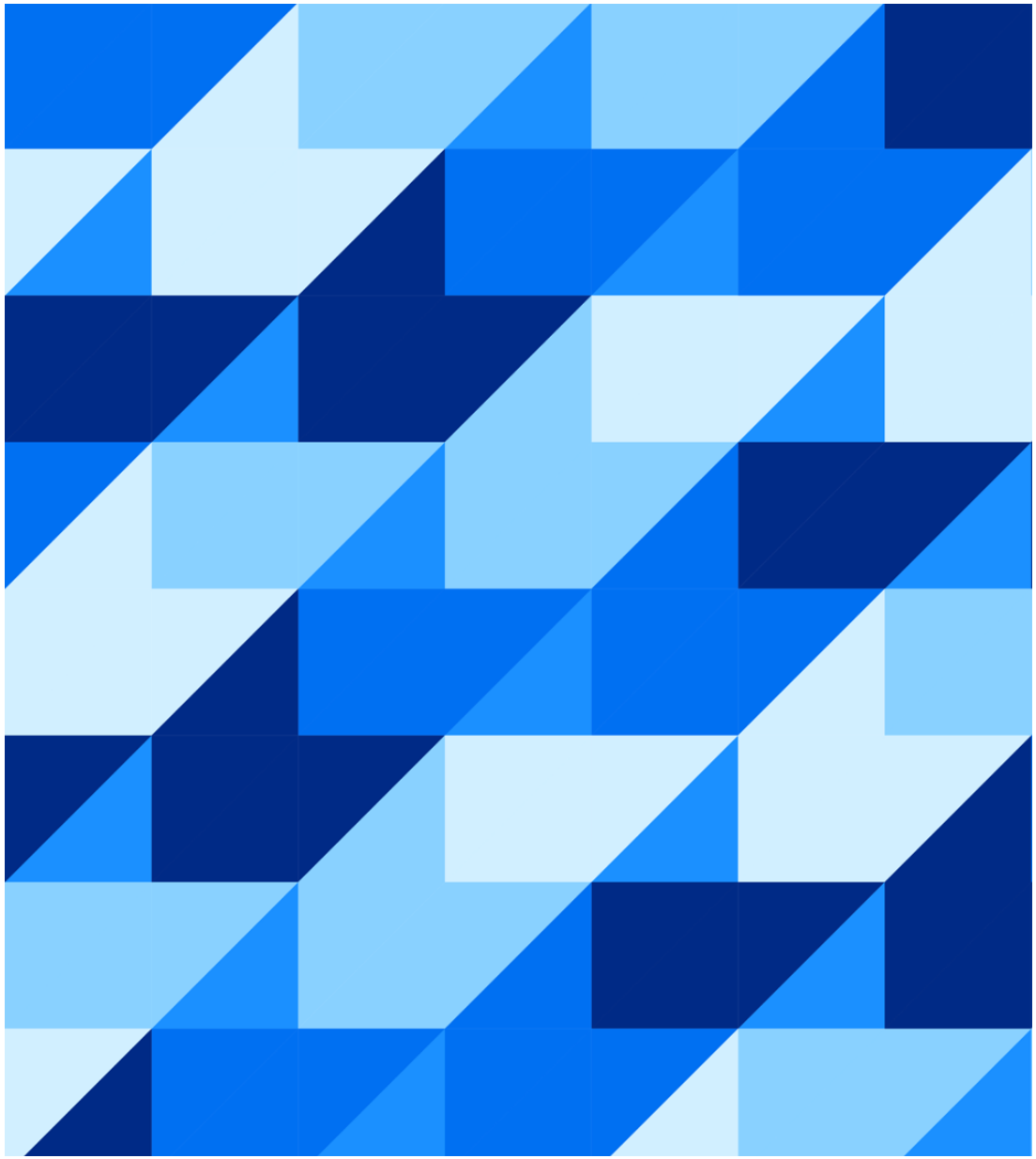
思爱普 SAP

GenVI: multi-modal generative AI solution for Visual Inspection

Xinyan Chen, SAP
November 29, 2024

Public



Agenda

1 Introduction to Visual Inspection

2 Zero-shot Referring Segmentation with GPT4v Generated Prompt

3 Towards more detailed auto-labelling mechanism with semantic feature extraction

4 Generating Minority Class Defect Detection Data
from Visual Inspection Dataset using Self-Supervised Defect Generator

Transforming from using naked eye to AI empowered algorithm

Motivation

Overcome human cognitive limits, improve productivity and quality significantly.

