



## Research paper

# Automatic Visual Inspection System based on Image Processing and Neural Network for Quality Control of Sandwich Panel

Vahid Torkzadeh and Saeed Toosizadeh\*

Department of Computer Engineering, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran.

## Article Info

### Article History:

Received 24 July 2021

Revised 12 December 2021

Accepted 21 February 2022

DOI:10.22044/jadm.2022.11002.2247

### Keywords:

Sandwich Panel, Dipping, Buckling, Image Processing, Convolutional Neural Network.

\*Corresponding author:  
s.toosi@mshdiau.ac.ir (S. Toosizadeh).

## Abstract

In this work, an automatic system based on the image processing methods using the features based on convolutional neural networks is proposed in order to detect the degree of possible dipping and buckling on the sandwich panel surface by a colour camera. The proposed method, by receiving an image of the sandwich panel, can detect the dipping and buckling of its surface with an acceptable accuracy. After a panel is fully processed by the system, an image output is generated in order to observe the surface status of the sandwich panel so that the supervisor of the production line can better detect any potential defects at the surface of the produced panels. An accurate solution is also provided in order to measure the amount of available distortion (depth or height of dipping and buckling) on the sandwich panels without the need for expensive and complex equipment and hardware.

## 1. Introduction

Today, with the improvement of imaging equipment and image processing algorithms, a new branch in the quality control and precision tools has emerged. Every day, we observe the use of advanced imaging systems for size measurement, calibration, mechanical joints control, production quality growth, etc. Using the image processing techniques can fundamentally change the product lines.

The sandwich panels are of great importance in the production of fixed and mobile Conex, stable prefabricated buildings, and prefabricated baths and toilets. Compared to the other industries, there are unique challenges in the design of an automatic control system for this industry, among which, are a variety of colors and patterns produced in the industry as well as the varying production line's settings and parameters depending on the customer's order. As a result, defining a set of fixed contracts and settings in which the production line works optimally and with a good quality is practically impossible. Further, unlike most visual automatic control systems that have the task of quality control in

discrete productions (e.g. car production lines and foodstuffs), the production line in this industry is continuous, and the product is constantly generated with different dimensions and features (based on the customers' order), by performing different operations at separate stations (such as forming and injection). A few example cases that require quality control and inspection in the sandwich panel production process are as follow:

- Color and thickness, efflorescence, dipping, and fracture
- No dipping and buckling on the surface of sandwich panel
- Alignment of the bottom and top sheets of the sandwich panel
- Conformity templates width with order.

Due to the wide range of variables affecting the sandwich panel production on the one hand, and the lack of a comprehensive study on the feasibility of using computer visual methods and artificial intelligence to control the quality of sandwich panel on the other hand, the present work only focuses on detecting the presence of

dipping and fractures on the produced sandwich panels as the first step for designing and producing a comprehensive automatic control system. The purpose of this work is to present an automatic system based on the machine vision and image processing methods that can properly detect and locate the defects within a sandwich panel. In this work, by the term “deficiency” in the sandwich panel production industry, we are referring to finding undesirable dipping and buckling in the production line.

In general, the quality control in the industry is performed in several different ways: 1) Manually, by human labor controlling the entire production process. 2) Automatically, by a computer-based system that monitors the production process often by using the digital cameras or different sensors. 3) Semi-automatically, by human forces interacting with a computer-based defects detection system [1-3].

The computer vision-based automatic detection methods can be either reference-based or non-reference-based. In the reference-based method, both the specific sample of defects and flawless samples are modelled. In the case of non-reference methods, there is no prior knowledge of the defective samples, and the only available knowledge is of perfect and flawless samples.

In the non-reference methods, the anomalies are those cases that do not match the expected pattern. From the machine learning perspective, the reference-based methods are implemented by a multi-class classifier, while the non-reference methods are implemented by a single-class classifier, which holds the normal (flawless) data in the training phase and applies the normal and non-normal data in the test phase.

In this work, a non-reference method is presented in order to detect and locate the anomalies in the sandwich panel. The proposed method is region-based, i.e. the extent to which a sub-region is abnormal is measured in comparison to a reference dictionary of normal sub-regions derived from the training images. The anomaly of a sub-region is calculated using the visual similarity between the test sub-region and a normal sub-region of the reference dictionary set. The lower the similarity between the two sub-regions, the greater the anomaly of that sub-region. Further, the degree of visual similarity between two sub-regions is the Euclidean distance between the feature vectors extracted from the two intended sub-regions. Finally, a sub-region is classified as abnormal if the degree of its anomaly is greater than the threshold obtained by estimating the similarity distribution between the

normal sub-regions. Since the proposed method is region-based, it allows for the simultaneous detecting and locating of the defects.

The novelties of this work can be described in three general terms:

1. Unlike the existing methods, this work does not require any prior knowledge of the defective samples, and has no dependence on the type of defects found in the sandwich panels. This is a very useful and important point because collecting such data is so costly, time-consuming, and sometimes impossible.
2. The proposed method does not depend on the color or pattern used in the sandwich panels.
3. The proposed method can properly estimate the distortion within the sheets of the sandwich panels (the depth or the height of dipping and buckling).
4. Further, the proposed solution is very low-cost, robust, and flexible to be set up and used in an existing production line, and it has been already deployed successfully in two sandwich panel production lines of a factory.

## 2. Literature Review

Defect detection is a very extensive issue in various research areas such as computer vision, signal processing, and communication [4], and it has been investigated in various applications such as industrial damage detection [1], product texture defect detection [5], medical defect detection [6], and textual defect detection [7].

In general, the anomaly detection methods can be divided into five general categories: the probability-based, distance-based, reconstruction-based, domain-based, and information theory-based methods.

The probability-based methods assume that the low-aggregated regions in the training dataset are less likely to contain a normal sample. Therefore, these methods try to estimate the density of the normal class, and use it in order to estimate the probability of normality or abnormality of an input test data.

The distance-based methods assume that the "normal" data is clustered precisely, while the unusual data is far from such "normal" data. These methods require an accurate technique in order to measure the similarity between two data samples. These methods utilize the Nearest Neighbourhood (NN) algorithms [8, 9] and the data clustering techniques.

The reconstruction-based methods consider a regression model that is trained from the normal data. In this case, the unusual patterns are

identified by estimating the error between the actual value of the regression and the regression output obtained with the "normal" data during the training process.

The domain-based methods, similar to the distance-based methods, try to describe a domain containing the "normal" data by defining a boundary around the "normal" class. In such cases, the anomaly detection is determined by whether a test pattern is in the "normal" class.

Finally, the last set of methods are based on the information theory. These methods calculate various metrics of the information theory such as entropy or Kolmogorov complexity for the normal images. They measure how much a test data has changed the content the information of a normal image. Refer to [10] for more detailed information in this area.

If one considers the problem of defect detection in foam sheets, it can be said that there is a high economic importance to detect the physical problems of foam plates including curvature or dipping, especially in the early stages of production. In the recent years, the methods have been proposed to detect surface defects of the products in the industrial and commercial applications by employing the image processing and computer vision techniques. A number of these proposed methods are based on the statistical properties. For example, the statistical histogram features have been employed in order to identify surface failures of steel strips [11]. The other statistical methods used include the co-occurrence matrix of grey surfaces [12, 13] and morphological operations [14-17]. Further, an approach based on multivariate image analysis, which itself uses multivariate statistical analysis as the basis, has been proposed to detect the metal surface failures [18]. Usually, the statistical methods determine the presence or absence of failure but provide no information about its location.

On the other hand, there are a number of proposed frequency-based methods that can be used, for example, Fourier analysis [19, 20] and Gabor conversion [21-23]. Further, a group of frequency-domain-based methods exist that use the wavelet analysis in order to identify the failures and surface cracks [24-26].

