Chaofan ZHOU, Meiqin LIU, Senlin ZHANG, Ping WEI, Badong CHEN, 2023. A graph-based two-stage classification network for mobile screen defect inspection. *Frontiers of Information Technology & Electronic Engineering*, 24(2):203-216. https://doi.org/10.1631/FITEE.2200524

# A graph-based two-stage classification network for mobile screen defect inspection



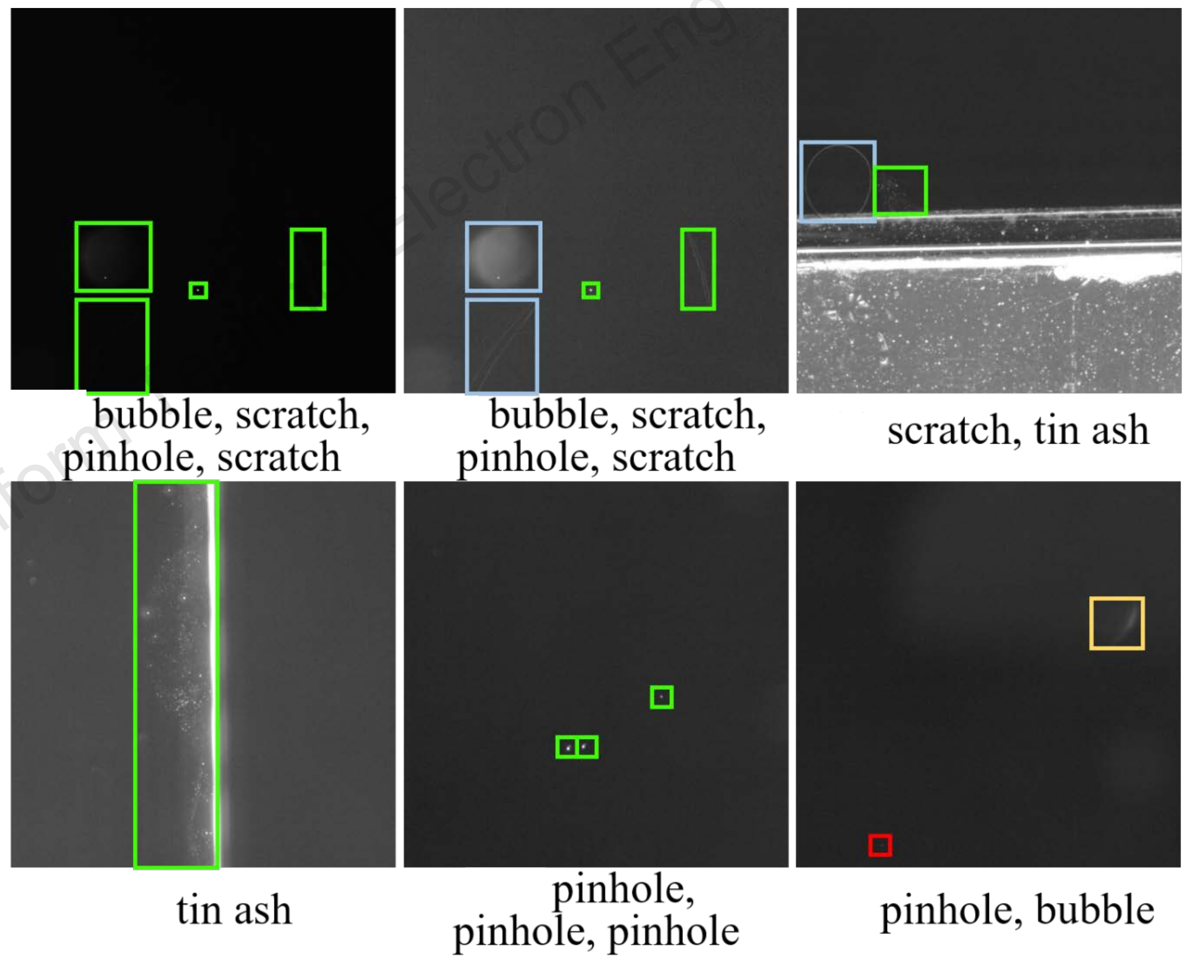
**Key words:** Graph-based methods; Multi-label classification; Mobile screen defects; Neural networks

Chaofan ZHOU
E-mail: zhouchaofan@zju.edu.cn
ORCID: https://orcid.org/0000-0003-0807-6539

Corresponding author: Meiqin LIU
E-mail: liumeiqin@zju.edu.cn
ORCID: https://orcid.org/0000-0003-0693-6574

