# Online, Real-time and Robust Detection and Localization of Foreign Objects on Paper Surface using Machine Vision and Clustering

Muhammad Nabeel Tahir[1], Zubair Khalid[2] and Adeem Aslam[3]
[1]*Department of Electrical and Computer Engineering, Rutgers University*, NJ, USA
[2]*School of Science and Engineering, Lahore University of Management Sciences*, Lahore, Pakistan
[3]*Department of Electrical Engineering, University of Engineering and Technology*, Lahore, Pakistan
nabeel.tahir@rutgers.edu, zubair.khalid@lums.edu.pk, adeem.aslam@uet.edu.pk

2022 28th International Conference on Mechatronics and Machine Vision in Practice (M2VIP) | 979-8-3503-3433-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/M2VIP55626.2022.10041095

***Abstract*—To serve the needs of surface inspection during the high-speed paper production process, this paper focuses on the design of the machine vision system and the development of associated algorithms for computationally efficient robust detection and localization of unwanted foreign objects. We address the practical challenges during the design of the architecture of an overall system that consists of cameras, illumination arrangement, and computing platforms. For the processing of images, we use machine vision for feature extraction followed by clustering for the quantification and localization of defects. We also conduct simulations and analyse detection error and localization error to evaluate the performance of the proposed system and algorithms. The proposed system has been deployed in the industry for surface inspection.**

***Index Terms*—Machine vision, surface inspection, paper production, clustering, feature extraction, detection, localization.**

## I. Introduction

With the advent of new technologies, many web or sheet manufacturing industries are shifting their manual or offline inspection processes for quality control to autonomous, robust, fast and efficient intelligent systems. The inspection process has found its applications in e.g. paper web, nonwoven, fabric, plastic, metal and wood industries [1], [2]. In these applications, during complex industrial process, defects such as unwanted objects, holes, and scratches arise at the surface of the product due to inconsistencies in the process design parameters, contamination of raw materials, or failure to follow operating procedures by the production staff. As a result, the total production cost (or efficiency) increases (or decreases) substantially as the removal of such defects requires extra human resources and time. In worst-case scenarios, the defective material needs to be disposed of which adds an extra toll to the production cost. In industries where any contamination or defect in the product can directly or indirectly affect organic life require an efficient inspection and defect detection system. In such industries, early detection of such defects can help the production team to reduce the cost by accurately re-calibrating the production parameters or instruments or improve the quality of raw materials to avoid the complete disposal of the final product.

This work was supported by Higher Education Commission of Pakistan under its National Research Program for Universities (NRPU).

A typical paper (or board) manufacturing process produces the webs with larger width and at higher machine speeds as compared to other processes which produce flat sheets or products. The paper manufacturing industry produces the paper with web widths ranging from $1-9m$ at speeds as high as $30\,\mathrm{m/s}$ $(1800\,\mathrm{meter-per-minute})$. In this context, we focus on the design and development of a machine vision system to detect and localize defects such as the presence of foreign objects on the paper surface during the production process while serving the quality requirements and addressing the practical constraints.

Although web inspection is a difficult task and requires extraordinary precision and accuracy, a number of techniques have been proposed to detect and localize these defects on the web surface using image processing. A co-occurrence matrix-based texture feature extraction and statistical self-organizing map-based method have been proposed by [3]. A stationary wavelets transform-based technique has been proposed in [4], which divides an image into sub-images and uses a fusion technique to segment the defects in an image. Other studies which use statistical signal processing methods such as Gaussian-Markov random field texture features extraction, singularity detection, and blob algorithm have been proposed in [5]–[7]. Apart from statistical signal processing-based algorithms, multiple techniques have been published based on machine vision algorithms [**?**], [**?**], [**?**], [8]–[10]. These algorithms were based on intelligent image processing and use color, geometric features, and texture to detect the defects on the surface.