The important point in the frequency domain-based methods is that sometimes they require very complex conversions from the location domain to the frequency domain.

Another set of proposed methods are the model-based approaches that specifically use fractal geometry [24-26]. In these methods, it is

relatively difficult to design and construct the models due to the lack of a standard way of modelling.

Also another set of proposed methods can be mentioned that tackle this problem by applying the machine learning algorithms and artificial intelligence as well as their combination with the previous methods. In [27], an expert fuzzy system consisting of six stages has been developed in order to evaluate the surface quality of steel strips. In [28], a control system for detecting metal surface failures has been proposed, which is specifically capable of detecting cracks and gaps on aluminum surfaces. Jia *et al.* [29] have developed another system to detect surface failures of steel plates. This system, which runs on a real-time basis and is based on support vector machines (SVM), is able to accurately detect 94.4% of the failures. In another approach, using the energy and entropy properties calculated by expanding matching wavelet package and employing a neural network, a system has been designed to identify and classify various failures of stainless-steel sheets [30]. The surface failure detection is not limited to the metal production industry, and is applicable to other industries including textile. For example, in [31], a system has been proposed that is capable of identifying and classifying fabric failures using support vector machines and Principal Component Analysis (PCA). It can also be used in the stone and ceramic industry. In a proposed system, defects on ceramic tiles surfaces are identified using the simple image processing techniques such as edge detection and morphological operations [32].

### 3. Proposed Methods

In this work, a system is designed that can detect the anomalies in the sandwich panels such as dipping and buckling. The system performance is designed according to the factory production line routine, in which a sample sandwich panel without anomaly is initially produced and approved by a quality supervisor. The prototype-produced sandwich panel is considered as the reference sample. Then the factory production line produces a large number of sandwich panels similar to the prototype. At this stage, there is no need for a human supervisor, and the system automatically detects the abnormal regions using the trained model based on the reference sample. Finally, the identified regions are presented as a labelled image to the quality supervisor. The advantage of this system is that there is no need for any changes in the production line, and no

costly hardware and tools to operate it. A simple color camera is also used to set up the system.

### 3.1. Formal definition of problem

Having a frame image  $I$  of the sandwich panel with the  $w \times h \times d$  dimensions, from the values of natural numbers  $w, h, d$ , where  $w$  and  $h$  are the image resolution and  $d \in \{1,3\}$  is the number of channels in the colour image (3 channels in RGB colour images), the quality control issue of the sandwich panel can be formally defined as follows:

Assuming the variable  $\mathcal{A}_I$  as a binary map of frame  $I$  with the  $w \times h$  dimensions, each point denoting the presence or absence of a defect, then the value of each  $P$  pixel of this matrix can be defined as Equation 1:

$$\mathcal{A}_I(P) = \begin{cases} 0 & \text{if pixel } P \text{ is not an anomaly in image } I \\ 1 & \text{if pixel } P \text{ is an anomaly in image } I \end{cases} \quad (1)$$

with image  $I$  and the  $\mathcal{A}_I$  binary map, the problem of detecting dipping and buckling in the sandwich panel can be defined as automatic detection of the  $\tilde{\mathcal{A}}_I$  binary map, provided that  $\tilde{\mathcal{A}}_I$  is an accurate and precise estimation of  $\mathcal{A}_I$ .

In most of the studies carried out to detect anomaly, the basic assumption is that the proposed method is trained using the normal flawless inputs, while during the test the system input can be both the normal and abnormal images, and the proposed model should be able to recognize and separate both of them. Since creating an input dataset of the anomaly cases is extremely difficult, costly, time-consuming, and sometimes impossible, this assumption exists in the proposed method of this study.

Hereafter, as a general agreement, the set of normal images used to train the proposed system will be denoted by  $E^{train}$ . Similarly,  $E^{val}$  denotes the set of normal validation images used, and finally, the set of images used in the test phase, which includes the normal and abnormal images, will be identified by  $E^{test}$ .

The proposed method for detecting dipping and buckling in a sandwich panel, a model,  $M(\theta)$ , learns the degree of normality of an image using the  $E^{train}$  set images, where the  $\theta$  variable denotes the free parameters of the model. Then using this model, which is fully trained on the normal images, each  $x$  input in the test phase is assigned a score  $\phi_I(x)$ , as its degree of anomaly. In this case, the higher the score  $\phi_I(x)$ , the more likely the input will be an anomaly. Therefore, the final input classification will be done with a  $T$  threshold, i.e. if the value  $\phi_I(x) > T$ , the input  $x$

will be labelled "abnormal"; otherwise,  $x$  denotes a normal sample. In such a situation, the equation  $\phi_I(x) = T$  will be the decision border of the problem in question.

In the proposed method, the  $T$  threshold value will be obtained using the  $E^{val}$  dataset that includes the inputs without anomalies.

Finally, as the last step of the proposed method, after detecting and locating the anomaly regions, the rate of dipping (depth of dipping) or buckling (height of buckling) will be calculated.

### 3.2. Proposed method based on image processing

In this work, the proposed method for detecting the anomaly of foam sheet surfaces is a region-based method, i.e. the input image (image  $I$ ) is divided into  $R$  sub-regions or better to say sub-windows with  $w_r \times h_r$  dimensions. These  $R$  sub-windows are obtained using the grid sampling method uniformly with step lengths. Then the proposed algorithm first calculates the anomalies of each of these sub-windows, and by combining the model response for each sub-window, it will obtain a map ( $\phi_I$ ) of the anomaly rate of each pixel of the original image sub-windows. By applying a threshold to this map, the  $\tilde{\mathcal{A}}_I$  map will be obtained, in which an estimate of the binary map is of  $\mathcal{A}_I$ .

In the proposed method, the anomaly rate of each sub-window is obtained by calculating the similarity between each sub-window and the reference dictionary  $D$ , which is made of normal areas. This dictionary is created using the images in the  $E^{train}$  collection. This external similarity is obtained by calculating the mean Euclidean distance between the feature vectors of each sub-window and the feature vectors extracted from the most similar sub-area of the dictionary. Since each sub-area is selected with step length  $s$  and the value of  $s$  is smaller than the sub-window ( $s < w_r, s < h_r$ ), each pixel in image  $I$  can belong to several sub-regions. Therefore, the anomaly of each pixel similar to Figure 1 will be obtained by averaging the anomaly of each sub-area. Finally, the final binary map ( $\tilde{\mathcal{A}}_I$ ) is obtained by thresholding the value of  $T$  on the map  $\phi_I$ .

Figure 2 shows a schematic representation of the  $\phi_I$  map obtained with windows with  $w_r = h_r = 3$  and step length  $s = 2$  and the corresponding map  $\tilde{\mathcal{A}}_I$ .

### 3.2.1. Feature extraction

In order to estimate the similarity between two images (in this work, sub-windows extracted from video frames) in the image processing domain, their features were compared. In the field of image processing, several methods have been proposed in order to extract the visual feature of the image. These features can be divided into the two categories of handcrafted features (or pre-designed features) and trained features (or deep features). A human expert extracts the handcrafted features by a pre-designed algorithm such as SIFT and SURF [33-35].

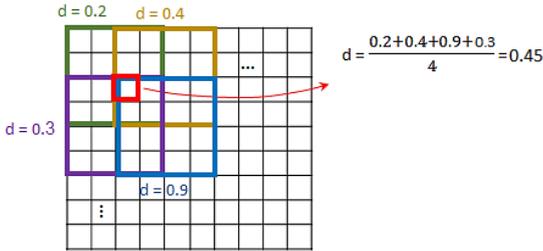


Figure 1. Calculating anomaly rate of each pixel based on estimated anomaly rate for each sub-window. In this example, size of sub-regions window is  $w_r = h_r = 4$ , and the step length is  $s = 2$ . The  $d$  values also represent the average apparent similarity value between each sub-window and the most similar sub-region in  $D$  dictionary per sub-window.