# Motivation

- Accurate and efficient defect inspection is significant for mobile phone screen industrial processes.
- There are some challenging issues brought by the characteristics of screen defects.
  - The problem of interclass similarity and intraclass variation (as shown in blue boxes)
  - The difficulty in distinguishing low contrast, tiny-sized, or incomplete defects (as shown in red box, red box, yellow box)
  - The modeling of category dependencies for multi-label images

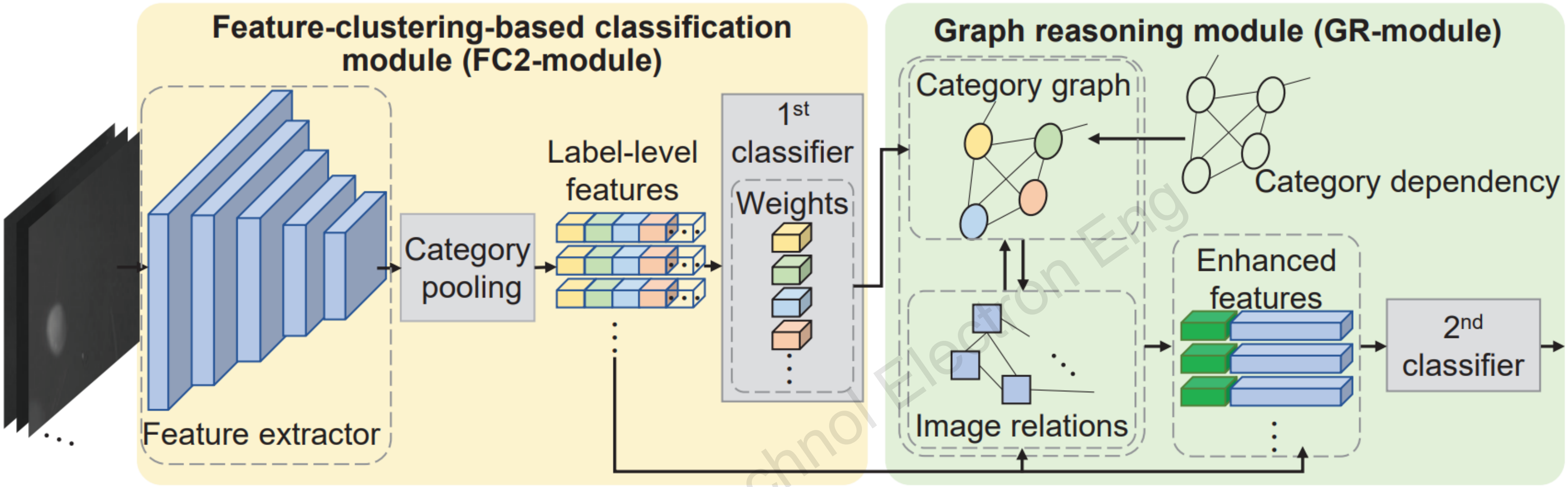



**Samples of mobile screen defects**

# Main idea

- A two-stage defect inspection network is proposed. which consists of two components, i.e., a feature-clustering-based classification module (FC2-module) and a graph reasoning module (GR-module).
- The graph in the GR-module enhances the image features with the help of category features to increase the interclass distance and the intraclass correlation in the feature space.
- Image-wise relations are applied to improve features of images that have low contrast, tiny-sized, or incomplete defects. When humans are not sure about the objects in the image being viewed, they refer to other images in which similar objects occur.
- We model the category dependencies using the graph convolution network (GCN) and the co-occurrence patterns of defects.

# Method




**The proposed two-stage network**

## Pipeline:

(1) The input images are sent to the feature-clustering-based classification module (FC2-module) to obtain label-level features and distilling weights.