Recently, the high-performance computing capability has enabled the use of artificial intelligence (AI) and deep learning in industrial developments and processes. Multiple deep-learning-based methods have been introduced to accurately detect and localize the defects on the web surface. Convolutional denoising autoencoder-based deep neural networks, generative adversarial network (GAN), and convolutional neural network (CNN) have been adopted in [11], [12] and [13] respectively. We also note the recent works which use different variants of CNN-based deep neural networks for surface inspection [14]–[18]. Since the methods based on deep learning or statistical signal processing are computationally intensive and require graphics processing based computing platforms, these techniques are impractical when real-time processing is of primary importance.

In this paper, we present the design of the machine vision system and develop associated algorithms for real-time, online (deployable on the machine) and robust detection of foreign objects on the surface for a high-speed paper production process. Our algorithms employ machine vision and unsupervised learning for the detection and localization of defects. We use a machine vision based algorithm for computationally efficient extraction of features from the acquired image. Once the features are extracted, we use clustering for the localization and quantification of defects. We also present the overall architecture of the system that addresses the practical challenges. Furthermore, we analyse detection error and localization error to evaluate the performance of the proposed system and algorithms. Numerical experiments reveal that the proposed system is capable of robust detection and localization of foreign objects. We have carried out this work in collaboration with the paper production industry where the proposed solution has been deployed.

The rest of the paper is organized as follows. We provide the problem description in Section II, where we also discuss the practical deployment challenges and mathematical formulation of the problem. After presenting the proposed system architecture and our approach for feature extraction, detection and localization in Section III, we discuss the practical challenges and analyse the performance of the proposed system in Section IV before making the concluding remarks in Section V.

## II. Problem Formulation

### A. Problem Under Consideration

Product quality control is an integral and most important part of an industrial production process. It aims at maintaining the quality during production and identification of defects if any in post-production to meet the standards set by the governing authorities. In the paper industry, a typical process uses the raw material in large volumes in either solid or semi-solid (thick liquid) form stored in large containers. Since ensuring the quality of such raw material and keeping it safe from external contamination is a very challenging task, the possible chances of contamination of raw material from unwanted foreign objects could arise at the time of mixing different materials or at the time of pressing.

We consider design and development of a (machine) vision-based system and associated algorithms for 1) online (deployed on machine), 2) real-time (efficient detection for instantaneous reporting) and 3) robust detection of foreign objects on the surface of the paper illuminated using front lighting.

### B. Mathematical Modeling

Once translated into an image space, an image with $S$ number of defects can be modeled as

$$\tilde{I}(x, y) = I(x, y) + \sum_{i=1}^{S} g(\alpha_i x - x_i, \beta_i y - y_i) + n(x, y), \quad (1)$$

where $I(x, y)$ represents the defect free image and $\tilde{I}(x, y)$ represents the image which is contaminated by $S$ number defects, each represented by $g(*)$ and has a different size, characterized by $\alpha_i$ and $\beta_i$), and different location given $x_i$ and $y_i$ for $i = 1, 2 \ldots, S$. Here $n(x, y)$ represents the noise introduced by the environment or conversion of a $3D$ scene into the image space. It is assumed that the noise introduced in the image is a Gaussian noise of mean $\mu$ and standard deviation $\sigma$.

The problem under consideration is to detect $S$ number of foreign objects (or defects) and localize each of the objects, that is, find the location $x_i, y_i$ and scaling coefficients $\alpha_i, \beta_i$ for $i = 1, 2 \ldots, S$.