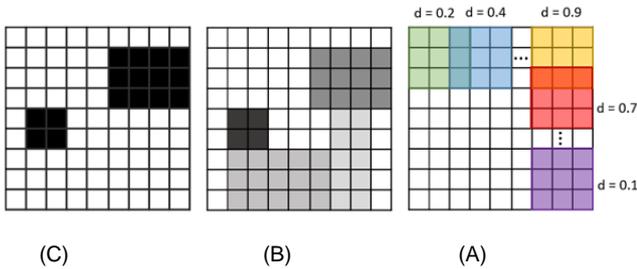


Figure 2. Schematic illustration of a sample of map  $\phi_1$  and its corresponding binary anomaly map ( $\tilde{A}_1$ ) A: An example of a sub-windows extraction procedure with  $w_r = h_r = 3$  and step length value  $s = 2$ . The  $d$  values also represent mean apparent similarity between each sub-window with the most similar sub-regions within  $D$  dictionary: Map  $\phi_1$ . In this example, the greater the numerical value of the anomaly (the distance), the darker the color of the area is shown. C: The corresponding binary map ( $\tilde{A}_1$ ).

Despite the relatively acceptable performance of the pre-designed feature-based approaches, due to the many difficulties in these approaches, the researches are moving towards using deep models. The trained or deep features are those extracted from the image by a trained convolutional neural network. In this method, there are no predictable ways for extracting the features from an image, and they are extracted using the image information and with a trained prior knowledge using other similar images [36,

37]. In the recent years, the convolutional neural networks have performed remarkably well in various applications such as voice recognition and image classification [38-43, 53-54].

The important point is that the neural network trained for a particular application can also be used for other applications. Because the network weights are trained in order to extract the key features of the image, and if the images do not differ fundamentally, the network can also be used for other applications. This is very effective in using these models in various applications because in some cases, network training is not possible from the beginning due to insufficient available training data. Therefore, using the pre-trained neural networks on very large datasets (such as Image Net containing over 1.2 million images in 1000 different classes [44]) and then using it as initialization of weights to fine-tune the model or using the main values of these models to extract the features from the images for the intended use is a normal and reasonable procedure in solving the problems using the neural network models [38].

In this work, the features based on the CNN networks were used by applying the AlexNet and ResNet-18 architectures. The reason for using these two architectures is that the scene identification and object recognition from the features in the images is a very complex issue. Thus the networks trained on these images are capable of detecting any particular pattern in the image (including very simple patterns within the surface of the sandwich panels). Like many existing methods that use this network for feature extraction [45, 46], the proposed method also uses FC7 hidden layer output for feature extraction. The dimensions of this feature vector is a 4096 dimensional vector. In the ResNet-18 network, the Average Pooling layer features are also used as the extracted features. This feature vector is a 512-dimensional one. The results of experiments on the existing data show that using the AlexNet architecture, the anomalies of the sandwich panels can accurately be identified, so this model provides better features for the work intended in this work and better matching with the input data in the problem ahead.

### 3.2.2. Reference dictionary construction

As mentioned earlier, the degree of anomalies of a sub-area is obtained by calculating the similarity between that sub-area and a reference dictionary  $D$  made of the normal sub-regions. The reference dictionary is constructed from a set of training images as  $E^{train} = \{E_1, \dots, E_L\}$ . For each

image  $E_l$ , the number of R windows  $\{P_1, \dots, P_R\}$  in  $w_n \times h_n$  dimensions is extracted by applying a uniform sampling strategy with step length  $s$ . Therefore, the total number of windows extracted for the whole  $E^{train}$  training set will be  $L_T = L \times R$ .

In the proposed method framework (convolutional neural network), the feature extraction module for each  $P_r$  window provides a vector of N dimensions, and this is the extracted properties vector  $f_r = \{f_1^r, \dots, f_N^r\}$  from the intended.

Then the feature vector dimensions obtained for all images in  $E^{train}$  are reduced to M ( $M < N$ ) using the Principal Component Analysis (PCA) [47]. M is the number of principal components within a data that represents a certain percentage of the variance in the data.

After reducing the dimension, each feature vector is normalized using Equation 2.

$$f_i^r = \frac{f_i^r - \mu_i}{\sigma_i}, \forall i \in (1, M) \quad (2)$$

where  $\mu_i$  and  $\sigma_i$  are the mean and variance of the data, respectively, and are defined as Equations 3 and 4, respectively.

$$\mu_i = \frac{1}{L_r} \sum_{r=1}^{L_r} f_i^r \quad (3)$$

$$\sigma_i = \frac{1}{L_r} \sum_{r=1}^{L_r} (f_i^r - \mu_i)^2 \quad (4)$$

The normalization process makes all the feature components have a zero mean and the same variance.

Finally, the reference dictionary is obtained by clustering all the reduced feature vectors within the training set into  $K_D$  clusters. In this work, we used the k-means clustering algorithm [48]. The set of all these  $K_D$  feature vectors form the Reference Dictionary (D).

Figure 3 illustrates the reference dictionary construction process schematically. As shown in this figure, windows are first extracted in the images with a specified step length. The feature vectors related to these windows are extracted using the CNN network, and their dimensions are reduced by the PCA method. After normalizing these vectors, the clustering operations are performed, and the cluster centres are considered as the key dictionary samples.

### 3.2.3. Anomaly detection

As mentioned earlier, the logic of the proposed method is that in order to detect the anomalies in an image, it must be estimated how far a sub-area of the original image is far from normality. For

this purpose, the concept of "normality" should be trained from one or more images without anomalies.

Therefore, we used V images from the  $E^{val}$  validation set of images that were not used to construct the reference dictionary. As a result, each image from the  $E^{val}$  set is processed similar to the image processing procedure used to construct a reference dictionary.

In this case,  $V_r = V \times R$  windows are extracted from the V image. Then the feature vectors with length N are extracted from each window and reduced to M dimensions using PCA. Eventually, these vectors are normalized. At the end of this procedure, the mean Euclidean distance between all the validation set feature vectors and the most similar sub-area of the reference dictionary m is calculated,  $\mathbf{d} = \{d_1, \dots, d_{V_r}\}$ .

Finally, the concept of "normality" is modelled using the borders of a Gaussian function with mean and variance, as Equations 5 and 6.

$$\mu_{anomaly} = \frac{1}{V_r} \sum_{r=1}^{V_r} d_r \quad (5)$$

$$\sigma_{anomaly} = \frac{1}{V_r} \sum_{r=1}^{V_r} (d_r - \mu_d)^2, \quad (6)$$

Using these borders, we can define a threshold for detecting a sub-region as normal or abnormal (Equation 7):

$$T_{nor} = \mu_{anomaly} + \eta \sigma_{anomaly}, \quad (7)$$

where  $\eta$  is a real positive constant number that modifies the problem settings. The lower this value, the more strongly the system depends on its previously trained data. At the test time, a sub-region of the image will also be considered anomaly if the mean Euclidean distance  $d_{test}$  with the most similar sub-regions of the reference dictionary m is more than the  $T_{nor}$  threshold.

### 3.2.4. Pre-processing

In order to maximize the optimization and reduce the processing load of the proposed system, the frame images are pre-processed before performing the proposed steps. The pre-processing procedure is very simple, and at the same time reduces the processing over load of the proposed system and minimizes the time required for real-time panels processing.