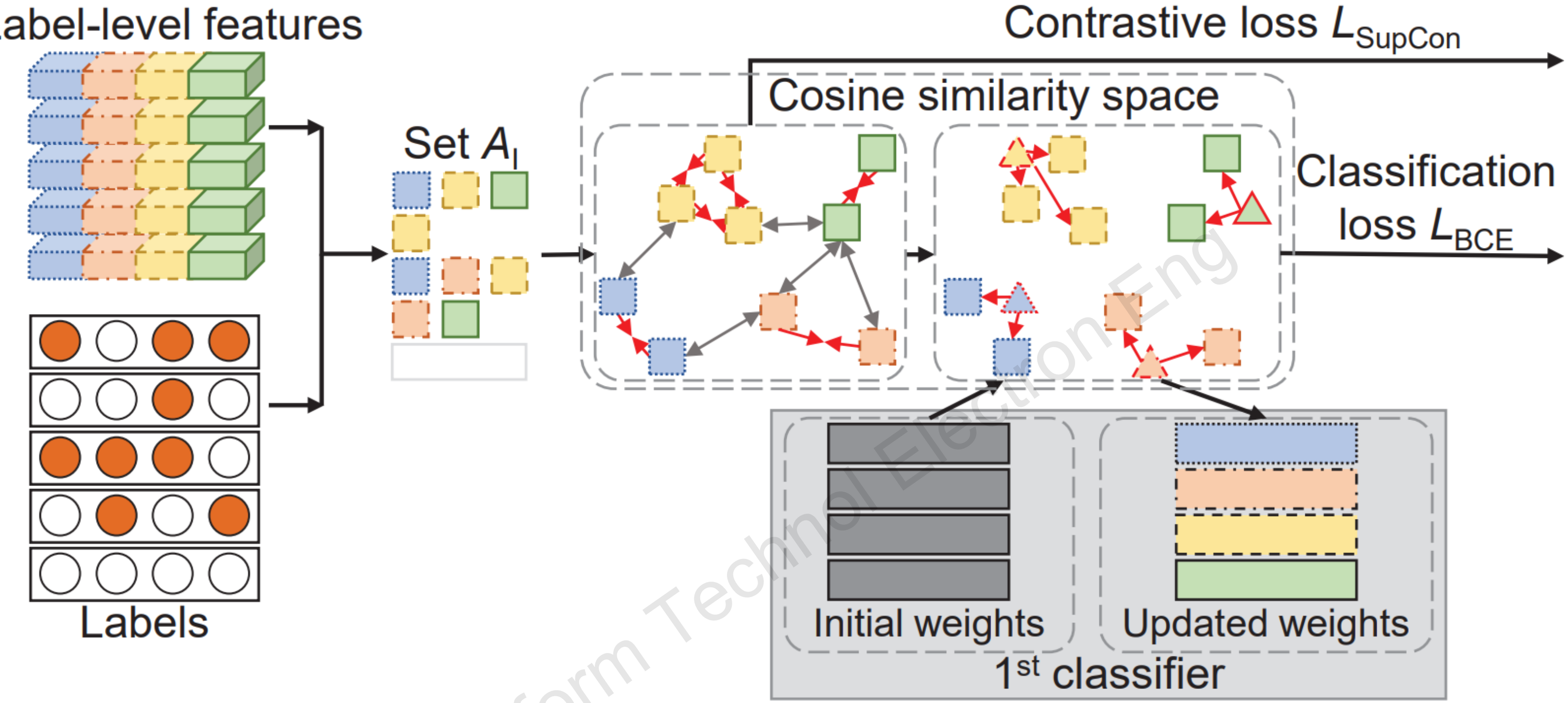
(2) Features and weights are fed into the graph reasoning module (GR-module) to enhance features.

(3) Enhanced features are used to improve the classification performance.

## Key point:

Based on graph, the GR-module exploits category-wise dependency, image-wise relations, and interactions between them. By doing so, the GR-module can incorporate category semantics and image relations to guide enhancement of image features.

# Method (Cont'd)




**The pipeline of updating the weights of the first classifier**



**Design reason：**

- To make graph better promote the model's performance, we use the distilling weights of the first classifier as the representations of category nodes.
- To make the weights of the classifier learn enough information about each defect category, the pipeline of updating the weights contains two parts:
  (1) contrastive learning, improved based on multi-label characteristics, which gathers label-level features with the same label;
  (2) a classifier that makes the weights represent the overall features of each cluster.

# Constructed dataset

**Table 1 Number of patches of each category in the training, validation, and testing sets**

| Category | Number of patches in the training set | Number of patches in the validation set | Number of patches in the testing set |
|---|---|---|---|
| Bubble | 91 | 21 | 33 |
| Pinhole | 1514 | 229 | 344 |
| Scratch | 1010 | 235 | 228 |
| Tin ash | 726 | 194 | 130 |
| Defect-free | 540 | 123 | 111 |
| Total patches | 2520 | 503 | 578 |



- The table shows the number of patches of each category in the training, validation, and testing sets.
- It is constructed by collecting image data from the real industrial pipeline.

