Leather Image Quality Classification and Defect Detection System using Mask Region-based Convolution Neural Network Model

Azween Bin Abdullah[1], Malathy Jawahar[2], Nalini Manogaran[3]*, Geetha Subbiah[4], Koteeswaran Seeranagan[5], Balamurugan Balusamy[6], Abhishek Chengam Saravanan[7]

Faculty of Applied Science and Technology, Perdana University, Kuala Lumpur, Malaysia[1]
Leather Process Technology Department, CSIR-Central Leather Research Institute, Chennai-600020, Tamil Nadu, India[2]
Department of CSE, S.A. Engineering College (Autonomous), Chennai-600077, Tamil Nadu, India[3]
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai-600127, Tamil Nadu, India[4, 7]
Department of CSE (AI and ML), S.A. Engineering College (Autonomous), Chennai-600077, Tamil Nadu, India[5]
Shiv Nadar (Institution of Eminence Deemed to be University), Uttar Pradesh – 201314, India[6]

***Abstract*—The leather industry is increasingly becoming one amongst the most important manufacturing industries in the world. Increasing demand has posed a great challenge as well as an opportunity for these industries. Quality of a leather product has been always the main factor in the setting of the market selling price. Usually, quality control is done with manual inspection. However, with human related errors such as fatigue, loss of concentration, etc., misclassification of the produced leather quality becomes a very serious issue. To tackle this issue, traditionally, image processing algorithms have been used, but, have not been effective due to low accuracies and high processing time. The introduction of Deep Learning methodologies such as Convolutional Neural Networks (CNNs), however, makes image classification much simpler. It incorporates automated feature learning and extraction, giving accurate results in lesser time. In addition, the usage of deep learning can also be applied for defect detection, which is, locating defects in the image. In this paper, a system for leather image classification and defect detection is proposed. Initially, the captured images are sent to a classification system, which classifies the image as good quality or defect quality. If the output of the classification system is defect quality, then a defect detection system works on the images, and locates the defects in the image. The classification system and the defect detection system are developed using Inception V3 CNN and Mask R-CNN respectively. Experimental results using these CNNs have shown great potential with respect to object classification and detection, which, with further development can give unparalleled performance for applications in these fields.**

Keywords*—Image leather classification; leather defect detection; Convolutional Neural Network; CNN; deep learning

I. Introduction

Image classification is a very important field in computer vision. Classification of images plays an important role in case of identification, segregation and decision-making processes in automated systems. With the ever-increasing complex tasks given to these computer vision-based tools and the increasing reliance on these systems, classifications systems must be robust and have high accuracy and detection rates, so as to perform error-free and without manual intervention. Classification systems are in high demand in the manufacturing sector, with one such industry being the leather industry [1].

Leather is a main material, used in the clothing and accessory industry for making fashion accessories, footwear, coats, amongst a host of other products. It is usually a residual product, made from the outer covering of animals such as skins and hides, commonly obtained as by-products of the meat industry. “Hide” is the outer covering of a large animal such as a buffalo, cow whereas “Skin” is that of a sheep, goat, etc. Leather, technically defined, is a natural protein polymer that is treated with tanning agents to make it resistant to enzymatic attack and putrefaction, as well as improve several physical properties.

Leather and leather products are increasingly gaining importance in the world [2]. The major task of leather processing industry is to convert the skin or hide into leather, from which, quality products can made. The leather, which is processed from raw to finish, is dried. The next step is the visual inspection and classification. Defects may arise in the processed sheet of leather due to various reasons such as wear and tear by the machines, heat, etc. These defects may vary in size and shape depending on the type of defect such as holes, scratches, bubbles, torn pieces, etc. The classification takes place according to the surface defect area of the defect, defect area, location, and sheet thickness amongst many other factors. The number of defects determines the price, with more defects fetching a lower price. Thus, classification is a very important process to be undertaken, to determine the economic value of the product.

The manual classification process, which is commonly used, is error prone. With the long working hours, operator fatigue and other reasons, these factors derail effective classification and quality assurance process. A simple classification error causes a loss of trust amongst customers and greatly devalues the product, implying significant losses for these leather producing factories. The defects, scratches, holes, bubbles, etc., thus need to be identified, located, and finally removed. Thus, defect detection is also essential.

The proposed leather defect detection system using advanced deep learning algorithms holds great potential to transform the leather industry. It can significantly improve efficiency, reduce costs, and enhance product quality through automated quality control. This system can quickly and accurately identify defects, streamlining production processes and saving time and resources. It also minimizes waste and rework, leading to cost reductions. Furthermore, the consistent adherence to quality standards and real-time defect detection ensures higher-quality leather products, fostering customer satisfaction and brand loyalty. These benefits are set to drive widespread adoption of advanced quality control technologies across the leather industry, enhancing the competitiveness of leather manufacturers globally.

The classification and defect detection system that is deployed, should be able to accurately classify the images and locate defects, if present in the images. With manual labour becoming increasingly expensive and error prone, the use of autonomous classification systems have been the most preferred, but least deployed at present, due to low accuracies and reliabilities. Thus, the system needs to be as autonomous and reliable as possible. Commonly used techniques for similar systems are image processing based techniques. The basic techniques include wavelet analysis, canny edge detection and thresholding. However, these techniques are pretty inaccurate due to their inapplicability on identification of different defects than on the one it was developed and tuned for. But, with the introduction of deep learning and most importantly, Convolutional Neural Networks (CNNs), image recognition tasks have been simplified. They also have excellent performance and high accuracy rates, which provide good reliability.

In this paper, we propose a system that combines classification and defect detection and developing a complete system by improving upon various algorithms. This paper's scope is to analyse existing methods of leather quality and defect classification and detection and propose a complete system for the classification and defect detection process using images [33-37].

The major contributions made in this paper are:

1) This paper proposes an efficient 2-stage based computer vision-based pipeline for automated intelligent leather defect detection.

2) The model uses Inception V3 in the first stage for identifying the leather images and classifying them as "Good" or "Defective", to effectively classify the defective leather.

3) The model deploys Region based Convolution Neural Network (R-CNN) for locating the defects, in the leather image.

4) The model operates in such a way that if only the image is classified as defective, the image is passed on to the second stage. This arrangement reduces the delay in processing and efficiently uses the resources.

5) The proposed model tests its performance on the real time data set acquired from Research labs. Since the number of defective images are practically less, augmentation procedure is employed and sufficient amount of defective leather images are obtained.

6) The performance of the proposed model is superior to the existing models and the results are promising.

The paper is structured in the following way. Firstly, a brief of image processing and CNNs is explored in Section II. Secondly, their applications in this field and related works that are carried out are explored in Section III. Thirdly, the methodology developed is described in Section IV. In Section V, the dataset, the experimental setup and the results are discussed. Section VI discusses the Conclusion and Future Scope of this research work.