In this pre-processing procedure, a local color mean filter is proposed that can detect relatively probable anomalous areas quickly but with a low accuracy at the very beginning of the process. The procedure for this pre-processing is also to use a local color mean equation, i.e. the input image is

divided into P sub-windows (W) with a specific size, and then the color mean within each area is calculated according to Equation 8.

$$AvgR_w = Mean(W(R)) \quad (8)$$

$$AvgB_w = Mean(W(B))$$

$$AvgG_w = Mean(W(G))$$

where W (R) is the value of the red channel in the sub-window W.

Then for each pixel  $P_i$ , the distance of that pixel is calculated by the values of the color mean sub-window in which they are located (Equation 9), and the mean distance of that pixel ( $D_{P_i}$ ) is obtained by Equation 10.

$$d_r = abs(P_i(R) - AvgR_w), \quad (9)$$

$$d_b = abs(P_i(B) - AvgB_w),$$

$$d_g = abs(P_i(G) - AvgG_w).$$

$$D_{pi} = \frac{d_r + d_b + d_g}{3} \quad (10)$$

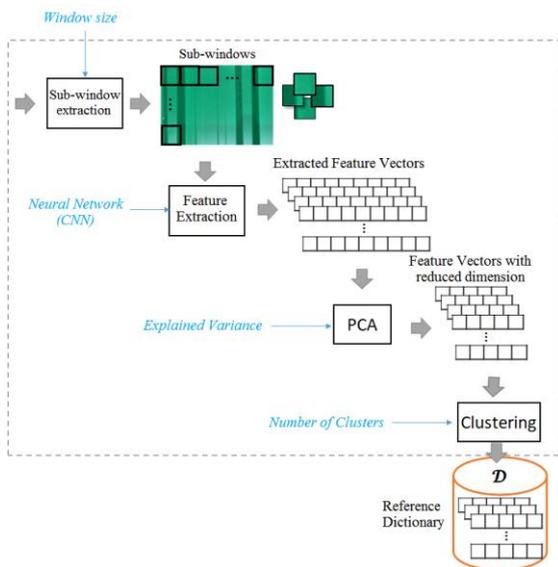


Figure 3. General process of reference dictionary construction.

Finally, if the value of this distance per pixel is greater than a  $T_{Preprocess}$  threshold value, that pixel is probably selected as the case of anomaly. By doing this on every pixel in the image, a binary mask image having the same dimension as the input image is obtained that determines whether the pixel is likely to be abnormal. Finally, a dilation morphological operation is applied to the mask.

Since this step is merely a pre-processing and this method is not very accurate for detection, a very non-rigid threshold was used to remove only the completely specific normal areas from future processing, thereby reducing the processes

described in the previous sections. To this end, in the next steps the extracted sub-windows ( $\{P_1, \dots, P_R\}$ ) instead of the whole image will only be considered around these probable estimated areas in the pre-processing so that all these regions are included (Figure 4).

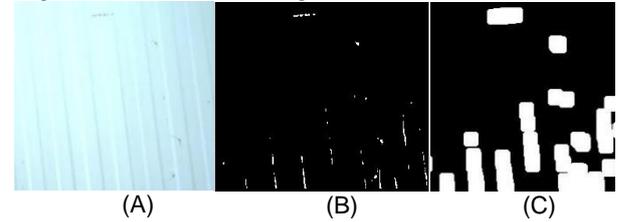


Figure 4. Pre-processing applied to a sample image. A: Frame original image. B: Output after applying color mean filter and thresholding C: Output after applying dilation morphology operations.

### 3.3. Measuring dipping and buckling

In the previous sections, a method has been developed that is capable of detecting the defects and anomalies in the sandwich panels. In this section, a method will be presented in order to estimate the extent of dipping and buckling (depth or height of distortions). In the following, a method based on color images and laser is presented in order to detect the depth of dipping. This method will use depth data information for detection. We have already presented and discussed this method in our previous paper [52], and the interested reader is encouraged to study that paper regarding more specific details. Below, we provide a concise description of the method.

The method presented in the previous chapters is image processing-based and very fast that does not require much training data. In this method, the model can also learn the surface of any object with any texture, pattern or even color, and is able to detect the defects in any different and complex texture. However, one of the problems of this method is that the anomaly must be completely distinct from the background in order to include a different pattern for detection by the image processing system presented. For example, a white buckling on a smooth white sandwich panel surface cannot be clearly distinguished unless it produces a different edge or pattern in the resulting image. Therefore, in order to solve this problem, the proposed laser-based method framework will be designed in such a way that in addition to estimating distortions, it will be able to detect the location of the anomalies. The experiments in the results section show that the shortcomings of this method for detecting anomalies also result in a combination of the two laser-based and image processing methods to present an optimum output. In the results section,

a combination of the two proposed methods will be discussed as the final system presented.

The goal of this project consists of scanning a panel's surface and extract 3D information to detect abnormal dipping and buckling. Utilizing color and depth cameras (RGBD) like Microsoft Kinect and mobile laser scanners are the first options that come to mind to solve this problem. Since the sandwich panel is moving on the production line during monitoring, the proposed system uses a simpler and more economical technology that only examines a single line of panel at a time. The proposed system uses a simple color camera module (8M high resolution camera module-CM8M30M12C) and a linear laser (HLM1230- 5mW 655 nm Focusable Laser Module). The linear laser is orthogonal to the sandwich panel's surface, and the laser beam image located in the camera's scope is processed to check flexion and curvature. Finally, using a calibration method, the laser beam curvature is mapped to the size of dipping and buckling in mm units to give the production line controller a better framework of the result of sandwich panel surface detection.

The four major components of our proposed method are as follow:

- (A) Calibration of the camera to resolve the distortion of the taken images
- (B) Laser linear beam detection in the image
- (C) Detection of flexion and curvature of the laser beam on the surface of the sandwich panel
- (D) Mapping the flexion and curvature of the laser beam to the extent of the dipping and buckling in mm units.

In the following, each component will be briefly described.

### 3.3. Camera calibration

The images captured by a majority of the existing cameras contain abnormal curvatures and distortions due to the limitations of the lenses used. We first need to resolve this problem by camera calibration (i.e. determining camera's internal parameters); otherwise, the measurement of laser beam curvature would be highly inaccurate.

To this aim, we take a set of images of a standard object with a checker-board pattern from different angles, and then using the calibration methods, the internal parameters of the camera are determined [49, 50]. Having these parameters, which consist of two parameters ( $k_1$  and  $k_2$ ) of the camera's distortional model, the distortions captured by the camera are corrected using Equations 11 and 12.

$$x = x_d(1 + k_1r^2 + k_2r^4) \quad (11)$$

$$y = y_d(1 + k_1r^2 + k_2r^4) \quad (12)$$

Here,  $(x, y)$  is the location of the original image modified pixel  $(x_d, y_d)$ , and  $r$  is equal to:

$$r^2 = x_d^2 + y_d^2 \quad (13)$$

The interested reader is referred to [49, 50] in order to find out more about the details of the calibration method we use.

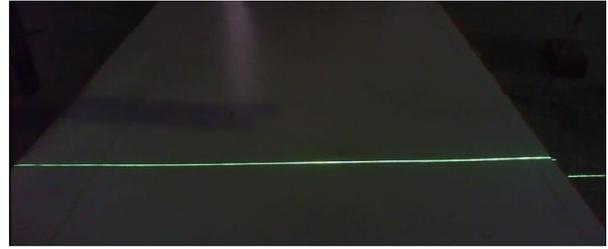


Figure 5. Sample image taken by camera, with low exposure time.