# Major results

**Table 2 Comparisons with state-of-the-art methods on a multi-label screen defect classification task**

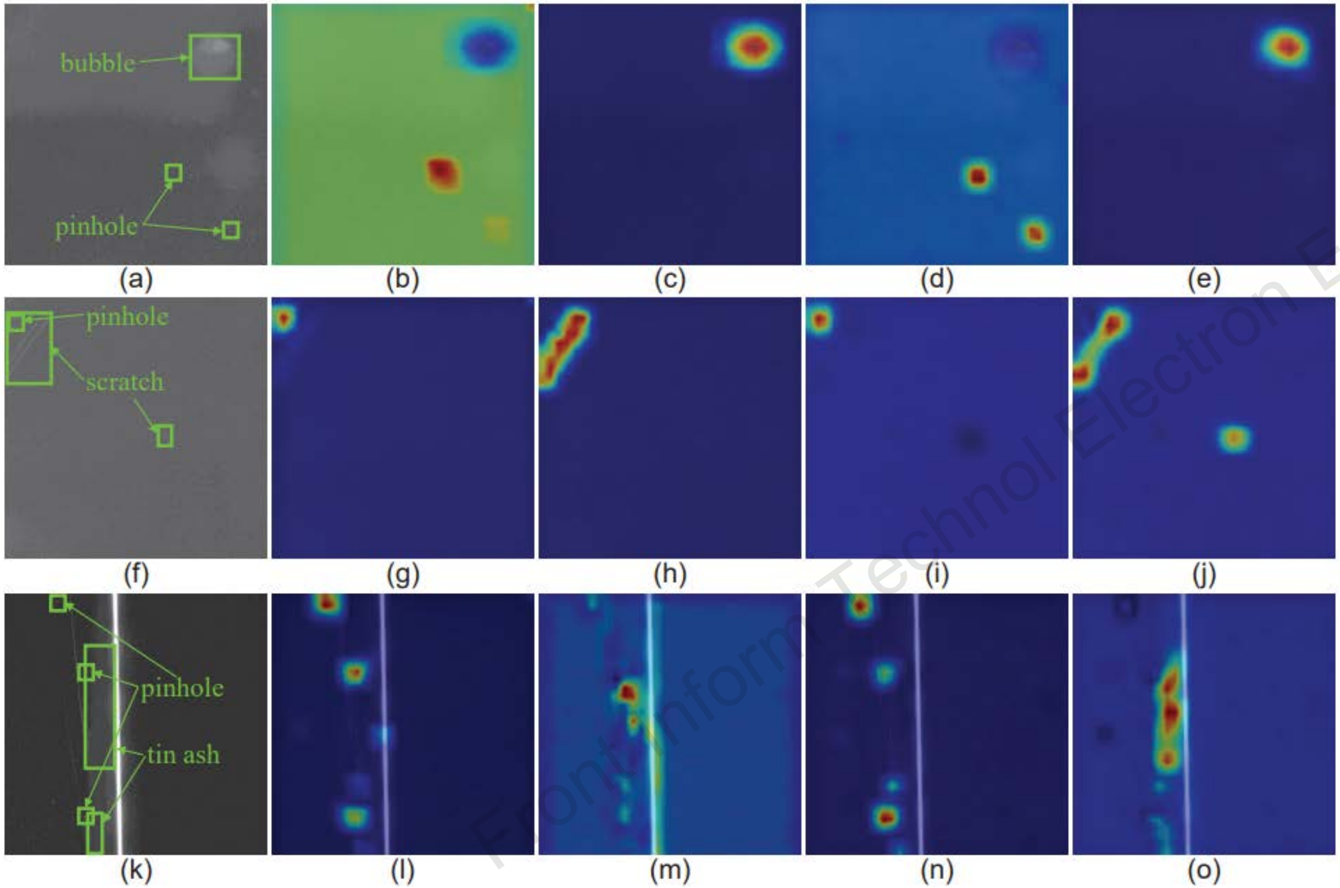
| Method | Accuracy | Precision | Recall | $F$-measure |
|---|---|---|---|---|
| ResNet-50 | 0.952 | 0.966 | 0.919 | 0.942 |
| VGG-16 | 0.967 | 0.963 | 0.827 | 0.890 |
| Inception-v3 | 0.971 | 0.963 | 0.953 | 0.958 |
| KSSNet | 0.953 | 0.828 | 0.713 | 0.767 |
| TDRG | 0.972 | 0.963 | 0.959 | 0.961 |
| SewerMl | 0.887 | 0.820 | 0.727 | 0.771 |
| Ours | **0.977** | **0.975** | **0.970** | **0.973** |

The results in bold are the highest



- Our approach achieves the highest accuracy (97.7%), precision (97.5%), recall (97.0%), and *F*-measure (97.3%) on the mobile screen defect dataset.
- Compared with the second place of each metric, our model outperforms with a margin of 0.5% on accuracy, 0.9% on precision, 1.1% on recall, and 1.2% on *F*-measure.
- These results prove that the proposed method is feasible.