II. Overview

A. Convolutional Neural Networks

Convolutional Neural Networks or CNNs are deep learning neural network algorithms widely used in computer vision applications. CNNs are now commonly used in image based applications where automatic feature extraction, object detection, image classification and image segmentation operations are required.

The Convolutional Neural Network structure is made up of three general layers. They are the input layer, hidden layers and an output layer. The hidden layers may contain multiple layers such as convolutional layers, activation or RELU layers, pooling layers, normalization layers and fully connected layers. As a result of the masking of the input and the output by the activation function and the final convolution in these layers, these are called as hidden layers.

A Convolutional Neural Network typically contains multiple convolutional layers, followed by pooling layers, normalization layers and a fully connected layer. The main function in the CNN is performed by the convolutional layers. These layers often consist of several kernels of multiple sizes which are used to apply the convolutional or dot product operation on their input, for producing feature maps. A pooling layer collects these feature maps from the previous layer and performs a max pooling or selection of a maximum activation from a small neighborhood feature region. Thus, the output feature dimensions are reduced. The next layer, the fully connected layer, is usually placed at the near end of the convolutional neural network and the output of this layer, is the high-level abstract classification or data for further processing, which is inference derived from the image input. Since the neural network is a self learning network, the convolutional layers and the fully connected layers contain neurons, whose weights are adjusted as the network goes through the training phase to set the optimal weights for accurate outputs [3][4].

Some of the best known CNNs in the field of computer vision are AlexNet [5], GoogLeNet [6] and ResNe t[7][8]. Each of these CNNs have a different architecture, thus giving different results for the same application. Thus, the architecture of the CNN must be carefully chosen to maximize the usefulness of it for the specific application.

B. Image Processing Techniques

Image processing techniques are the traditionally used methods for computer visions applications. They involve basic processing techniques such as segmentation, extraction, colour based identification techniques, shape and texture based segmentation amongst a variety of other techniques that target a specific feature of the object or objects that are to be identified in the image. These techniques are applied on a training image set to generate a standard feature set of the target object to be detected. These feature sets are compared to features extracted from images which have objects to be detected or classified using similarity comparison metrics or machine learning techniques such as support vector machines (SVMs).

1) Image processing techniques vs. convolutional neural networks: Since image processing based techniques are fairly simple to develop and tested, multiple methods using hybrid techniques have been proposed for the classification of leather quality and defect identification in them, separately. These methods use a technique or a combination of techniques in quality grading and defect identification. Some of the more advanced techniques that are used in the process are colour based defect detection, wavelet based image analysis for grading the quality of leather, background subtraction and thresholding for defect identification amongst a host of other basic techniques. These techniques, however, do not work well in many generalized cases due to the specific features that are extracted from the training image set, which may not work well in different environments and conditions. Now, with the introduction of deep learning algorithms such as CNNs, the development of classification systems for leather grading and defect detection systems have become simpler. Due to the automated feature learning by these CNNs, the features pertaining to the leather qualities for the grading systems, and the features pertaining to the defects for the defect detection systems can be extracted easily through use of specific CNN algorithms which makes the learning and detection of the texture in case of leather quality and the shape, size and details of the defects in the defect detection system much simpler. As a result of the very generalized approach that CNNs take, it is poised to work accurately in multiple scenarios.

III. Related Work

A method for the detection and classification of surface defects is proposed by Choonjon Kwak et al. [9]. In this method, a two-step segmentation process using thresholding and morphological processing is used for defect location whereas statistical first order and second order features, and geometrical features are used for the classification process. A three stage sequential decision tree classifier is involved in the classification process, mainly, for classifying five types of defects such as lines, stains, holes, knots and wears. The proposed method achieves a good accuracy rate of 91.25 %. In addition to these techniques, wavelet features are also commonly used for identifying defects. A leather defect detection proposal is given by Sobral et al. [10] which uses wavelet transform with a bank of optimized filters, with each filter tuned for detecting one type of defect. Shape and the wavelet sub-bands of the filers are selected to maximize the ratio of feature values on defect regions and normal regions. This defect detection methodology produces fast and accurate results. Another such method for identification of leather defects is proposed by Jawahar et al. [11] using wavelet feature extraction technique. The leather images are captured and processed in the frequency domain using wavelet transform, for easy analysis of uncorrelated pixels and edges and easy highlighting of the frequency separation and image variation, respectively. Further, the defect are identified using wavelet statistical features and wavelet co-occurrences matrix features like Entropy, Energy, Contrast, Correlation, Mean, Cluster Prominence Standard Deviation and Local Homogeneity. The identified defects are classified through a support vector machine classifier, thus, giving robust results. Fu Qiang He [12], also proposed a wavelet based feature extraction technique. A multiresolution approach with energy, entropy matrices for defect detection is used. The defect distinguishing process uses a wavelet band selection procedure, which automatically determines the number of resolution level for decomposition of sub images, removing repetitive texture patterns from the image. This is followed with an adaptive binary algorithm for separating the defective regions. Experimental results prove the fastness and the efficiency of this method. Sze Teng Liong et al. [13] present a deep learning architecture based method for defect detection on leather images. The detected defects are marked using a robotic arm and chalk. This proposed methodology is able to achieve a very high segmentation accuracy rate of 91.5 % on the training data and 70.35 % on the test data.

Kasi et al. [14] propose using auto adaptive edge detection algorithm for defect identification. The methodology uses an edge detection algorithm which focuses on the leather edges with continuity and clarity, removing irrelevant edges and identifies leather defects and achieves good results. Malathy Jawahar et al. [15] propose a leather surface defect detection system which uses a multi-level thresholding algorithm for segmenting defective and non-defective regions from the captured leather surface images. Next, a texture feature extraction algorithm is used for quantification of the leather surface defects. A neural network classifier is then used for classification. The methodology used achieves a good accuracy rate of 90 %. Fan Dahuang et al. [16] propose a leather surface detection algorithm by using ultra-high definition images captured by a camera. A segmentation algorithm that is based on the saturation channel characteristics of the captured image is used for obtain region of interest (ROI). These regions are further used to detect the presence of defects in them using a method of automatic detection and location using image gradients. Experimental results prove that the method detects and locates defects effectively and precisely while outperforming many other traditional methods.

Jang-Woo Kwon et al. [17] propose a texture analysis method for leather quality classification. The proposed method grades the leather quality by using extracting density and defects from a black image. The defects are extracted by using differentiation of histogram distribution from the image pixel, considered as a window and a search window. The evaluation of leather grade is done by using this differentiation process and offers good results. Santos Filho et al. [18] also propose a leather quality classification methodology. Initially, the regions of interest (ROI) is considered, segmented and filtered in three steps. Next, using this ROI, haralick texture features is extracted using a gray level co-occurrence matrix (GLCM). Various descriptors such as energy, homogeneity, contrast, cluster shade, maximal correlation coefficient are used. Experimental results show that the system performs efficient classification of goat leather with accuracy rate reaching 93.22 %. Winiarti et al. [19] propose for the utilization of a pre-trained deep convolutional neural network (CNN) for extracting discrete features from tanning leather image, for classification of leather types. The extracted features are passed into a support vector machine (SVM) which performs the classification process, thus, resulting in good performance.