### 3.4. Laser beam detection in image

Following calibration, the laser beam should be detected in the camera's captured images. Due to the brightness of the laser beam, the light-sensitive sensor of the camera would be saturated in its normal condition. By adjusting the camera's optical sensor exposure time to light, the image is dimmed to the point where the sensor is out of saturation.

By increasing the exposure value (following Equation 14), which results in decreasing its time, the extra light and noise in the environment decreased sharply, while the laser beam became brighter than the image (Figure 5).

$$EV = \log_2 \frac{N}{t^2} \quad (14)$$

Finding the laser beam location that does not require the processing of the entire image, and only part of the image that includes the laser beam, can be processed. In order to locate the laser beam in the image, given the high intensity of the laser light, the easiest way is to apply a threshold to the image and to determine the approximate position of the laser beam. In order to eliminate noise and thinning, this operation along with the application of morphological operators can provide a good approximation of the laser beam location of the image.

Since we require a method to accurately detect the slightest curvature and change the angles in a laser beam image to locate dipping and buckling on the sandwich panel's surface, the method used in the proposed system, as well as the simple solution presented above, attempts to locate the laser

beams in the image with below the pixel accuracy. Given that the laser beam thickness is substantially larger than one pixel and in segments orthogonal to the laser beam, generally has a local maximum, this method tries to reach a continuous and sub-pixel accuracy for locating the laser beam by analyzing the laser pixels' intensity in the neighborhood.

Therefore, for each pixel on the laser beam image, a Gaussian function, i.e.  $y(x)$ , is fitted on the pixels in its neighborhood, i.e.  $f(x)$ , which are orthogonal to the laser beam at the selected pixel. The Gaussian function center (that can be in decimal and sub-pixel scale due to the continuity of the function) is considered as the exact position of the laser in that neighborhood (Figure 5). In optimal fitting of the Gaussian function, given the possibility of being trapped in local maxima due to the use of a decreasing gradient, choosing a suitable initial guess for the Gaussian function parameters can have a significant impact on the accuracy and speed of fitting. In the proposed system, the initial assumption for the Gaussian center in each segment would be the pixel identified in the threshold operation.

Amplitude of Gaussian ( $A$ ) is the light intensity of that pixel, and the Gaussian sigma ( $\sigma$ ) is dependent on the number of pixels associated with the laser beam in the investigated segment.

$$y(x) = Ae^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (15)$$

In the presented method, we use the exponential function of the difference error, with the aim of minimizing the effect of high values of the difference on finding the parameter optimal values:

$$E(g(e)) = -\frac{1}{N} \sum_{i=1}^N e^{-\left(f(x) - Ae^{-\frac{(x-\mu)^2}{2\sigma^2}}\right)^2} \quad (16)$$

By deriving the error function for unknown parameters and by applying the gradient descent, the initial guess can be propelled to the desired result (Equation 17).

$$\begin{pmatrix} \sigma_{new} \\ \mu_{new} \\ A_{new} \end{pmatrix} = \begin{pmatrix} \sigma_{old} \\ \mu_{old} \\ A_{old} \end{pmatrix} - \gamma \begin{pmatrix} \frac{\partial E(g(e))}{\partial \sigma} \\ \frac{\partial E(g(e))}{\partial \mu} \\ \frac{\partial E(g(e))}{\partial A} \end{pmatrix} \quad (17)$$

$$\text{if } \text{sgn}(\gamma_{old}) \cong \text{sgn}(\gamma_{new}) : \gamma_{new} = 0.8\gamma_{old} + 0.001$$

Selecting a constant value for the speed coefficient of the gradient descent makes fitting time-consuming.

This problem especially occurs when the data has a long tail, which relatively increases the overall value of the error function, and the rest of the data seems to have a low value. This issue makes the progress towards a solution to be relatively slow.

In order to solve this problem, if we increase the constant speed of the gradient descent, it will initially move faster towards the target but will skip over it several times when it reaches the local maximum (the over-shoot phenomenon).

An appropriate coefficient of variation must be such that is initially high and then approaches zero when reaching the solution. For this purpose, using the interpolation method to find the root and the constant coefficient, if the partial derivative is changed, we multiply its coefficient by less than one to decrease the speed coefficient of the gradient descent in each divergence.

Finally, the  $\mu$  value represents the location of the laser beam in the segment under investigation.

### 3.4.1. Laser beam curvature detection

By analyzing the laser beam that is accurately located in the image, the location of the dipping and the buckling would be determined accurately. For this purpose, it is sufficient to first extract a number of short lines (10-20 pixels) from the laser beam. In the second step, each part of the extracted line is considered as independent and the slope of the straight line that best describes that part is calculated. It is expected that on the intact surface of the sandwich panel (that occupies a very significant part along each laser beam), the slope of the extracted lines does not change unusually.

In the third and final steps, by comparing the slope of the obtained parts with the mean, the locations of the panel that are defective can be accurately determined. The segments whose slope is more than one threshold ( $T_{L1}$ ) from the mean slope value are selected as the candidate areas for the anomaly. The entire surface of each panel is scanned according to the method described with high speed and accuracy to determine dipping and buckling.

### 3.4.2. Measuring amount of dipping and buckling

The described system could determine the relative degree of error in the curvature of the laser beam image but does not have a standard unit (i.e. ISI unit). In order to solve this limitation, we calibrate the system's measurement unit with a standard

panel so that the panel covers the width of the laser beam in image, and includes a checker-board pattern with precisely determined dimensions (in millimetres) on its vertical wall. The checker-board pattern points, as shown in Figure 6, are found that indicate what height are the camera pixels, located on laser beam area, in mm. Using this dataset, a leading artificial neural network is trained that has two neurons in the input layer (corresponding to coordinate of pixel) and a neuron in the output layer (corresponding to height in mm). Given the representational power of artificial neural networks with a hidden layer [51], 4 neurons would be considered for the hidden layer in this network. At runtime, after finding points in the image where the laser beam has an abnormal curvature, the input of the neural network will be fed with coordinates of those points. The height of the laser beam would be accurately determined by this network at the given point. The human supervisors and quality control experts would be able to use this feature to easily evaluate the product quality and take any optional and necessary actions based on the output of the system (for example, if the distortion exceeds a  $T_{L2}$  threshold, the production line will be stopped). To the best of our knowledge, there are no other studies or works on measuring the size of dipping or buckling on the surface of sandwich panels. Hence, this removes the need for performing a comparative study, and also makes our approach the first of its kind.

## 4. Experiments

### 4.1. Implementation details

Using a validation dataset and a grid search strategy, the proposed system in this work was tested by varying the following parameters to select the best value of the parameters:

As mentioned earlier, by examining the outputs of AlexNet and ResNet-18 neural networks, it was concluded that the features presented by the AlexNet network in the current application and the data used in this work performed better for extracting anomaly regions from the sandwich panels' images. Since the sub-windows extracted from the sandwich panel images are very simple images with plain patterns, there is no need for a complex procedure to extract features for this data. However, the complexity of the ResNet-18 network over the AlexNet network may be the reason of poor performance of the features extracted from the network. Therefore, in this work, we used the features extracted by the AlexNet network as the feature vectors for reference dictionary construction as well as in the

test phase. The input of this network is a  $227 \times 227$  image, so the input sub-windows are initially resized.

The extracted sub-windows were evaluated in  $10 \times 10$ ,  $25 \times 25$ ,  $50 \times 50$ , and  $100 \times 100$  dimensions. The remarkable point is that the larger the extracted sub-windows, the less complex the time and computational load of the system, and the less accurate the detection accuracy. Since large windows contain much of the normal data, the similarity between the normal and abnormal windows is increased, and as a result, finding the decision boundary is more difficult. The most desirable result in the current system is the  $25 \times 25$  sub-window.