### C. Practical Deployment, Challenges and Constraints

For the problem formulated above, we design the machine vision system with the consideration of deployment at the Paper Mill of Bulleh Shah Packages (largest paper board producer of the region), where paper of different widths ranging from $1.5\,m$ to $1.6\,m$ is produced at the machine speed of 400 meters per minute. To cover the width range and identify the defects as small as $1\,\mathrm{mm} \times 1\,\mathrm{mm}$, we use two cameras with resolution $3\,\mathrm{MP}$ each along with four light sources. The 3MP cameras provide the best resolution under the light conditions present at the deployment site. The lower resolution cameras did not provide satisfactory results while higher resolution cameras increased the processing time. The advantage of using the two cameras is that it allows the system to scan smaller area with better resolution and accuracy for defect identification. In addition to ensure the robustness of detection and localization during the development of system and processing algorithms, we address the practical constraints such as 1) maintaining uniform and sufficient illumination of the surface using external light sources, 2) real- time processing of high resolution video stream, and 3) transfer of high-resolution video without network congestion or video frame corruption. The system was deployed on a machine with speed 400 meters per minute. The speed of the machine varied from 0-400-0 meters per minute as one role of paper web is processed.

## III. Proposed Machine Vision System and Image Processing Algorithms

We design the machine vision system with architecture depicted in Fig. 1. The video streams from the two cameras are transmitted to a high performance computing system installed near the machine using gigabit power over ethernet (PoE) switches. Each frame grabbed from a camera is transmitted to a separate thread of our application dedicated to process the video stream of either left or right camera. The raw video frame in each thread is first pre-processed to remove the noise elements and then passed to the detection algorithm stage for feature extraction, defect detection and defect localization. Once a defect has been detected, a signal is generated along to alert the machine operator staff and related information is displayed in the user interface of the developed application.

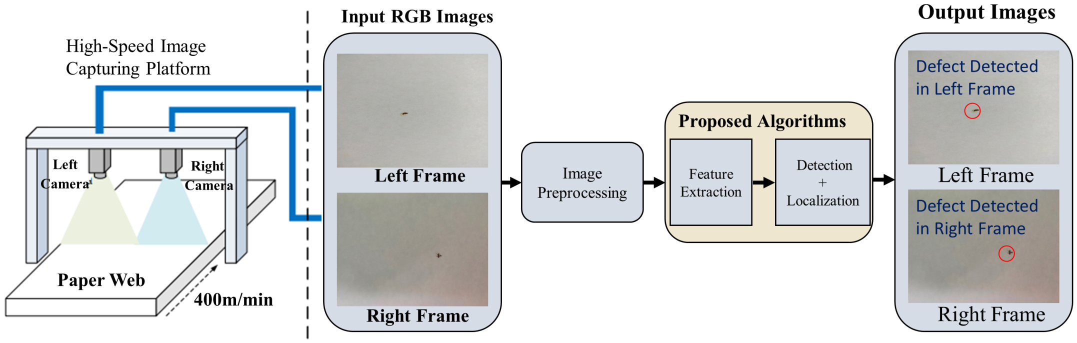
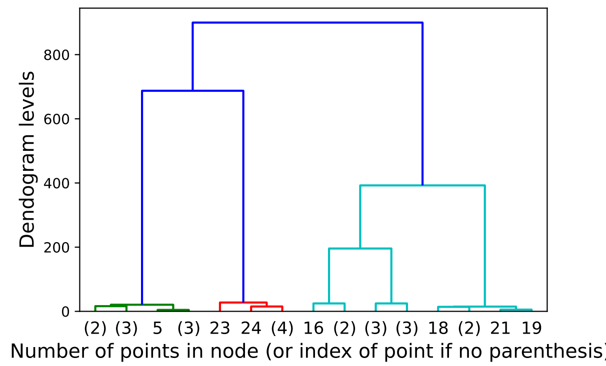



Fig. 1: System Architecture of Proposed Defect Detection System.

(a)




(b) Dendrogram for Hierarchical Clustering

Fig. 2: (a) Image with foreign objects the and the detected features, and b) associated dendrogram used for clustering of features and localization of defects.