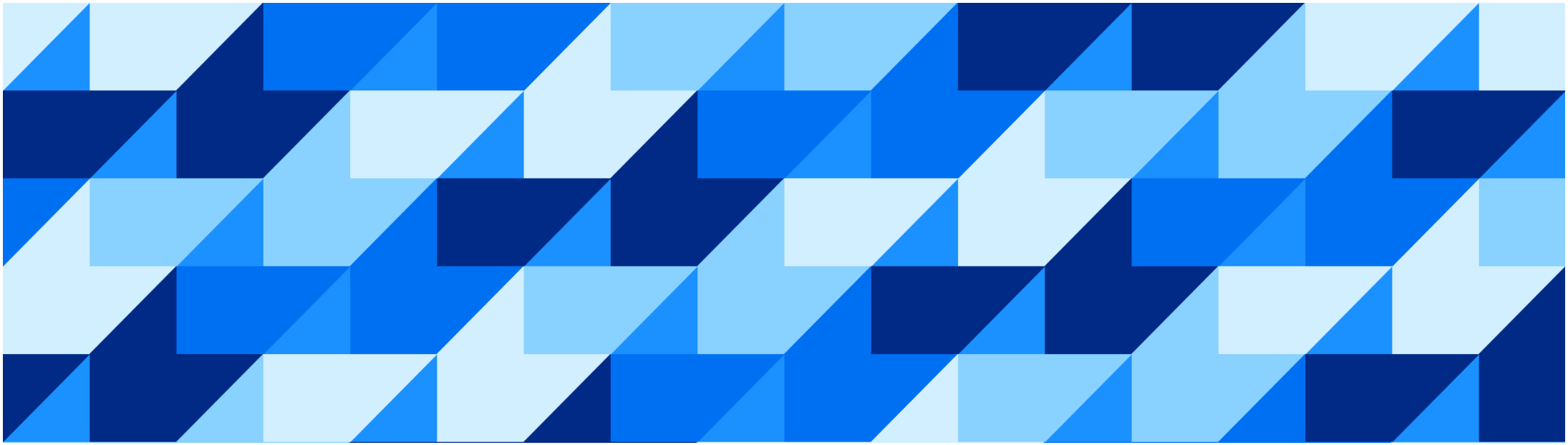
Current AI Solution

End-to-end AI algorithms for:
- Classification, Object detection, and Segmentation