In the defect classification category, Ding et al. [20] propose a hierarchical classification technique and a defect extraction technique using image processing. Using geometric feature representation, defects are initially divided into dots, lines and surfaces. Next, the dominating characteristics can be collected by analyzing this data through extracting texture, gray and geometry from the defects. Each type of defect is considered individually and their representative characteristics are chosen and dimensions reduced for establishing a database. Characteristics are further converged by clustering in the database. It is then used for defect classification. This methodology of classification results in a high performance system being able to achieve more than 90 % classification accuracy. Jian et al. [21] propose a methodology to classify leather surface defects using a Feed-forward Neural Network (FNN) and a decision tree combination. This allows for the optimal attribute selection and defect classification. The method performs efficient and fast classification of defects. In a comprehensive approach for leather defect classification, Sze Teng Liong et al. [22] propose a combination of deep learning and image processing techniques. Initially, the leather images are partitioned into small patches and pre-processed using a canny edge detection technique. This enhances the visualization of the defects. Next, the extraction of notable features takes place through the use of an artificial neural network (ANN) and convolutional neural network (CNN). Experimental results show that the network achieves a good 80.3 % classification accuracy rate. In another method that identifies defects but on yarn-dyed fabric, Junfeng Jing et al. [23] propose using a convolutional neural network (CNN), modelled on a modified AlexNet architecture, which replaces local response normalization layers with batch normalization layers, to improvise efficiency in computation and classification. Defect extraction takes place through the multiple layers of the CNN and the final classification is got as the end result from the CNN. Experimental results show a good performance of this classification method. Choojong Kwak et al. [24] propose another system for the surface defect detection and classification. Thresholding and morphological processing are used to detect and locate visual defects. Five types of defects that are lines, holes, wears, knots and stains are focused on. For classification of defects, a two multi layered perceptron model is used, with one, two hidden layers. Comparison results with a decision tree approach shows higher classification accuracy achieved by this model. Hoang-Quan Bong et al. [25] propose a leather defect detection and classifying system, focusing on defects such as scars, scratches and pinholes. Several image processing techniques are cascaded and applied for the image feature extraction and defect position location. These collected features are used for the classification of the defect type using a support vector machine (SVM). The method offers good accuracy and speed.

From the analysis of existing systems and methodologies [32], it can be seen that either the system is developed to work on a specific type of leather or is tuned to detect a specific type of defect. In either of these applications, the proposed system cannot work with various other types of leather images and identify any type of defect in them. Also, the dataset that has been used in the majority of the works have not been released or is not available publicly. As a result, this work proposes to develop a highly efficient and powerful leather classification and defect detection system that is not specifically tied to one type of leather and defect.

IV. Proposed Methodology

An efficient system is proposed for the classification of leather images and detection of defects in them. It consists of two single-channel convolutional neural networks, one for classification and the other for locating and marking defects.

The first channel for classification uses Inception V3 [26] deep convolutional neural network for identifying saliency features in the leather images and classifying them as either good or defective. The second channel for locating defects uses Mask R-CNN [27]. If the resultant output of the first channel CNN is defective, the image is passed on to the second channel CNN that locates the defects in the image. By using this methodology, an efficient and faster processing of the leather images classification can be done. This avoids the wastage of crucial time by not trying to locate defects, which takes a considerable amount of time, when the image is of a good quality leather sample. The robustness of this approach rests on the fact that the classifier is able to classify the images with a high accuracy. The flowchart of the proposed system is shown in Fig. 1.

A. Classification CNN

Classification is the process of systematic arrangement of a certain thing in groups or categories according to established conditions or criteria. With respect to images, the main task of classification is the acceptance of the input image and processing it, to find the class of the image. For example, considering a classification of pets as dogs or cats, when an image containing a cat is given as input, the classification system should classify the image as of cat category. Since, images are an array of pixels, recognizing any specific object category by a computer, which cannot understand semantic information in an image is a difficult task. This is where convolutional neural networks are used.

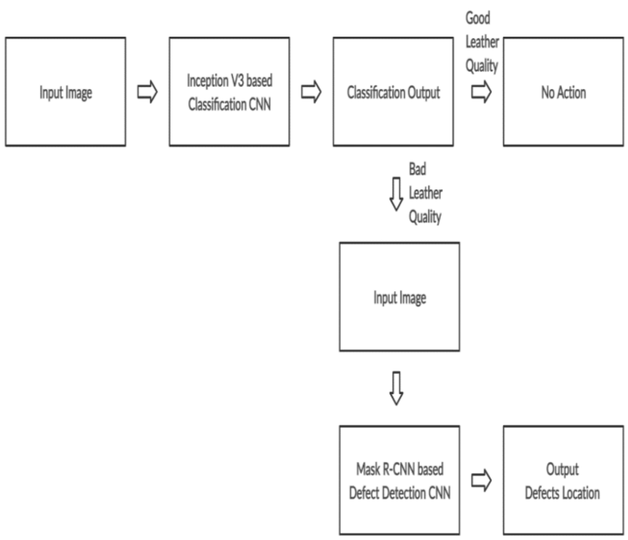


Fig. 1. Flowchart of the proposed system.

Here a deep convolutional neural network based on Inception V3 is proposed. Inception V3 is an improved Inception CNN architecture with 48 layers. It is more efficient than VGGNET and has a computational cost of only about 2.5 times higher than that of GoogLeNet (Inception V1, 27 layers).

For the gathering of all of the salient features in the image, a deep convolution network with multiple filters and layers is required. Since the deepening of the network costs computationally, the Inception module is developed, which has multiple filters for performing multiple convolution operations on the input at the same level of the network, widening it.

The Inception CNN architecture, given in Fig. 2, is based on the inception modules or blocks, which consists of multiple filters in different sizes (1x1, 3x3 and 5x5) for performing convolution operation on the input, at the same level. Each block's end consists of a global average pooling for output concatenation from a single block. Two auxiliary classifiers are attached to the network, for softmax operation on the outputs for total loss adjustment purposes [28].

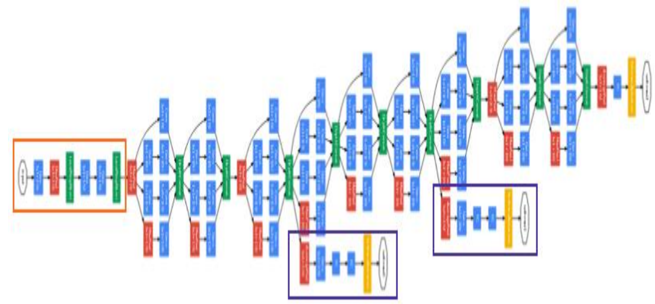


Fig. 2. Inception V3 architecture.