Figure 6. Determination of checker-board pattern points, used as training data for neural network learning.

Reference dictionary size (number of clusters in the clustering algorithm): Different values of 5, 25, 50, 100, and 150 were evaluated for the number of categories in the reference dictionary. The greater the number, the longer the time required to estimate the similarity between a test data sub-window with the reference sub-windows and the better the performance and efficiency of the system. The most optimal result in the current system was with the dictionary size  $K_D = 50$ .

Tables 1 and 2 show the selected values for the proposed system parameters, which have had the best performance on a held-out validation set.

### 4.2. Results of proposed system

Since the sandwich panel is produced in discrete segments (such as 10 m or 20 m), a component is required to identify the beginning and end of each segment in order to continuously operate the proposed system on the production line. This component provides the system supervisor the result of processing each sandwich panel separately and independently. For this purpose, the thickness seen by the camera is averaged at each width. Given that, this value passes a certain

threshold when the panel enters and exits under the machine, the beginning and the end of each piece of sandwich panel is easily determined.

In order to evaluate the desired data system, videos of 35 different sandwich panels with lengths of 10 and 20 m and thicknesses of 7 and 20 cm were collected in various patterns and colors. In order to train the image processing system, videos containing normal panels (with quality supervisor approval) were extracted for each one of the 35 panels as the reference samples. The total duration time of these videos is 5 hours. Then a number of different dipping and buckling were manually added to the surface of each sandwich panel to test the system; and a one-hour video of these 35 different panels with different anomalies was used. The anomalies in the test videos are labelled for system accuracy and evaluation by a quality supervisor.

**Table 1. Different values of parameters in proposed system.**

Parameter	$K_D$	PoV (proportion of variance)	$w_r \times h_r$	s	$T_{preprocess}$
Value	50	97%	25x25	15	150

**Table 2. Different values of thresholds considered in second part of proposed system to detect depth of dipping and buckling.**

Threshold	$T_{L2}$	$T_{L1}$
Value	3mm	5

#### 4.2.1. Evaluation criteria

In order to calculate the accuracy of the system detection after obtaining the binary map estimation  $\hat{\mathcal{A}}_I$ , the Non-Maxima Suppression (NMS) was applied on the output. As a result, a number of points were selected as the more likely anomaly regions. Then the distance from these points to the nearest anomaly point in the binary map of  $\mathcal{A}_I$  Ground Truth (GT) was calculated. If this distance was less than a  $T_g$  threshold, that point was considered to be a “correctly labelled” anomaly. The  $T_g$  threshold value in this work was 10 pixels.

Finally, Equation 18 was used to calculate the precision identification (F1 Score).

$$F1Score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (18)$$

The precision and recall values are defined as Equations 19 and 20:

$$Precision = \frac{TP}{TP + FP} \quad (19)$$

$$Recall = \frac{TP}{TP + FN} \quad (20)$$

In the above equations, the TP variable represents the number of correctly identified anomalies. The TN variable represents the number of normal samples that are correctly classified in the normal class. Also the FP variable represents the samples that were normal and were wrongly detected as abnormal, and FN shows the number of anomaly samples that were wrongly identified as normal.

**Table 3. Identification precision (F1 Score) obtained for size of different sub-windows and reference dictionaries in the proposed system.**

Size of reference dictionary	Size of extracted sub-windows			
	10	25	50	100
5	80.6	80.1	78.2	70.6
25	88.4	87.8	83.4	79.2
50	92.3	<b>92.2</b>	90.6	86.9
100	92.7	92.4	90.8	87.1
150	92.9	92.7	91.1	87.9

#### 4.2.2. Results

Table 3 shows the identification precision (F1 Score) obtained for the proposed system of different values for the size of the selected windows and the reference dictionaries. As it could be seen in the results, for the  $10 \times 10$  window and the dictionary size of 150, the proposed system performed the best and achieved a 92.9 precision. As mentioned earlier, the smaller the size of the window, the greater the number of selected windows ( $w_r = h_r$ ) and the larger the reference dictionary ( $K_D$ ), the greater the number of samples to be compared with each window. Therefore, these two factors increase the performance time. Due to the precision of the window size of 25 and dictionary size of 50, and reduction of its performance time to at least one-third of the runtime required for the values  $w_r = h_r = 10$  and  $K_D = 150$ , these values were selected as the appropriate parameters in the proposed system.

Finally, we considered the proposed laser-based method in order to estimate the distortion. Due to the lack of a proper standard dataset for detecting the extent of dipping and buckling in the sandwich panels, or the proprietary nature of the other methods, it is not possible to reproduce or compare the results with other methods. However, in order to illustrate the efficiency of the proposed system, we compared it with a basic one. The basic method used to detect distortions is very similar to the proposed one, except that it detects the anomalies on the surfaces by fitting a line, not on the laser line segments, but on the entire laser line. The basic method considers those

points to be the anomalies whose distance to the matched straight line is greater than a threshold. The results of the experiments showed that the proposed laser-based system was able to achieve an accuracy of 89.7% for detecting anomaly, whereas the accuracy of the basic method was 76%. Moreover, the overall error of this system in estimating the rate of dipping and buckling was 0.9 mm.

As a final issue, the experiments were conducted to combine the two proposed methods (image processing-based and laser-based). As mentioned earlier, image processing is a very fast method that does not require much data for training, and has no dependence on the color or pattern of the sandwich panels. However, this method may have some difficulty to detect in some situations where the dipping and the buckling on a flat surface does not cause any discoloration or visible edges in the image. In contrast, the laser-based method does not have such problems. This method is somewhat sensitive to the pattern on the sandwich panels (due to the comparison of the slope of the obtained fragments with the mean), and may sometimes fail to properly detect the existing anomalies. Therefore, considering this method as a compliment, some experiments have been considered to combine the results of this method, i.e. the values of the anomaly per pixel in the two methods are merged together. In the image processing-based method, anomaly score was shown using the  $\phi$  map, which indicated its score value as an anomaly sample for each pixel in the image. Also in the laser-based method, the slope distance of the obtained segments per pixel with the mean slope value was considered as the degree of anomaly score of each pixel (map  $\Psi_I$ ). Eventually, the final score for each pixel  $p$  was obtained by combining the two methods, similar to the Equation 21.

$$AnomalyScore_p = \alpha\Phi_I(p) + (1 - \alpha)\Psi_I(p) \quad (21)$$

where  $\alpha$  is the weight coefficient of these two methods integrations.  $\alpha$  value is considered 0.7 by testing on a validation set.

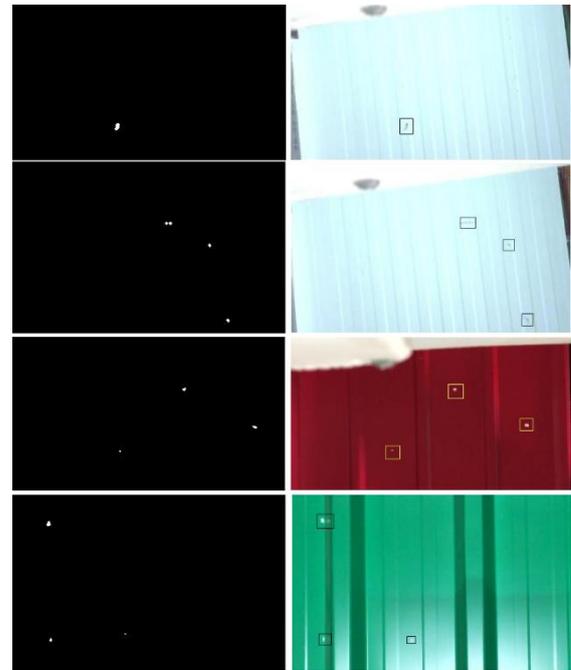
Table 4 indicates the results obtained using each one of the two methods and their combination. As the results obtained show, a combination of the two methods has been able to increase the detection accuracy of dipping and buckling of the sandwich panels surfaces.