Train from well-labelled datasets

Task specific model for certain application scenarios

Zero-shot Referring Segmentation with GPT4v Generated Prompt



Zero-shot Referring Segmentation with GPT4v Generated Prompt

Motivation

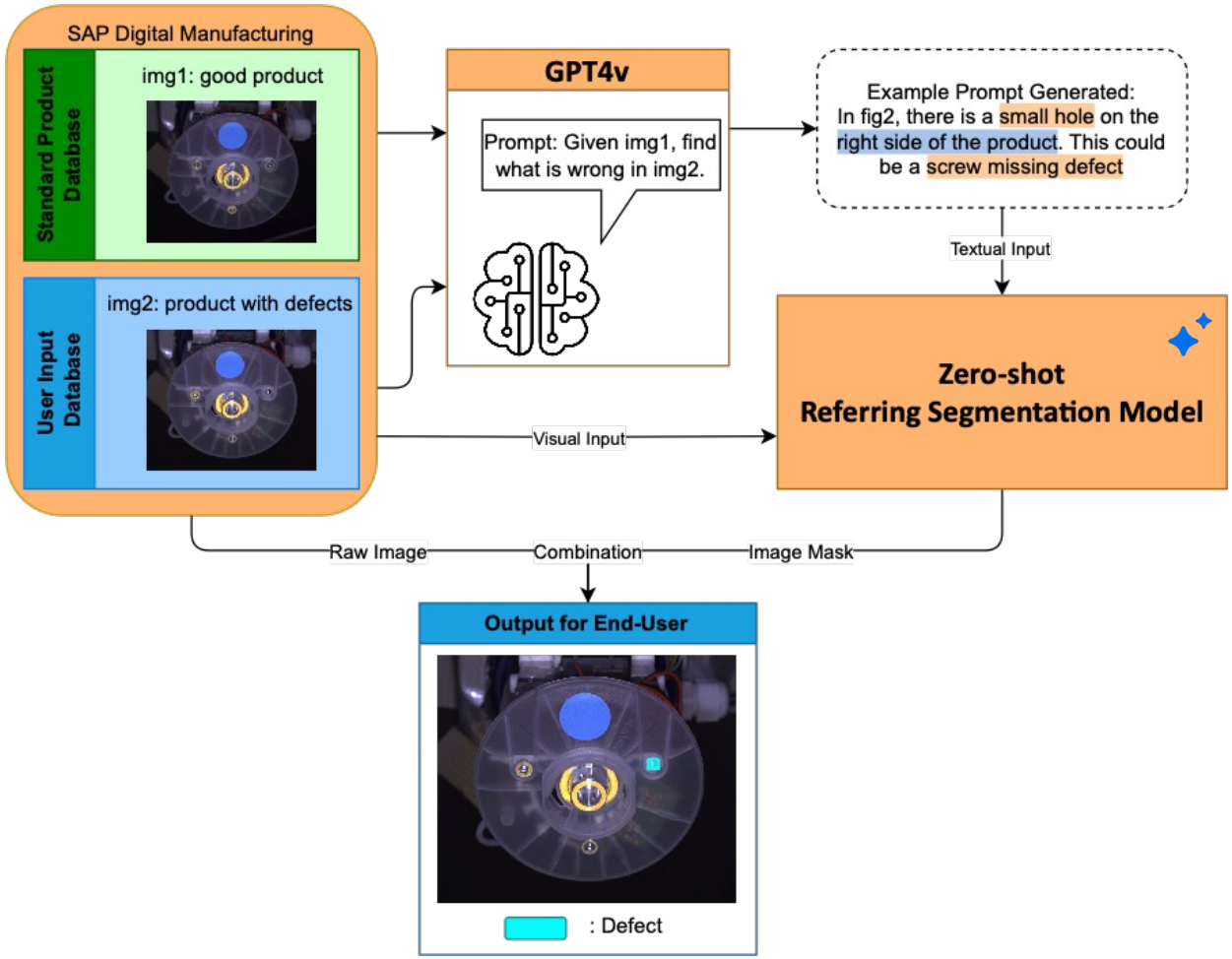
Current methods for visual inspection are mostly based on end-to-end object detection & segmentation models which are trained fully supervised by well labelled comprehensive dataset.

They have the following drawbacks:

•Well labelled dataset will cost lots of labor force

•Require annotators to have specific expertise

Zero-shot Referring Segmentation with GPT4v Generated Prompt

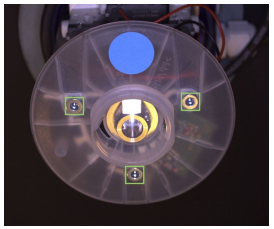


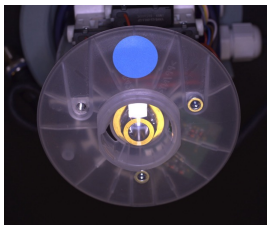
Zero-shot Referring Segmentation with GPT4v Generated Prompt

Experiment: Using GPT4v + Segmentation model, Two-stage

Image → **GPT4V** → Identified defects *(defect labels)* → **Segmentation Model** → Paint the location of the defect

Prompt → **GPT4V**





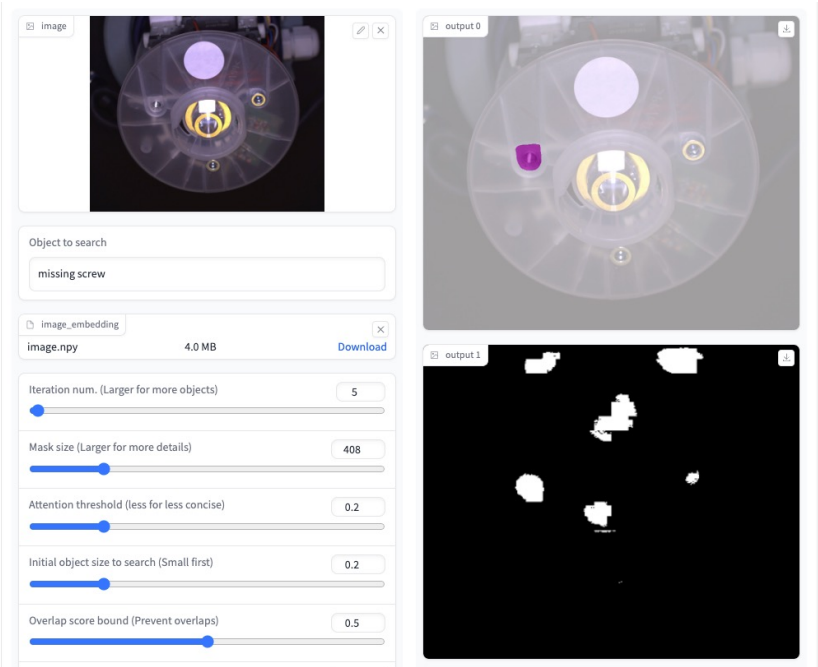
Prompt 1:
"The image above (img1) is a correctly ensembled factory product, which consists of three screws (right, left, bottom) and one translucent cap. We have draw three bounding boxes to indicate the location of the three screws."