The output of this classifier network is the input image's classification as good quality or defect quality. If the output of this neural network comes out to be as defective quality, then the image is pushed to the next channel, the defect locating CNN, to locate the defects in that image.

B. Defect Detection CNN

Any image can have multiple objects present in it. Object detection is the process of locating each of the object in an image. Images can have multiple objects with same class labels and different objects with different class labels.

Region based CNNs (R-CNNs) have been the most effective type of CNNs due to their architectural design proposing highly confident prospective object regions. R-CNNs work by proposing a bunch of regions or boxes in an image using Selective Search algorithm and checks whether these regions contain any object of interest. Mask R-CNN is the CNN architecture that is proposed for being used in the defect detection CNN channel. It accurately locates the defects in the image and draws optimum bounding boxes around it. The architecture of the CNN is given in Fig. 3.

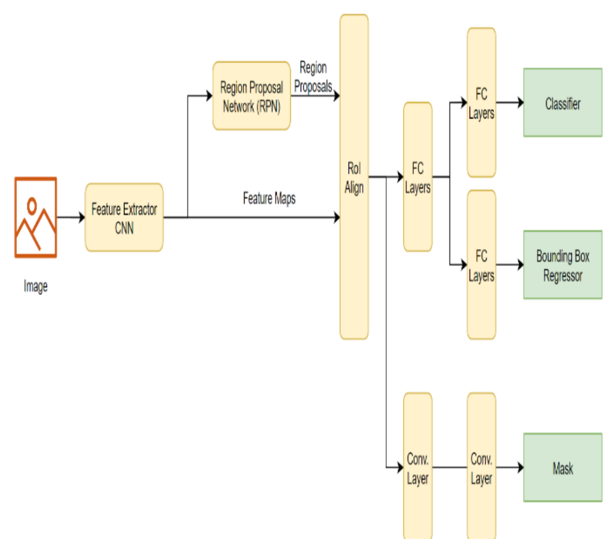


Fig. 3. Mask R-CNN architecture.

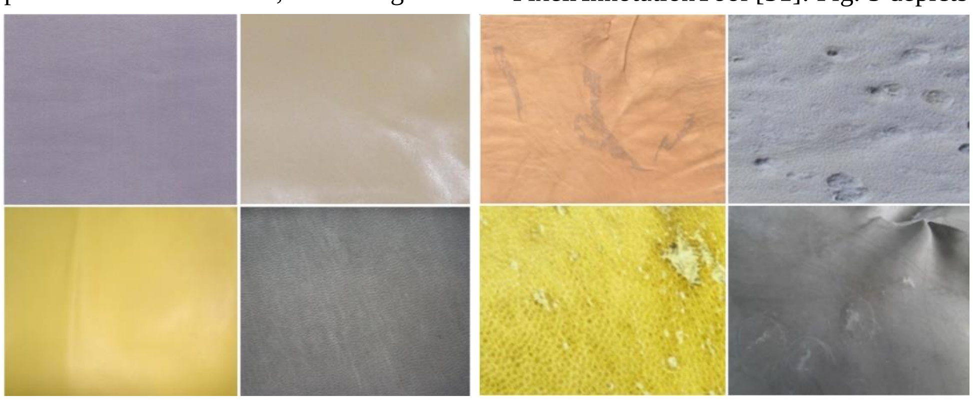
The Mask R-CNN architecture consists of an initial CNN which is a feature extractor. The feature extractor that we use in this work is Inception V2. This feature extractor CNN works on the image to produce feature maps. These feature maps correspond to the salient features detected in the image. The feature map is parallelly passed to a Region Proposal Network (RPN) to generate regions of interest, wherein prominent candidate objects of interest in the feature maps are identified and marked. These marked regions on the feature maps is passed to an ROI Align layer, along with the original generated feature maps, to align the proposed candidate regions of interest with the feature maps effectively. This effective mapping is the reason why Mask R-CNN is preferred over other R-CNNs such as Faster R-CNN, Fast R-CNN, etc. Another reason is the optimization of the bounding boxes, to minimize the area required to be removed due to a defect. These aligned maps are then passed to fully connected layers for class label prediction, bounding box fit and masks.

Thus, the output from this CNN is the image in which the defects are identified and each of them are marked with bounding boxes. Masked outputs are disabled due to non-use in this particular scenario.

Emphasizing the potential for collaboration with industry stakeholders, technology providers, or research institutions is crucial for further refining and validating the proposed leather defect detection system using deep learning. Collaborative efforts can facilitate the exploration of real-world implementation scenarios, allowing for comprehensive testing and validation of the system's effectiveness across diverse production environments. Moreover, knowledge exchange between academia and industry can lead to insights and innovations that advance automated quality control practices not only within the leather industry but also in other sectors. By fostering collaborative partnerships, stakeholders can collectively contribute to the evolution and adoption of cutting-edge technologies, driving continuous improvement and enhancing operational efficiencies in quality control processes on a broader scale.

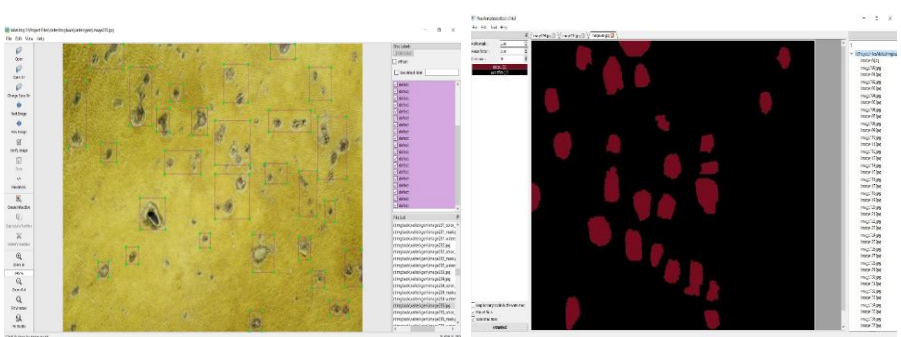
V. EXPERIMENTAL TESTS AND ANALYSIS

The availability of useful leather datasets that can be used for training and testing techniques have always been a huge issue. Most of the already handful of available datasets, have either a low quantity of images that can be used for training and testing, or do not contain quality images. The dataset that has been generated by us, using data augmentation method. The images are available in [29], samples of which are shown in Fig. 4. Originally, it consisted of 428 good quality and 354 defect quality leather images collected from leather industry. By using, image augmentation techniques such as scaling, cropping, flipping, grayscale, brightness, contrast, saturation, etc., a total of 29716 images are produced and packed in the dataset. The original 354 defective images used and each of the defect in them are manually annotated by a tool called LabelImg [30] and mask created by a tool called PixelAnnotationTool [31]. Fig. 5 depicts the process.