Figures 7 and 8 illustrate the system output sample for an experimental panel sandwich (manually created on its surface some dipping and

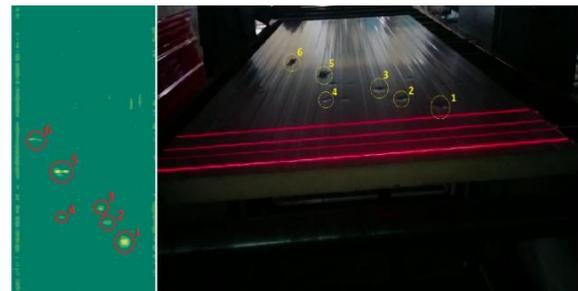
buckling). As it can be seen, the location of almost all distortions has been correctly identified.

**Table 4. Results of image-based and laser-based methods and their combination.**

	Image processing-based method	Laser-based method	Compared (i.e. basic) method	Combining our two proposed methods
Identification precision (F1 Score)	92.2	89.7	76	95.6



**Figure 7. An example of a system output. Right column: framed images with areas identified as anomaly. Left Column: binary mask ( $\mathcal{A}_I$ ) for anomalous areas in each frame.**



**Figure 8. An example of designed system performance. Right image is taken with the camera from the panel. Some points of the dipping and buckling are numbered on it. Left image is obtained by laser scanning the panel and its analysis. This image denotes the points corresponding to the panels dipping and buckling that have been correctly detected.**

## 5. Conclusions

In this work, a system was designed and implemented for the quality control of surface of the sandwich panel for detection of abnormal dipping and buckling. We showed that with very inexpensive hardware tools, and employing image

processing and machine vision techniques, very small dipping and buckling (even less than 1 mm in depth or height) could be accurately identified. Since the cost of sandwich panel production is relatively high, the designed system helps to detect the possible anomalies in the sandwich panel production line, resulting in an unusual dipping and buckling on the surface of the panel as quickly as possible and without the need for manpower. Then the necessary steps can be taken to resolve it. Two main solutions were proposed for this purpose. In the first step, an image processing system was developed that was able to detect the anomalies on the sandwich panels. In the second step, a method was presented that was capable of detecting and measuring the amount of distortion, i.e. the size of any dipping and buckling. For this purpose, a laser-based system was developed without changing the configuration of the production line and without the use of complicated hardware. However, it must be noted that the laser-based solution is not able to detect all the defects, and it might produce many false-positives when the texture of the sandwich panel is not flat and has large structural patterns. Finally, we showed that combining both the image processing-based and laser-based methods could greatly improve the performance of the system in anomaly detection; however, this comes at the cost of increasing the computational complexity and latency of the proposed methods, which should be taken into account due the fact that this system should perform in real-time.

## References

- [1] A. Kumar, "Computer-vision-based fabric defect detection: A survey," *IEEE transactions on industrial electronics*, Vol. 55, pp. 348-363, 2008.
- [2] A. Kumar and G. K. Pang, "Defect detection in textured materials using Gabor filters," *IEEE Transactions on industry applications*, Vol. 38, pp. 425-440, 2002.
- [3] D. A. Wheeler, B. Bryczynski, and R. N. Meeson Jr, *Software Inspection: An Industry Best Practice for Defect Detection and Removal*: IEEE Computer Society Press, 1996.
- [4] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, Vol. 41, p. 15, 2009.
- [5] P. Navarro, C. Fernández-Isla, P. Alcover, and J. Suardíaz, "Defect detection in textures through the use of entropy as a means for automatically selecting the wavelet decomposition level," *Sensors*, Vol. 16, p. 1178, 2016.
- [6] D. Micucci, M. Mobilio, P. Napoletano, and F. Tisato, "Falls as anomalies? An experimental evaluation using smartphone accelerometer data," *Journal of Ambient Intelligence and Humanized Computing*, Vol. 8, pp. 87-99, 2017.
- [7] M. W. Berry and M. Castellanos, "Survey of text mining," *Computing Reviews*, Vol. 45, p. 548, 2004.
- [8] S. Boriah, V. Chandola, and V. Kumar, "Similarity measures for categorical data: A comparative evaluation," in *Proceedings of the 2008 SIAM international conference on data mining*, 2008, pp. 243-254.
- [9] V. Chandola, S. Boriah, and V. Kumar, "Understanding categorical similarity measures for outlier detection," *Technology Report, University of Minnesota*, 2008.
- [10] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal Processing*, Vol. 99, pp. 215-249, 2014.
- [11] L. Weiwei, Y. Yunhui, L. Jun, Z. Yao, and S. Hongwei, "Automated on-line fast detection for surface defect of steel strip based on multivariate discriminant function," in *2008 Second International Symposium on Intelligent Information Technology Application*, 2008, pp. 493-497.
- [12] Z. Chao, Z. Desen, and X. Li, "Textural defect detection based on label co-occurrence matrix," *JOURNAL-HUAZHONG UNIVERSITY of SCIENCE and TECHNOLOGY NATURE SCIENCE EDITION*, Vol. 34, p. 25, 2006.
- [13] P. Caleb and M. Steuer, "Classification of surface defects on hot rolled steel using adaptive learning methods," in *KES'2000. Fourth International Conference on Knowledge-based Intelligent Engineering Systems and Allied Technologies. Proceedings (Cat. No. 00TH8516)*, 2000, pp. 103-108.
- [14] K.-L. Mak, P. Peng, and K. F. C. Yiu, "Fabric defect detection using morphological filters," *Image and Vision Computing*, Vol. 27, pp. 1585-1592, 2009.
- [15] A. Cord, F. Bach, and D. Jeulin, "Texture classification by statistical learning from morphological image processing: application to metallic surfaces," *Journal of Microscopy*, Vol. 239, pp. 159-166, 2010.
- [16] D.-H. Shi, T. Gang, S.-Y. Yang, and Y. Yuan, "Research on segmentation and distribution features of small defects in precision weldments with complex structure," *NDT and e International*, Vol. 40, pp. 397-404, 2007.
- [17] R. Anand and P. Kumar, "Flaw detection in radiographic weldment images using morphological watershed segmentation technique," *Ndt and E International*, Vol. 42, pp. 2-8, 2009.
- [18] S. Zhou, D. Liang, and Y. Wei, "Automatic detection of metal surface defects using multi-angle lighting multivariate image analysis," in *2016 IEEE International Conference on Information and Automation (ICIA)*, 2016, pp. 1588-1593.