Prompt 2:
"We have presented img1 which is correctly assembled. And the last image shows a test image. Please first check whether there exist a screw in each bound box area, then check whether the cap is there. Finally, please tell me:
Are there any defect in the image?
If so, what are the defects in the image?
and where are the defects located?"

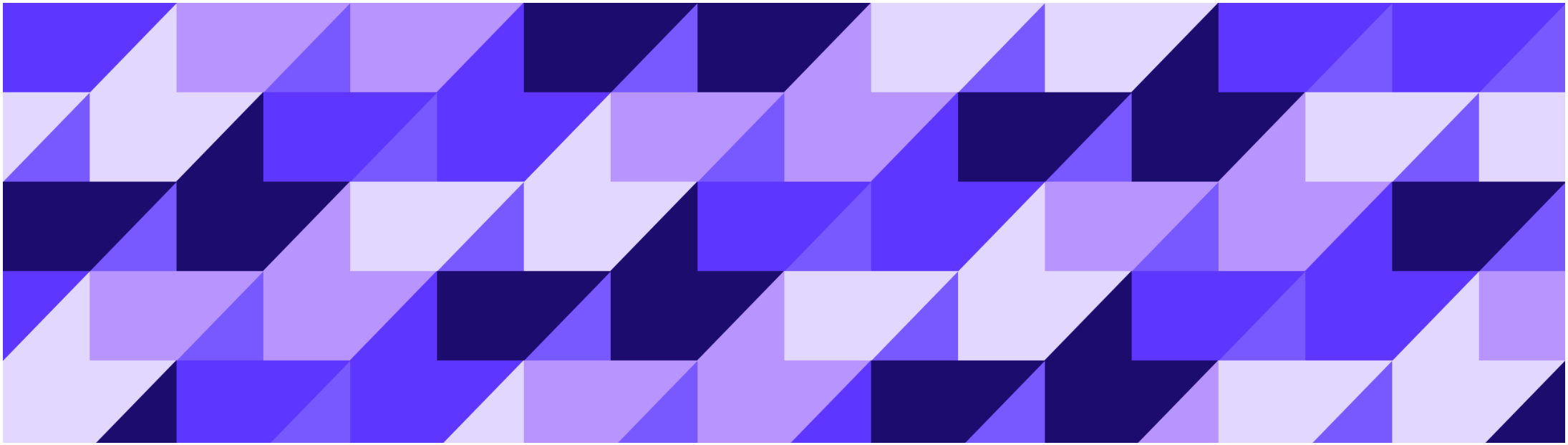
Response:

In the test image provided, there is a screw present in the right and bottom bounding box areas, but the left bounding box area is missing a screw. The cap is present in the center of the image. The defects in the image are as follows:

Missing screw in the left bounding box area. The defect is located on the left side of the image where the bounding box is drawn but no screw is present.