(a) (b)

Fig. 4. Sample dataset images of (a) Good quality leather and (b) Defect quality leather.



(a) (b)

Fig. 5. Dataset (a) Annotation using labelImg (b) Mask creation using pixel annotation tool.

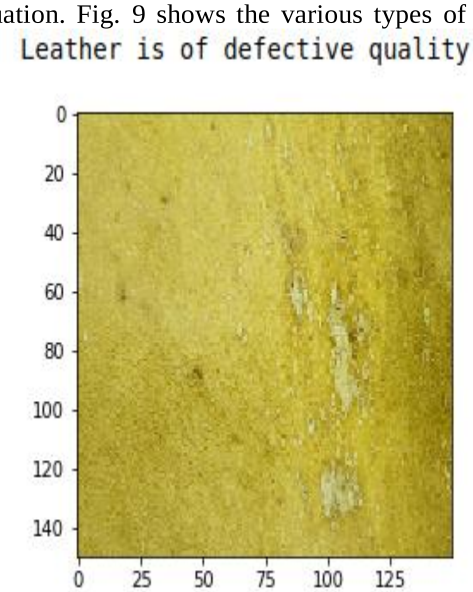
The experiments were carried out on a machine with Ubuntu 16.04 as the Operating System, 16 GB RAM, Intel i5 processor and a Nvidia GTX 1060 GPU with 6GB VRAM. The integrated leather quality classification and defect detection system was implemented using Python, Tensorflow and Keras. Since the system uses two single channel convolutional neural networks, one for classification and one for defect detection, a two-step training process is followed with unique datasets.

The first classification CNN is fed with the categorized images with 23772 training images with 13004 good quality and 10768 defect quality leather images. The validation set of 5944 images consists of 3260 good quality and 2684 defect quality leather images. The network is trained for 200 epochs with the image batch size being 32. This network achieves a classification accuracy of 99 % on the training data and 84.6 % on the validation data. The second defect detection network consists of an object detection CNN and as a result expects annotated images as the training dataset. Thus, as a result, a total of 224 defect leather quality images were annotated and corresponding masks created, and the same augmentation techniques of colour, saturation, rotation, hue change, flipping, etc., were performed on a total of 530 epochs for a total of 1,20,000 training steps, with batch size set as 1. The training set consists of 162 images and the validation set consists of 62 images, both containing annotation and mask data for each image. This network obtained a detection rate of 99.6 % on the training set, 99 % on the validation dataset and 99 % on a 20 image test dataset. The system results from the experiments are listed in Table I, with the samples shown in Fig. 6, defects detected shown in Fig. 7, training and validation accuracy graphs shown in Fig. 8 and various types of loss functions shown in Fig. 9. Fig. 7 on a closer inspection reveal that the defective spots are identified with high precision and accuracy. At some spots the accuracy is around 60%, where in a suitable selection of threshold value can help to detect the defects with increased precision in a real time situation. Fig. 9 shows the various types of loss the model incurs during the training phase of the proposed RCNN model. X-axis indicates the samples in the epoch and y-axis indicates the loss value. It can be noted that in all the loss graphs, the loss has been steadily decreasing and even approaches 0, beyond certain epochs. This means the model has converged to the leather images in detecting the defects. Hence it can be observed that the proposed model performs a promising automated defect detection of the leather from machine vision approach.

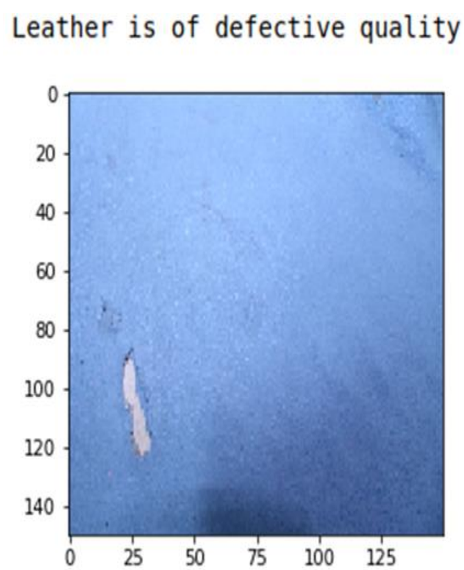
The proposed leather defect detection system offers several advantages over previous approaches, primarily through the integration of advanced deep learning techniques. Unlike traditional methods that rely on manual inspection or basic machine vision algorithms, the proposed system leverages sophisticated deep learning models such as the Inception model for classification and the RCNN model for detection. This enables the system to achieve higher accuracy in defect identification and localization, addressing the limitations of human subjectivity and the inability of older systems to detect subtle or complex defects reliably. Moreover, the automated nature of the proposed system improves efficiency, reduces labor costs, and ensures consistent and scalable defect detection across different production environments. Overall, the integration of deep learning technologies in the proposed system leads to superior performance, increased reliability, and enhanced quality control capabilities compared to previous approaches.

TABLE I. RESULTS OF THE PROPOSED SYSTEM

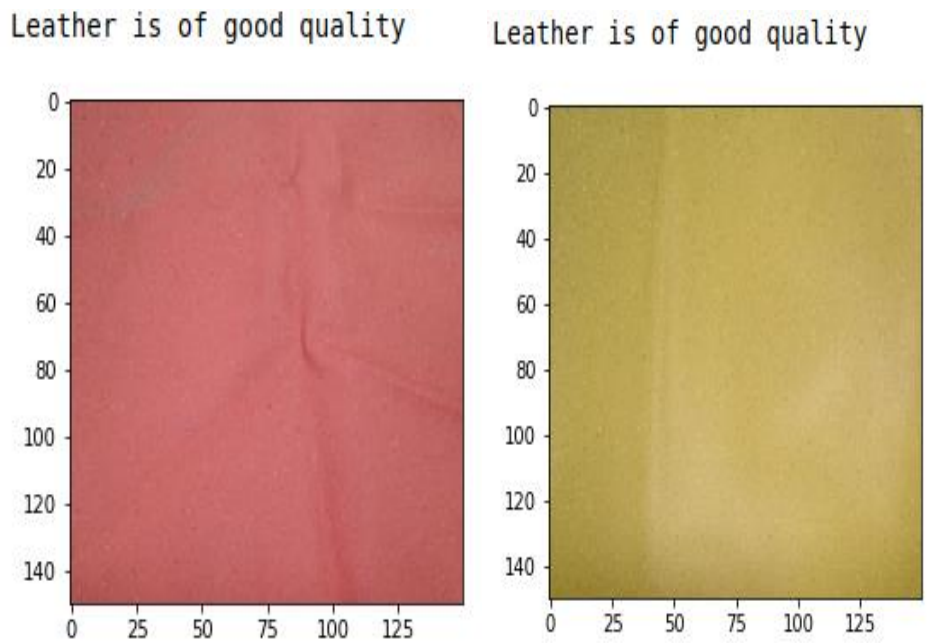
Methodology	Dataset	Total number of Images	Accuracy Rate
Classification	Train	23772	99 %
Classification	Validation	5944	84.6 %
Defect Detection	Train	162	99.6 %
Defect Detection	Validation	62	99 %
Defect Detection	Test	20	99 %