## A. Pre-Processing

We introduce the pre-processing stage to remove the unwanted frequency and noise components introduced in the image acquired from the camera due to light fluctuations and hardware processing (sensors noise, quantization noise) as we are using commercial of-the-shelf low resolution video cameras and a single camera is covering a significantly large area of the paper. In the pre-processing stage, we first convert an image to a gray scale image from a coloured image and normalized to pixel range 0-255. The normalization technique improves the dynamic range of the image signal which increases the consistency in the image pixels by stretching the range of pixel intensity values. Second, we apply Gaussian smoothing to carry out low pass filtering using a zero mean discrete Gaussian filter of dimensions (11,11) which performs a symmetric smoothing in all directions to ensure bias-free edge detection in a particular direction. We note that the Gaussian smoothing also decreases the probability of false detection of defects.

## B. Feature Extraction Algorithm

The computer vision based detection techniques use unique set of features (edges, corners) as building blocks to construct or identify the desired objects. The biggest challenge in the development of such detection algorithms is to keep the computational cost low to ensure the detection or localization of the required set of features in real-time. We use corners in an image as features and employ the FAST corner extraction algorithm proposed in [19] as it outperforms its alternatives in terms of real-time image processing and detection accuracy in

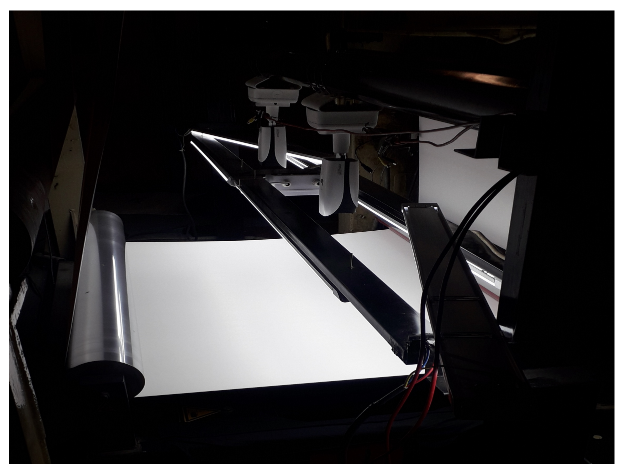
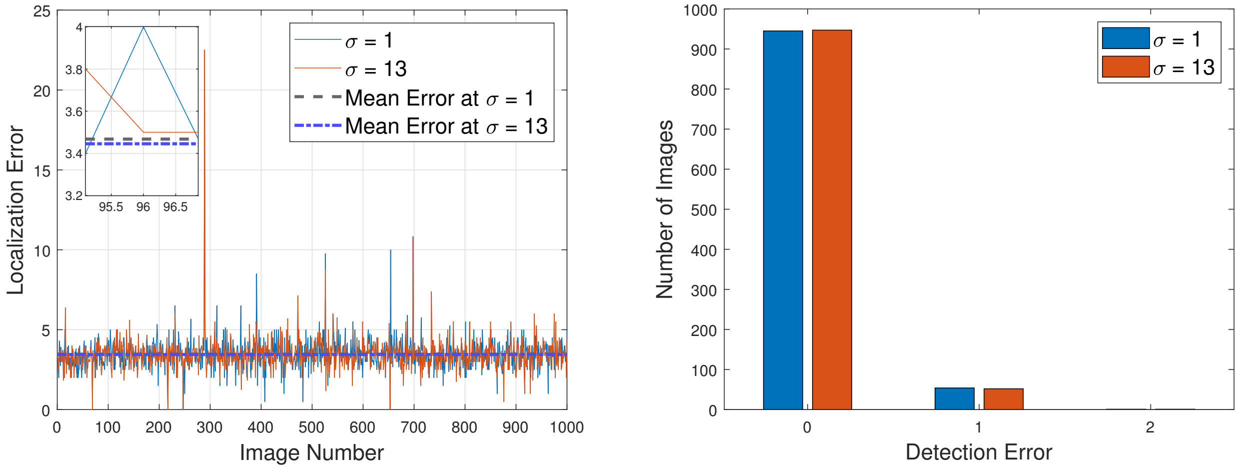


Fig. 3: Deployment of Proposed System at Paper Mill.