Towards more detailed auto-labelling mechanism with semantic feature extraction



Towards more detailed auto-labelling mechanism with semantic feature extraction

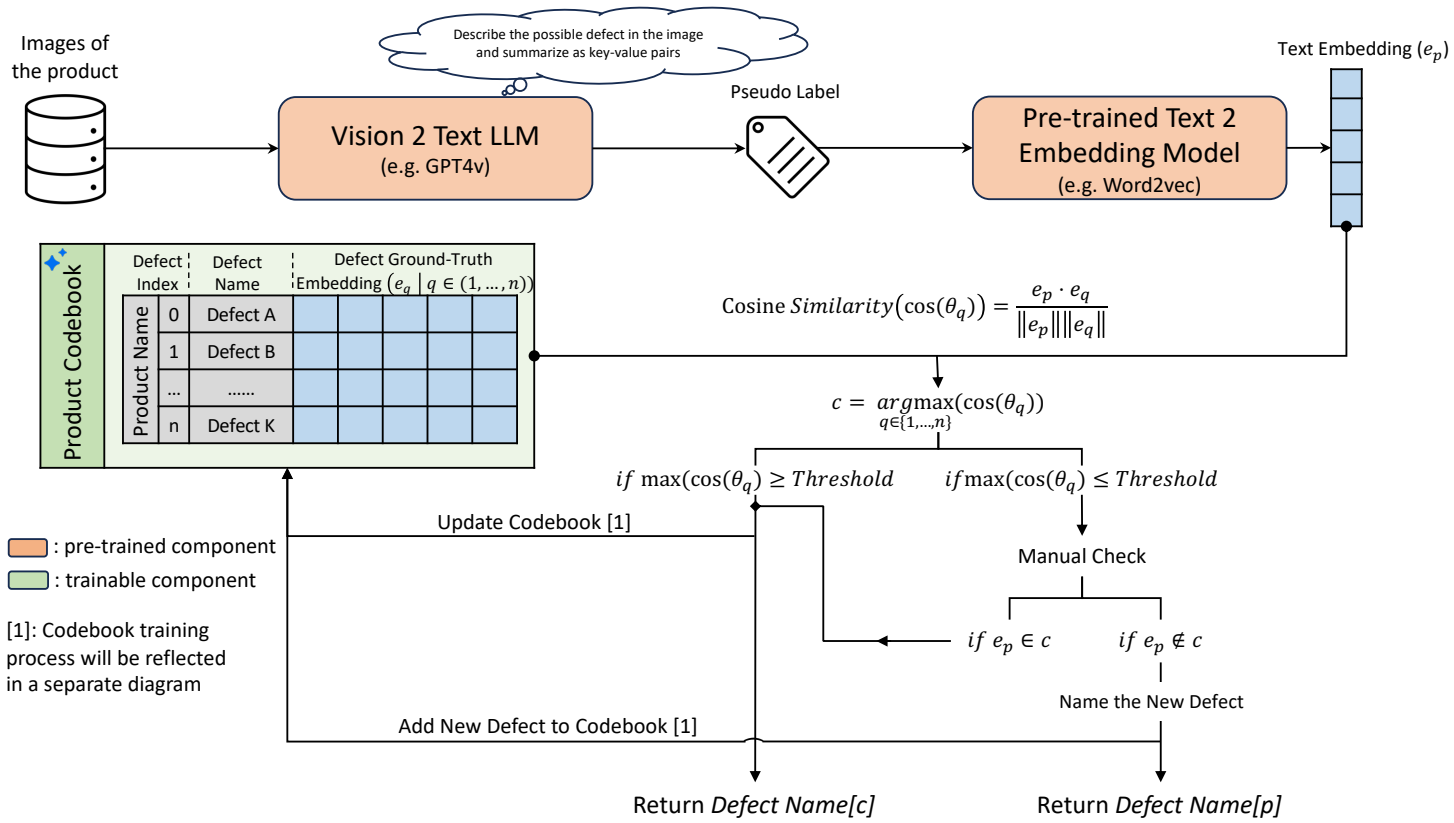
Motivation

Since most of the classification model in Visual Inspection are trained by supervision, we need an efficient auto-labelling process for images come from production to reduce the manual labelling cost. The current auto-labelling process is mostly based on semi-supervised learning algorithms.