(a)

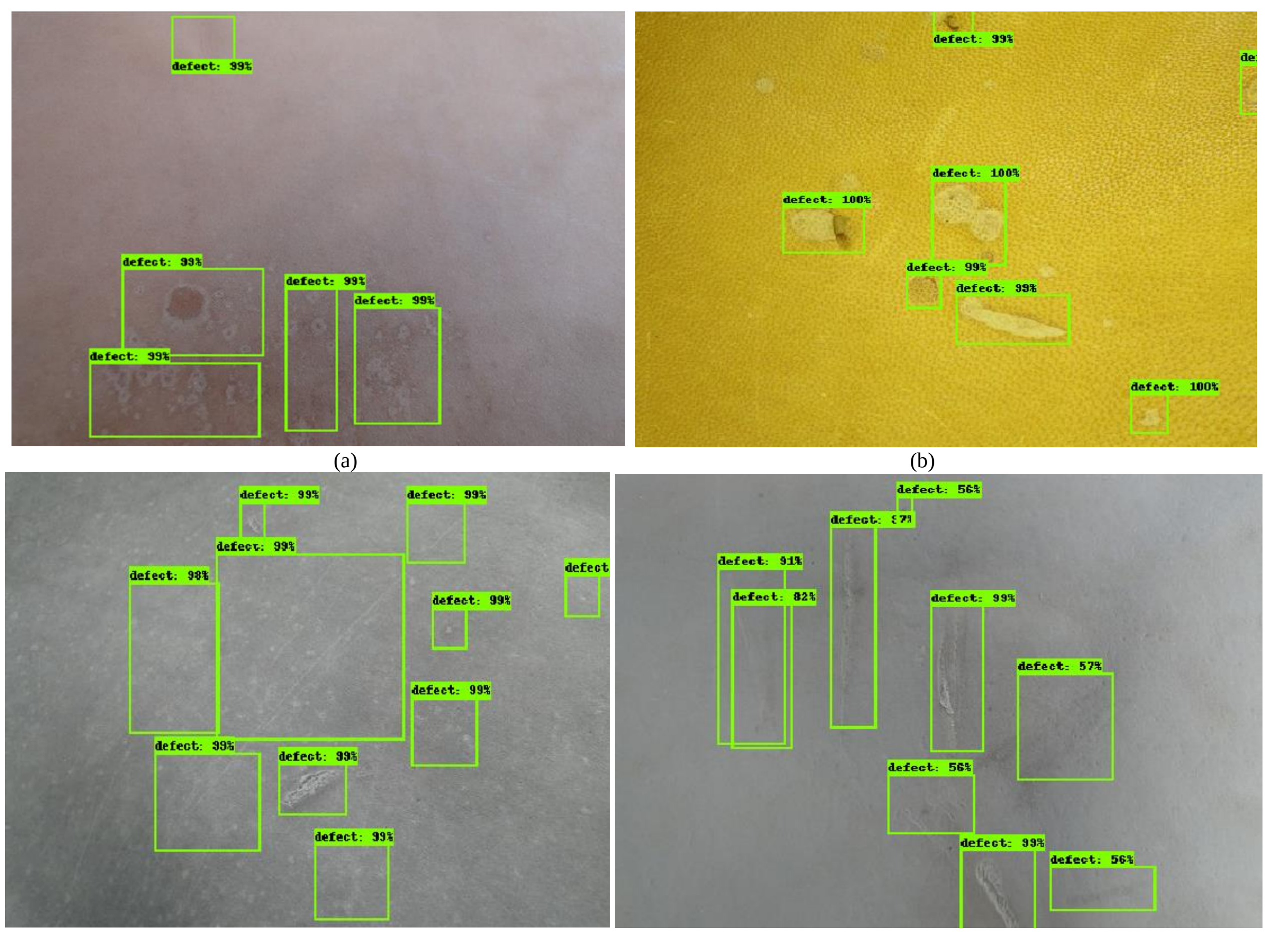


(b)



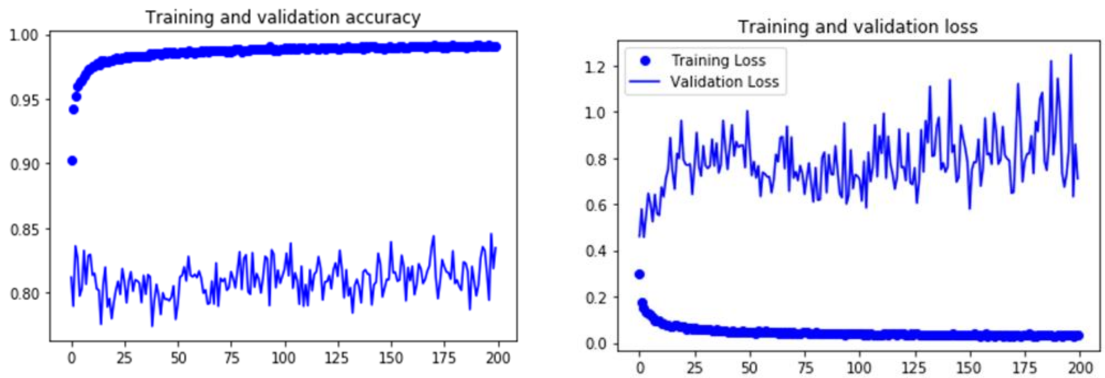
(c) (d)

Fig. 6. (a) Defective quality leather sample 1; (b) Defective quality leather sample 2; (c) Good quality leather sample 1; (d) Good quality leather sample 2.