- [19] C.-h. Chan and G. K. Pang, "Fabric defect detection by Fourier analysis," *IEEE transactions on Industry Applications*, Vol. 36, pp. 1267-1276, 2000.
- [20] B. Zuo and F. Wang, "Surface cutting defect detection of magnet using Fourier image reconstruction," *Computer Engineering and Applications*, Vol. 52, pp. 256-260, 2016.
- [21] J. L. Raheja, S. Kumar, and A. Chaudhary, "Fabric defect detection based on GLCM and Gabor filter: A comparison," *Optik*, Vol. 124, pp. 6469-6474, 2013.
- [22] D. Choi, Y. J. Jeon, J. P. Yun, S. W. Yun, and S. W. Kim, "An algorithm for detecting seam cracks in steel plates," *World Acad Sci Eng Technol*, Vol. 6, pp. 1456-1459, 2012.
- [23] Y.-J. Jeon, D.-c. Choi, J. P. Yun, C. Park, and S. W. Kim, "Detection of scratch defects on slab surface," in *2011 11th International Conference on Control, Automation and Systems*, 2011, pp. 1274-1278.
- [24] H. Zheng, B. Jiang, and H. Lu, "An adaptive neural-fuzzy inference system (ANFIS) for detection of bruises on Chinese bayberry (*Myrica rubra*) based on fractal dimension and RGB intensity color," *Journal of food engineering*, Vol. 104, pp. 663-667, 2011.
- [25] M. Yazdchi, M. Yazdi, and A. G. Mahyari, "Steel surface defect detection using texture segmentation based on multifractal dimension," in *2009 International Conference on Digital Image Processing*, 2009, pp. 346-350.
- [26] J. Blackledge and D. Dubovitskiy, "A surface inspection machine vision system that includes fractal texture analysis," 2008.
- [27] Z. Jiuliang, L. Weiwei, Y. Feng, L. Jun, Z. Yao, and Y. Yunhui, "Research on surface quality evaluation system of steel strip based on computer vision," in *2009 Third International Symposium on Intelligent Information Technology Application*, 2009, pp. 32-35.
- [28] H. Zheng, L. X. Kong, and S. Nahavandi, "Automatic inspection of metallic surface defects using genetic algorithms," *Journal of materials processing technology*, Vol. 125, pp. 427-433, 2002.
- [29] H. Jia, Y. L. Murphey, J. Shi, and T.-S. Chang, "An intelligent real-time vision system for surface defect detection," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, 2004, pp. 239-242.
- [30] C. Lee, C.-H. Choi, J. Choi, Y. Kim, and S. Choi, "Feature extraction algorithm based on adaptive wavelet packet for surface defect classification," in *Proceedings of 3rd IEEE International Conference on Image Processing*, 1996, pp. 673-676.
- [31] A. Kumar and H. C. Shen, "Texture inspection for defects using neural networks and support vector machines," in *Proceedings. International Conference on Image Processing*, 2002, pp. III-III.
- [32] H. Elbehiery, A. Hefnawy, and M. Elewa, "Surface defects detection for ceramic tiles using image processing and morphological techniques," 2005.
- [33] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, Vol. 60, pp. 91-110, 2004.
- [34] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*, 2006, pp. 404-417.
- [35] L. Ding and A. Goshtasby, "On the Canny edge detector," *Pattern Recognition*, Vol. 34, pp. 721-725, 2001.
- [36] C. Cusano, P. Napoletano, and R. Schettini, "Intensity and color descriptors for texture classification," in *Image Processing: Machine Vision Applications VI*, 2013, p. 866113.
- [37] P. Napoletano, "Hand-crafted vs learned descriptors for color texture classification," in *International Workshop on Computational Color Imaging*, 2017, pp. 259-271.
- [38] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: an astounding baseline for recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2014, pp. 806-813.
- [39] P. Napoletano, "Visual descriptors for content-based retrieval of remote-sensing images," *International journal of remote sensing*, Vol. 39, pp. 1343-1376, 2018.
- [40] S. Bianco, L. Celona, P. Napoletano, and R. Schettini, "On the use of deep learning for blind image quality assessment," *Signal, Image and Video Processing*, Vol. 12, pp. 355-362, 2018.
- [41] C. Cusano, P. Napoletano, and R. Schettini, "Combining multiple features for color texture classification," *Journal of Electronic Imaging*, Vol. 25, p. 061410, 2016.
- [42] A. Bhandare, M. Bhide, P. Gokhale, and R. Chandavarkar, "Applications of convolutional neural networks," *International Journal of Computer Science and Information Technologies*, Vol. 7, pp. 2206-2215, 2016.
- [43] O. Abdel-Hamid, A.-r. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional neural networks for speech recognition," *IEEE/ACM Transactions on audio, speech, and language processing*, Vol. 22, pp. 1533-1545, 2014.
- [44] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009, pp. 248-255.
- [45] Y. Zhou, H. Nejati, T.-T. Do, N.-M. Cheung, and L. Cheah, "Image-based vehicle analysis using deep neural network: A systematic study," in *2016 IEEE*

*International Conference on Digital Signal Processing (DSP)*, 2016, pp. 276-280.

[46] J. Yue-Hei Ng, M. Hausknecht, S. Vijayanarasimhan, O. Vinyals, R. Monga, and G. Toderici, "Beyond short snippets: Deep networks for video classification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 4694-4702.

[47] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, Vol. 2, pp. 37-52, 1987.

[48] E. W. Forgy, "Cluster analysis of multivariate data: efficiency versus interpretability of classifications," *biometrics*, Vol. 21, pp. 768-769, 1965.

[49] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 22, 2000.

[50] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *cvpr*, 1997, p. 1106.

[51] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, Vol. 4, pp. 251-257, 1991.

[52] Torkzadeh V, Toosizadeh S. Automatic visual inspection system for quality control of the sandwich panel and detecting the dipping and buckling of the surfaces. *Measurement and Control*. 2019 Sep; 52(7-8):804-13.

[53] Fallahzadeh, M., Farokhi, F., Harimi, A., Sabbaghi-Nadooshan, R. (2021). Facial Expression Recognition based on Image Gradient and Deep Convolutional Neural Network. *Journal of AI and Data Mining*, 9(2), 259-268. doi: 10.22044/jadm.2021.9898.2121

[54] Erfani, S. (2021). Automatic Facial Expression Recognition Method Using Deep Convolutional Neural Network. *Journal of AI and Data Mining*, 9(2), 153-159. doi: 10.22044/jadm.2020.8801.2018

## سیستم خودکار بازرسی بصری بر پایه پردازش تصویر و شبکه عصبی برای کنترل کیفی ساندویچ پانل

وحید ترک زاده و سعید طوسی زاده \*

گروه مهندسی کامپیوتر، واحد نیشابور، دانشگاه آزاد اسلامی، نیشابور، ایران.

ارسال ۲۰۲۱/۰۷/۲۴؛ بازنگری ۲۰۲۱/۱۲/۱۲؛ پذیرش ۲۰۲۲/۰۲/۲۱

### چکیده:

در این مقاله، یک سیستم خودکار بر پایه روش‌های پردازش تصویر، مبتنی بر ویژگی‌های استخراج شده از شبکه‌های عصبی پیچشی، با هدف شناسایی میزان فرورفتگی و برآمدگی روی سطوح ساندویچ پانل و استفاده از یک دوربین رنگی ساده ارائه شده است. روش پیشنهادی با دریافت یک تصویر از سطح ساندویچ پانل می‌تواند با دقت مناسب و قابل قبولی فرورفتگی‌ها و برآمدگی‌های روی سطح آن را شناسایی کند. پس از این که یک پانل به طور کامل توسط سیستم پردازش می‌شود، یک تصویر خروجی از وضعیت سطح ساندویچ پانل تولید می‌شود که در آن وضعیت خرابی‌های موجود روی سطح آن نمایش داده شده است و ناظر خط تولید می‌تواند با استفاده از این تصویر عیوب احتمالی را شناسایی و مکان یابی کند. همچنین یک راهکار با دقت بالا ارائه شده است که میزان خرابی‌های موجود روی سطح ساندویچ پانل را (عمق یا ارتفاع فرورفتگی و تورفتگی‌ها)، بدون نیاز به تجهیزات پیچیده یا سخت افزارهای گران قیمت، اندازه گیری می‌کند.

**کلمات کلیدی:** ساندویچ پانل، فرورفتگی، برآمدگی، پردازش تصویر، شبکه‌های عصبی پیچشی.