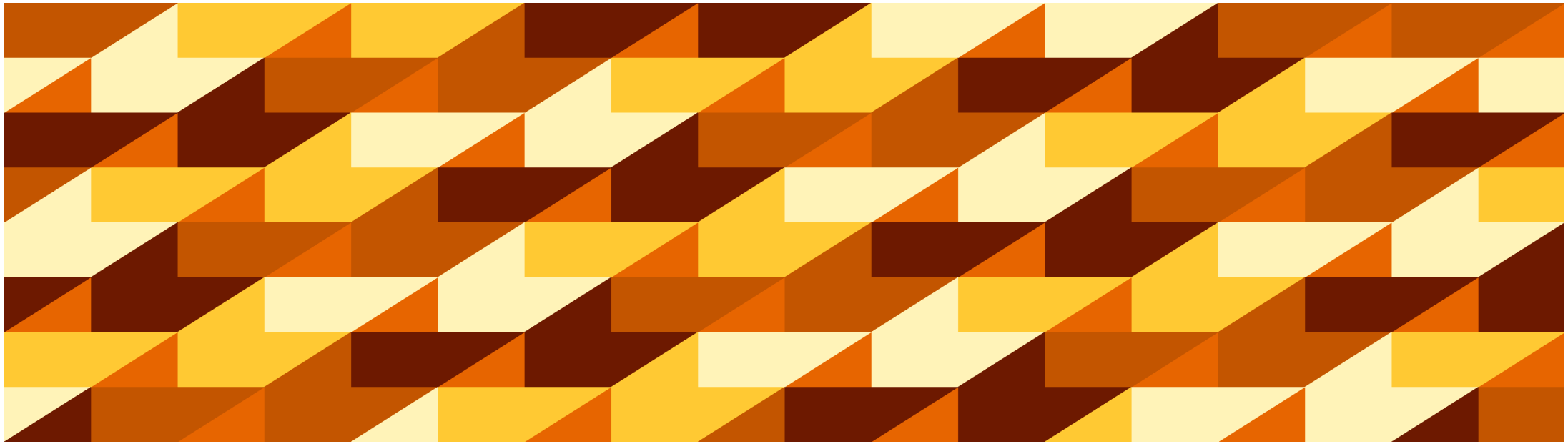
They have the following drawbacks:

- Limited to the existing classes in the labelled dataset and can only choose a label from the known classes
- Features that are directly extracted from product images are not precise enough

Towards more detailed auto-labelling mechanism with semantic feature extraction



Generating Minority Class Defect Detection Data using Self-Supervised Defect Generator



Generating Minority Class Defect Detection Data from Visual Inspection Dataset using Self-Supervised Defect Generator