# **Major results:** Network visualization




**Visualization for feature maps with a class activation map (CAM): the second and third columns show visualization results of ResNet-50, and the fourth and fifth columns exhibit those of our approach**

- Figs. (b) and (d): the maps of our approach can highlight regions of tiny-sized defects.
- Figs. (h) and (j): the maps of our approach can highlight regions of low contrast and incomplete defects.
- Figs. (m) and (o): the maps of our approach can highlight illegible regions caused by the characteristic of interclass similarity and intraclass variation.

# **Major results:** Generalization performance

**Table 9 Comparisons with the state-of-the-art methods on the steel surface defect dataset**

| Method | Accuracy | Precision | Recall | $F$-measure |
|---|---|---|---|---|
| ResNet-50 | **0.982** | 0.915 | 0.877 | 0.896 |
| VGG-16 | 0.974 | 0.850 | 0.848 | 0.849 |
| Inception-v3 | 0.979 | 0.912 | **0.928** | 0.920 |
| KSSNet | 0.953 | 0.839 | 0.617 | 0.711 |
| TDRG | 0.978 | 0.896 | 0.853 | 0.874 |
| SewerMl | 0.954 | 0.813 | 0.707 | 0.756 |
| Ours | **0.982** | **0.960** | 0.885 | **0.921** |

The results in bold are the highest



- Dataset: created by Severstal and is multi-label.
- Compared to other methods, our model achieves the best results on almost all metrics, which proves that it has good generalization capability.

# Conclusions

- Given the characteristics of mobile screen defect data from the real industrial pipelines and the difficulties in recognizing them, we proposed a two-stage defect inspection framework, consisting of the FC2-module and the GR-module.
- Our approach is mainly designed for three problems: (1) the problem of interclass similarity and intraclass variation, (2) the difficulty in distinguishing low contrast, tiny-sized, or incomplete defects, and (3) the modeling of category dependencies for multi-label images.
- The experimental results on the mobile screen defect dataset showed that our model has better defect inspection performances: 97.7% accuracy and 97.3% *F*-measure.
- Our approach can detect illegible defects and has good generalization capability.

**Chaofan ZHOU** received the B.E. degree in automation from Wuhan University, Wuhan, China, in 2020. She is currently working towards the Ph.D. degree in control theory and control engineering at Zhejiang University, Hangzhou, China. Her research interests include image processing, computer vision, and deep learning.



**Meiqin LIU** received the B. E. and Ph.D. degrees in control theory and control engineering from Central South University, Changsha, China, in 1994 and 1999, respectively. She was a Post-Doctoral Research Fellow with the Huazhong University of Science and Technology, Wuhan, China, from 1999 to 2001. She was a Visiting Scholar with the University of New Orleans, New Orleans, LA, USA, from 2008 to 2009. She was a Professor with the College of Electrical Engineering, Zhejiang University, Hangzhou, China, from 2001 to 2021. She is currently a Professor with the Institute of Artificial Intelligence and Robotics, Xi’an Jiaotong University, Xi’an, China, and also with the College of Electrical Engineering, Zhejiang University, Hangzhou, China. She has authored more than 200 papers in major journals and international conferences. Her current research interests include theory and application of artificial intelligence, multi-sensor networks, information fusion, and nonlinear systems.

**Senlin ZHANG** received the B.E. degree in control theory and control engineering from Wuhan University of Technology, Wuhan, China, and the M.E. degree in control theory and control engineering from Zhejiang University, Hangzhou, China, in 1984 and 1991, respectively. He is currently a Professor with the State Key Laboratory of Industrial Control Technology and the College of Electrical Engineering, Zhejiang University, Hangzhou, China. He has authored more than 200 papers in major journals and international conferences. His current research interests include unmanned undersea systems, intelligent systems, and multi-sensor networks.

**Ping WEI** received the B.E. and Ph.D. degrees from Xi’an Jiaotong University, Xi’an, China. From 2016 to 2017, he was a Postdoctoral Researcher with the Center for Vision, Cognition, Learning, and Autonomy (VCLA), University of California, Los Angeles (UCLA). He is currently an Associate Professor with the Institute of Artificial Intelligence and Robotics, Xi’an Jiaotong University. His research interests include computer vision, machine learning, and computational cognition.



**Badong CHEN** received the Ph.D. degree in computer science and technology from Tsinghua University, Beijing, China, in 2008. He is currently a Professor with the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an, China. His research interests include signal processing, machine learning, artificial intelligence and robotics. He has authored or coauthored over 200 articles in various journals and conference proceedings.