(a) (b)

Fig. 4: (a) Localization error against image index and (b) distribution of detection error for different values of variances of noise.

the low light and noisy conditions.

In order to find that a pixel $\mathbf{p}$ located at $(x, y)$ with intensity $\tilde{I}_p(x, y)$ is a corner/feature or not, we follow the following steps:

1) Select the pixel $\mathbf{p}$ as a target pixel.
2) Select an appropriate value of threshold $\Delta I$.
3) Select a Bresenham circle of radius 3 total 16 pixels around the point $\mathbf{p}$.
4) The selected pixel can be classified as a feature if there exist $P$ contiguous pixels in Bresenham circle which are all either darker than $\tilde{I}_p(x, y) - \Delta I$, or all brighter than $\tilde{I}_p(x, y) + \Delta I$. We use the default value for $P = 12$ reported in [19].
5) The computational cost of the algorithm is further reduced by comparing the value of target pixel with values of pixels $\{1, 3, 5, 9, 13\}$. If at least three of these pixel points satisfy the threshold criteria discussed in Step 4, the point $\mathbf{p}$ can be considered as a potential feature.
6) If the points $\{1, 3, 5, 9, 13\}$ do not satisfy the threshold criteria then $\mathbf{p}$ is not a potential feature. In case the pixel $\mathbf{p}$ satisfy the criteria, test all the pixels inside circle and check if at least $P$ contiguous pixels follow the criterion.
7) Repeat the procedure for all the pixels in the image.

### *C. Defect Detection and Localization using Clustering*

We use the identified features to accurately detect and localize the number of defects (foreign objects) present in an image. Given an image with multiple features associated with a single object, the algorithm can detect these features as illustrated in figure Fig. 2(a). To identify the exact location, taken as the center of the object, it is important to correctly assign the identified features to that object followed by the use of assigned feature points as reference to measure the location of the object. Since the features associated with each object can be considered as the smaller cluster of points centered around a common point, we can use clustering algorithm to find the center of these cluster thus providing the estimated location of an object. The number of clusters present in an image correspond to the number of objects and the centers of these clusters corresponds to the locations of the objects in the image.

We employ hierarchical clustering with Euclidean distance metric as decision criterion to cluster the features into defects. We obtain a cluster tree (dendrogram, illustrated in Fig. 2(b)), use agglomerative [20] clustering approach, also termed as bottom-up, and use single-linkage metric given by

$$\mathcal{L}(l, m) = \underset{\mathbf{f}_i \in c_\ell, \mathbf{f}_j \in c_m}{\operatorname{argmin}} \ \operatorname{dist}(\mathbf{f}_i, \mathbf{f}_j), \tag{2}$$

as decision criterion. In (2), $\mathcal{L}(l, m)$ denotes the minimum distance between the two clusters and $\mathbf{f}_i$ and $\mathbf{f}_j$ are the points belonging to clusters $c_\ell$ and $c_m$ respectively. The technique adopted here for clustering features into defects starts by considering each data point as a cluster and then computes the minimum distance between clusters. If the distance between two clusters is less than threshold $\Delta t$, the clusters are merged together. The value of $\Delta t$ was selected by trial and error method and later was hard coded. This process is repeated until all the clusters are farthest than the minimum threshold or all the clusters are merged to one cluster.