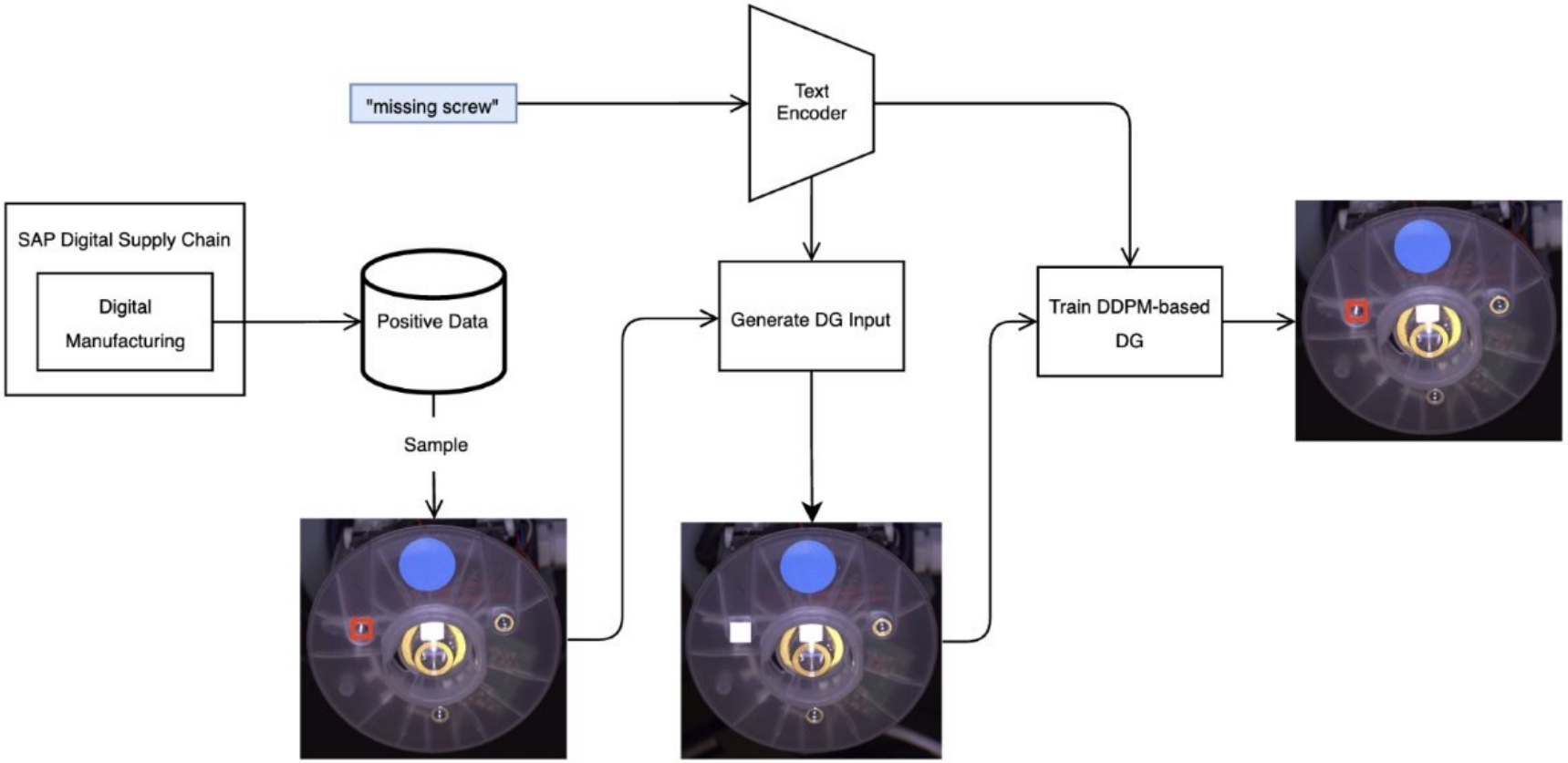
Motivation

In many cases of visual inspection, it is very difficult to have sufficient images showing defective products. In consequence, AI/ML models trained are prone to overfitting and unable to compete with human capabilities in identifying defective products. An extreme example is model may predict all data to be non-defective, which still results in a high accuracy.

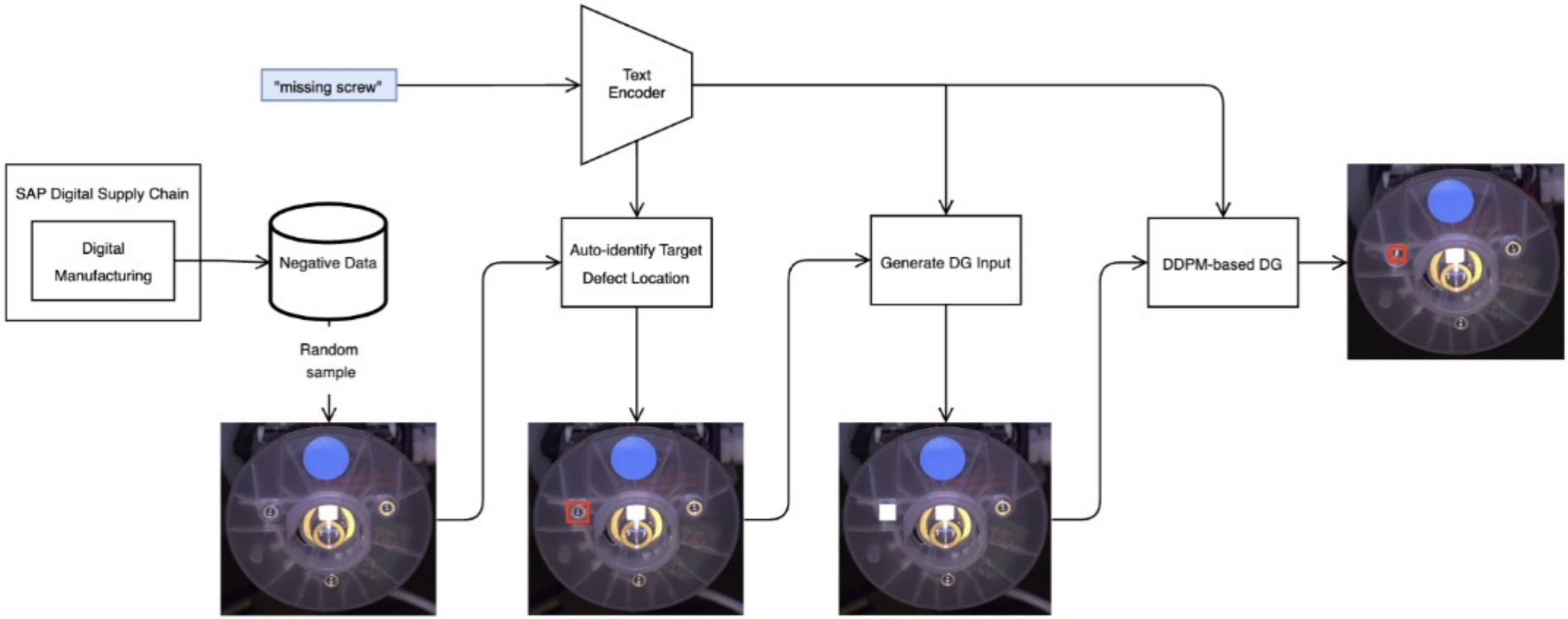
Challenges:

- VI data is very imbalanced

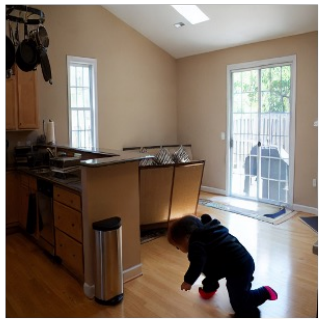
Generating Minority Class Defect Detection Data from Visual Inspection Dataset using Self-Supervised Defect Generator



Generating Minority Class Defect Detection Data from Visual Inspection Dataset using Self-Supervised Defect Generator

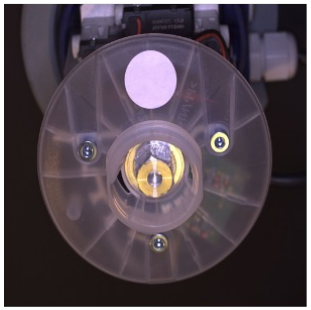


Generating Minority Class Defect Detection Data from Visual Inspection Dataset using Self-Supervised Defect Generator

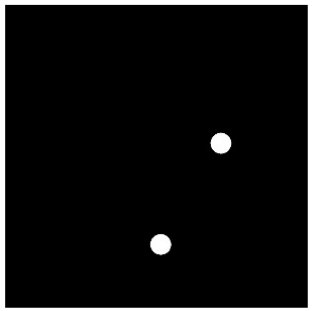
Original	Mask	Prompt	Result	Findings
	+ 	baby crawling on the floor		Performs well, mask bounding box can be used as pseudo-label for painted object.
		5 scratches on the wooden floor surface		Unable to paint defect (scratch)

Generating Minority Class Defect Detection Data from Visual Inspection Dataset using Self-Supervised Defect Generator

Original



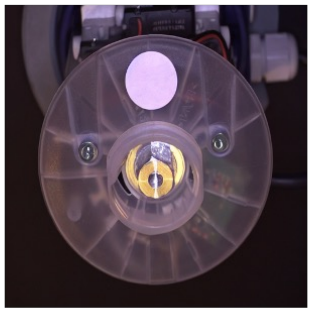
Mask



Prompt

two missing screws which are silver holes.

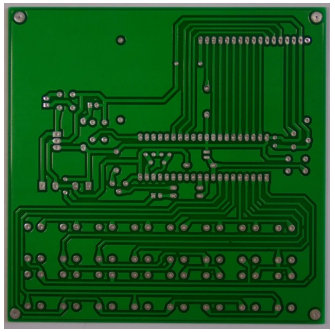
Result

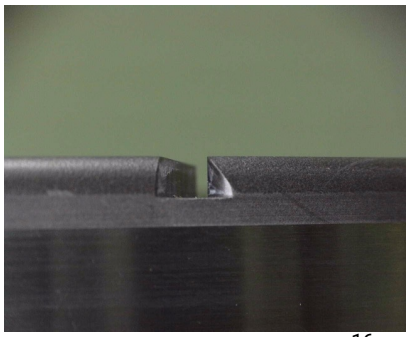


Findings

1. Does not perform well for small masks, paints based on surrounding instead of prompt
2. Unable to inpaint defect (missing screw)

- Working with the VISION dataset, a large and diverse visual inspection dataset.
- Finetune the image inpainting model to inpaint industrial defects and small objects to gain more control over image generation.





Thank you.

Contact information:

Xinyan Chen
xinyan.chen@sap.com