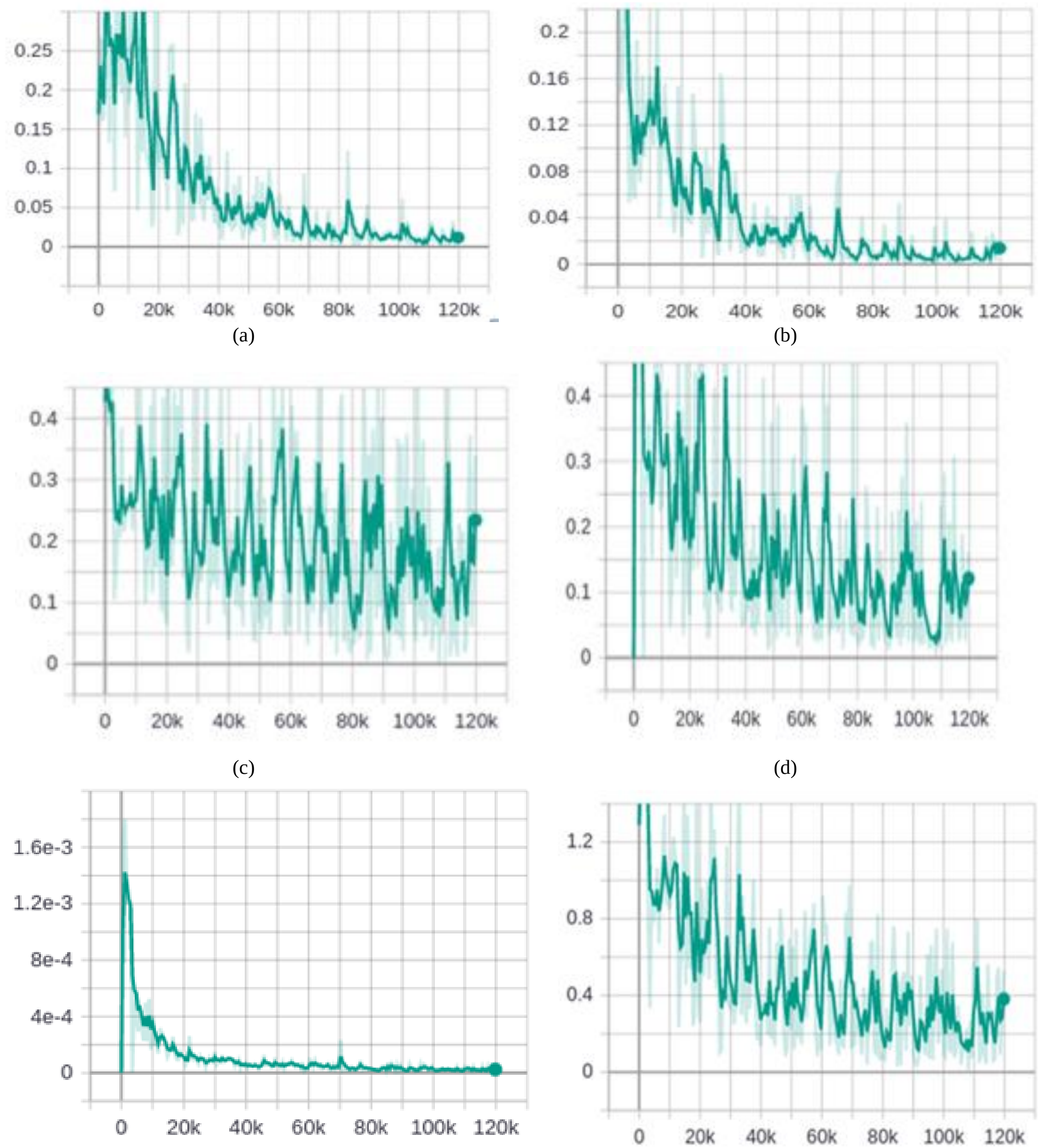
(c) (d)

Fig. 7. Leather image defect detection sample results.



(a) (b)

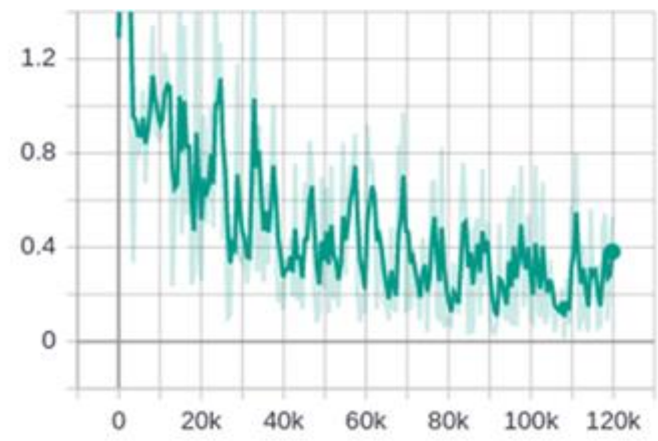
Fig. 8. Graphs for training and validation phases of the proposed model – (a) Accuracy graph; (b) Loss graph.



(a) (b)

(c) (d)

(e) (f)



(g)

Fig. 9. Various loss graphs during defect detection training of the proposed model – (a) RPN localization loss; (b) Objectness loss; (c) Classification loss; (d) Localization loss ; € mask loss; (f) Clone loss; (g) Total loss.

VI. Conclusion

This paper presented a comprehensive system for the classification and detection of defects in leather images using convolutional neural networks. Initially, the basic image processing approaches and their applications were discussed in this field. With the introduction of deep learning methods, convolutional neural networks are the most preferred way in this field due to their unique automated feature extraction and learning approaches. The recent works using both of these approaches was discussed. The proposed system uses a double channel method, where one channel has a Inception V3 convolutional neural network which classifies the leather image as “good quality” or “defect quality” while the other channel has Mask R-CNN which detects and locates the defects in the leather image and draws bounding boxes for candidate defect regions, when the output of the first channel is “defect quality”. Experimental results show that the system achieves a high performance with the classification CNN achieving a 99 % accuracy rate on the training dataset and 84.6 % on the validation dataset. The defect detection CNN achieves 99.6 % defect detection accuracy on the training dataset, 99 % on the validation dataset and 99% accuracy on the test dataset. Thus, the system is proved to be able to accurately classify leather images and identify and locate defects in images, where only those images which are classified as “defective”, are sent to the second channel defect detection convolutional neural network for detection of defects, thereby making the process much faster.

Researchers could make leather defect detection systems using deep learning even better in the future. They could add advanced technologies like reinforcement learning and transfer learning. Reinforcement learning helps the system make good decisions and adapt to new situations. Transfer learning uses knowledge from related areas to improve defect detection accuracy. The systems could also be used in other industries besides leather manufacturing, like textiles or car manufacturing. This would allow automating quality control processes to boost efficiency across different sectors. Exploring these avenues can enhance automated defect detection systems' capabilities and versatility. This could lead to more industries adopting the technology and drive innovation.

Acknowledgment

Fuad A. M. Al-Yarimi extend his appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through the General Research Project under grant number (R.G.P1/181/43).

References

[1] Kohli, Parag. (2013). Leather Quality Estimation Using an Automated Machine Vision System. IOSR Journal of Electronics and Communication Engineering. 6. 44-47. 10.9790/2834-0634447.

[2] Barik, Debabrata. (2019). Introduction to Energy From Toxic Organic Waste For Heat and Power Generation. 10.1016/B978-0-08-102528-4.00001-8.