## IV. Analysis and Evaluation

### *A. Deployment and Addressing Practical Challenges*

We develop a stand alone system designed to capture the images of paper surface, process the images using proposed algorithms and display the information on a custom built software. To communicate with the existing network or relay information to control or production, we use the industrial Modbus communication protocol. The deployment of the proposed system in the industry is depicted in Fig. 3. We have ensured the uniform illumination of the surface by first measuring the light intensity at different positions of target using light intensity meter and then installing the uniform white lights over the camera platform. To avoid the delay in communication of video frames from camera to the computing platform (i7 8th Gen, 16 GB RAM), we use a local gigabit PoE communication network. High speed object capture using commercial cameras require appropriate selection of camera frames per second (FPS) and shutter speed. We use $\text{Shutter Speed} = \frac{\text{Machine Speed in m/s}}{\text{Allowed Motion Blur in mm}}$ and $\text{FPS} = \frac{\text{Machine Speed in m/s}}{\text{Length Covered in a Frame in m}}$ to set the shutter speed and FPS equal to $\frac{1}{4000}\text{s}$ and $40\,\text{FPS}$ respectively.

### *B. Accuracy Metrics and Evaluations*

We evaluate the the performance of the proposed system and associated algorithms by evaluating the the detection error and the localization error. We find the overall performance of the system to be satisfactory since the system has been successfully deployed in the industry. We conduct numerical experiments to analyse the performance of our proposed system. We generate $K = 1000$ images using the model given in (1). For each image, we randomly choose the number of defects, $S \in [1, 5]$, their locations and sizes. Then we add Gaussian noise with mean $\mu = 0$ and variances $\sigma = [1, 13]$ respectively. The analysis was performed for sigma values $1 - 13$, but to show the performance of variance only extreme values were chosen. We process each image using the proposed system architecture and algorithms. For each image, we compute the detection error defined as

$$E_{\text{detection}}(k) = |S_k - \tilde{S}_k|, \tag{3}$$

where $S_k$ and $\tilde{S}_k$ denote the actual number of objects present in the image and number of objects detected by the proposed system. The mean detection error captures the performance of the object detection module of the proposed system. In order to evaluate the localization accuracy, we define the localization error given by

$$E_{\text{Loc}}(k) = \frac{1}{\tilde{S}_k} \sum_{i=1}^{\tilde{S}_k} \|\mathbf{z}_{ki} - \tilde{\mathbf{z}}_{ki}\|_2, \tag{4}$$

where $\mathbf{z}_{ki}$ denote the actual location, $\tilde{\mathbf{z}}_{ki}$ denote the estimated location (center of the cluster $i$) of the object $i$ in an image $k$ and $\|\cdot\|_2$ represents the Euclidean norm. We plot the localization error against image index and the distribution of detection error for different values of noise variances in Fig. 4(a) and Fig. 4(b) respectively, which demonstrate the robustness of the proposed algorithms. The localization error for $\sigma = 1$ is slightly higher than for $\sigma = 13$ and the difference is negligible. The Gaussian filter applied in pre-processing stage reduces the variance in pixel values introduced by adding the noise for different $\sigma$ values. And as a result the localization error does not change by changing the variance of noise introduced as we are removing the similar noise in pre-processing stage.

## V. Conclusion

In this paper, we have developed a technique using machine vision and unsupervised machine learning for real-time defect detection and localization to support paper production industries in improving and monitoring product quality. The proposed system efficiently detects and localizes unwanted foreign objects in an image for a wide range of object sizes. During the development of the solution, multiple practical challenges related to high-resolution image processing, low light image acquisition, and computational resources have been provided. The proposed algorithm has the capability to accurately detect and localize multiple objects in real-time at different levels of noise signatures. In addition, the developed system deployed at partner industry follows the standard industrial protocols and serves as an integral paper web inspection system. In conclusion, a robust, efficient web inspection system has been designed, developed and deployed with the capacity to report the issues and defects in real-time for local industry. In the future, we plan to extend the scope of the system by adding detect classification features not only for the paper web inspection but for fabric, steel, and plywood industries.

## References


[1] J. E. Graf, S. T. Enright, and S. I. Shapiro, "Automated web inspection ensures highest quality nonwovens," *Tappi journal (USA)*, 1995.