[3] Milosevic, N. (2020) “Convolutions and Convolutional Neural Networks,” Introduction to Convolutional Neural Networks [Preprint]. Available at: https://doi.org/10.1007/978-1-4842-5648-0_12.

[4] Mueller, J. and Massaron, L. (2019) in Deep learning for dummies. Hoboken, NJ: John Wiley & Sons, Inc.

[5] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) “ImageNet classification with deep convolutional Neural Networks,” Communications of the ACM, 60(6), pp. 84–90. Available at: https://doi.org/10.1145/3065386.

[6] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S.E., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2014). Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1-9.

[7] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770-778. https://doi.org/10.1109/CVPR.2016.90.

[8] He, K., Zhang, X., Ren, S., Sun, J. (2016). Identity Mappings in Deep Residual Networks. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science(), vol 9908. Springer, Cham. https://doi.org/10.1007/978-3-319-46493-0_38.

[9] Kwak, C., Ventura, J.A., & Tofang-Sazi, K. (2001). Automated defect inspection and classification of leather fabric. Intell. Data Anal., 5, 355-370.

[10] Sobral, J.L. (2005). Leather Inspection Based on Wavelets. Iberian Conference on Pattern Recognition and Image Analysis.

[11] Jawahar, M., Chandra Babu, N.K., & Vani, K. (2014). Leather texture classification using wavelet feature extraction technique. 2014 IEEE International Conference on Computational Intelligence and Computing Research, 1-4.

[12] He, F.Q., Wang, W., & Chen, Z. (2006). Automatic Visual Inspection for Leather Manufacture. Key Engineering Materials, 326-328, 469 - 472.

[13] Liong, S., Gan, Y.S., Huang, Y., Yuan, C., & Chang, H. (2019). Automatic Defect Segmentation on Leather with Deep Learning. ArXiv, abs/1903.12139.

[14] Kasi, M.K., Rao, J.B., & Sahu, V.K. (2014). Identification of leather defects using an autoadaptive edge detection image processing algorithm. 2014 International Conference on High Performance Computing and Applications (ICHPCA), 1-4.

[15] Jawahar, M., Vani, K., & Chandra, N. (2019). Machine Vision Inspection System for Detection of Leather Surface Defects. Journal of The American Leather Chemists Association, 114, 10-19.

[16] Dahuang, F., Lei, D., & Jiehang, D. (2019). Automatic Detection and Localization of Surface Defects for Whole Piece of Ultrahigh-definition Leather Images. 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS), 229-232.

[17] Kwon, J., Choo, Y., Choi, H., Cho, J., & KiI, G. (2004). Development of leather quality discrimination system by texture analysis. 2004 IEEE Region 10 Conference TENCON 2004., A, 327-330 Vol. 1.

[18] Filho, E.Q., Sousa, P.H., Filho, P., Barreto, G.D., & Albuquerque, V.H. (2020). Evaluation of Goat Leather Quality Based on Computational Vision Techniques. Circuits, Systems, and Signal Processing, 39, 651-673.

[19] Winiarti, S., Prahara, A., Murinto, & Ismi, D.P. (2018). Pre-Trained Convolutional Neural Network for Classification of Tanning Leather Image. International Journal of Advanced Computer Science and Applications, 9.

[20] Ding, C. & Huang, H. & Yang, Y.. (2018). Description and Classification of Leather Defects Based on Principal Component Analysis. Journal of Donghua University (English Edition). 35. 473-479.

[21] Jian, L., Wei, H., & Bin, H. (2010). Research on inspection and classification of leather surface defects based on neural network and decision tree. 2010 International Conference On Computer Design and Applications, 2, V2-381-V2-384.

[22] Liong, S., Gan, Y.S., Liu, K., Binh, T.Q., Le, C.T., Wu, C., Yang, C., & Huang, Y. (2019). Efficient Neural Network Approaches for Leather Defect Classification. ArXiv, abs/1906.06446.

[23] Jing, J., Dong, A., & Li, P. (2017). Yarn-dyed fabric defect classification based on convolutional neural network. International Conference on Digital Image Processing.

[24] Kwak, C., Ventura, J.A., & Tofang-Sazi, K. (2000). A neural network approach for defect identification and classification on leather fabric. Journal of Intelligent Manufacturing, 11, 485-499.

[25] Bong, H., Truong, Q.B., Nguyen, H., & Nguyen, M.T. (2018). Vision-based Inspection System for Leather Surface Defect Detection and Classification. 2018 5th NAFOSTED Conference on Information and Computer Science (NICS), 300-304.

[26] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2015). Rethinking the Inception Architecture for Computer Vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2818-2826.

[27] He, K., Gkioxari, G., Dollár, P., & Girshick, R.B. (2017). Mask R-CNN. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42, 386-397.

[28] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A.A. (2016). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. ArXiv, abs/1602.07261.

[29] LeatherImageDataSet. (2023). https://github.com/AkshRSH/LeatherImageDS (accessed 20 December 2022).

[30] LabelImg, Github. (2020). https://github.com/tzutalin/labelImg (accessed 20 March 2020).

[31] PixelAnnotationTool, Github. (2020). https://github.com/abreheret/PixelAnnotationTool (accessed 20 December 2022).

[32] Jawahar, Malathy & Anbarasi, L. & Subbiah, Geetha. (2022). Vision based leather defect detection: a survey. Multimedia Tools and Applications. 82. 1-27. 10.1007/s11042-022-13308-x.

[33] S.-T. Liong, D. Zheng, Y.-C. Huang, and Y. S. Gan, “Leather defect classification and segmentation using deep learning architecture,” Int. J. Comput. Integr. Manuf., vol. 33, no. 10–11, pp. 1105–1117, Nov. 2020, doi: 10.1080/0951192X.2020.1795928.

[34] L. Jian, H. Wei, and H. Bin, “Research on inspection and classification of leather surface defects based on neural network and decision tree,” in 2010 International Conference On Computer Design and Applications, 2010, vol. 2, pp. V2-381-V2-384, doi: 10.1109/ICCDA.2010.5541405.

[35] S.-T. Liong, Y. S. Gan, Y.-C. Huang, K.-H. Liu, and W.-C. Yau, “Integrated Neural Network and Machine Vision Approach For Leather Defect Classification,” CoRR, vol. abs/1905.1, 2019, [Online]. Available: http://arxiv.org/abs/1905.11731.

[36] Y. S. Gan, S.-S. Chee, Y.-C. Huang, S.-T. Liong, and W.-C. Yau, “Automated leather defect inspection using statistical approach on image intensity,” J. Ambient Intell. Humaniz. Comput., vol. 12, no. 10, pp. 9269–9285, 2021, doi: 10.1007/s12652-020-02631-6.

[37] G. Pazzaglia, M. Martini, R. Rosati, L. Romeo, and E. Frontoni, “A Deep Learning-Based Approach for Auto