[2] L. Makelin, J. Moisio, and J. Jarvinen, "From hole detection to image processing and information technology," in *Proceedings of 50th Appita Annual General Conference*, vol. 1, 1996, pp. 817–822.

[3] J. Iivarinen, K. Heikkinen, J. Rauhamaa, P. Vuorimaa, and A. Visa, "A defect detection scheme for web surface inspection," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 14, no. 06, pp. 735–755, 2000.

[4] Z. Wang, "Web surface defect segmentation based on stationary wavelet transform," in *Materials Science and Information Technology*, ser. Advanced Materials Research, vol. 433. Trans Tech Publications Ltd, 2 2012, pp. 426–431.

[5] X. Huang, J. Dong, and M. Wang, "Paper web defection segmentation using gauss-markov random field texture features," in *2011 International Conference on Image Analysis and Signal Processing*, 2011, pp. 167–170.

[6] S. Qiu and J. Sun, "Web surface defect inspection based on singularity detection," in *2006 IEEE International Conference on Information Acquisition*, 2006, pp. 1364–1368.

[7] X. Qi, X. Li, and H. Zhang, "Research of paper surface defects detection system based on blob algorithm," in *2013 IEEE International Conference on Information and Automation (ICIA)*, 2013, pp. 694–698.

[8] Ç. Aytekin, Y. Rezaeitabar, S. Dogru, and I. Ulusoy, "Railway fastener inspection by real-time machine vision," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 7, pp. 1101–1107, 2015.

[9] H.-D. Lin and H.-L. Chen, "Automated visual fault inspection of optical elements using machine vision technologies," *Journal of Applied Engineering Science*, vol. 16, no. 4, pp. 447–453, 2018.

[10] R. Shanmugamani, M. Sadique, and B. Ramamoorthy, "Detection and classification of surface defects of gun barrels using computer vision and machine learning," *Measurement*, vol. 60, pp. 222–230, 2015.

[11] S. Mei, Y. Wang, and G. Wen, "Automatic fabric defect detection with a multi-scale convolutional denoising autoencoder network model," *Sensors*, vol. 18, no. 4, p. 1064, 2018.

[12] J. Heger, G. Desai, and M. Z. El Abdine, "Anomaly detection in formed sheet metals using convolutional autoencoders," *Procedia CIRP*, vol. 93, pp. 1281–1285, 2020.

[13] H. Zhang, Z. Chen, C. Zhang, J. Xi, and X. Le, "Weld defect detection based on deep learning method," in *2019 IEEE 15th international conference on automation science and engineering (CASE)*. IEEE, 2019, pp. 1574–1579.

[14] S. Arikan, K. Varanasi, and D. Stricker, "Surface defect classification in real-time using convolutional neural networks," *arXiv preprint arXiv:1904.04671*, 2019.

[15] F. Zhou, G. Liu, F. Xu, and H. Deng, "A generic automated surface defect detection based on a bilinear model," *Applied Sciences*, vol. 9, no. 15, p. 3159, 2019.

[16] L. Xiao, B. Wu, and Y. Hu, "Surface defect detection using image pyramid," *IEEE Sensors Journal*, vol. 20, no. 13, pp. 7181–7188, 2020.

[17] C.-Y. Tsai and H.-W. Chen, "Surfnetv2: An improved real-time surfnet and its applications to defect recognition of calcium silicate boards," *Sensors*, vol. 20, no. 16, 2020.

[18] B. Hu, B. Gao, W. L. Woo, L. Ruan, J. Jin, Y. Yang, and Y. Yu, "A lightweight spatial and temporal multi-feature fusion network for defect detection," *IEEE Transactions on Image Processing*, vol. 30, pp. 472–486, 2020.

[19] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Computer Vision – ECCV 2006*, A. Leonardis, H. Bischof, and A. Pinz, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 430–443.

[20] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, p. 264–323, Sep. 1999.