Deep Learning-Driven Machine Vision for Precise Leather Defect Detection

Dr. David Raj Micheal
Division of Mathematics
School of Advanced Sciences
Vellore Institute of Technology Chennai
Tamil Nadu – 600127
davidraj.micheal@vit.ac.in

S.Akshaya
Division of Mathematics
School of Advanced Sciences
Vellore Institute of Technology Chennai
Tamil Nadu – 600127
akshaya.s2023c@vitstudent.ac.in

***Abstract*—-The leather industry is a significant global sector that spans various markets, including fashion, automotive, footwear, furniture, and accessories. In this project we focuses on automated leather defect detection, a crucial application in industries that require high-quality leather products.The primary objective is to classify various types of leather defects accurately, using deep learning techniques, to improve quality control and reduce manual inspection efforts. This system is particularly valuable as leather defect detection is labor-intensive and error-prone, making automation beneficial for consistency and efficiency. By utilizing advanced convolutional neural networks (CNNs) and implements popular models, including a custom CNN, ResNet, VGG, and AlexNet, to analyze and classify six classes of leather defects: folding marks, grain off, growth marks, loose grains, non-defective, and pinholes. The models were evaluated based on their accuracy in identifying defects, and their training performance was visualized to compare effectiveness. This automated defect classification system aims to enhance quality assessment procedures in leather processing, offering a scalable solution that is less susceptible to human error.**

***Index Terms*—Keywords- Leather defect detection, image processing, Deep learning techniques, Convolutional neural networks (CNN), Accuracy.**

I. Introduction

Leather defect detection and classification using image processing involves the application of advanced computer vision techniques to automatically identify and categorize imperfections on leather surfaces. During the manufacturing process, various kind of defects can occur, compromising the quality and value of the leather. Traditional manual inspection methods are often inconsistent and subjective, leading to the need for more precise and efficient automated systems. A key approach in this process is the use of Convolutional Neural Networks (CNNs), a deep learning model designed for image analysis. CNNs are particularly effective for defect detection because they can automatically learn and recognize the complex texture patterns and small, ambiguous regions where defects typically occur. By training CNNs on labeled images of leather with various defects, the model can accurately detect and classify different types of imperfections. The purpose of the CNN-based model is to automate the detection and classification process, providing a faster, more consistent, and scalable solution compared to manual inspection. CNNs help localize the regions of interest related to specific defects and improve classification accuracy, ultimately enhancing the overall quality control in leather manufacturing.

II. Objectives

The objective of this study is to develop and implement a Convolutional Neural Network (CNN)-based model for the automated detection and classification of defects in leather surfaces. By leveraging advanced image processing techniques, the model aims to enhance the accuracy and efficiency of quality control in leather manufacturing. Specifically, the study seeks to train the CNN on a diverse dataset of labeled leather images to enable precise identification of various imperfections, thereby reducing reliance on traditional manual inspection methods. Ultimately, the goal is to provide a scalable and consistent solution that improves defect localization and classification, contributing to superior product quality and value in the leather industry.

III. Literature Review

P.K., Seelan, D.A.S. (2020) This study addresses the challenge of detecting leather defects by using texture analysis and neural network classification. It employs the Gray Level Co-occurrence Matrix (GLCM) to extract texture features and trains a perceptron neural network to identify common defects, achieving a 94.2% accuracy with a dataset of 1232 leather images. The method offers a viable solution for automating leather inspection in industrial environments.**M. Praveen Kumara and S. Denis Ashok(2020)** This study introduces a color image processing approach to improve leather defect detection by using multi-level thresholding in the 'Lab' color domain, enhancing defect visibility and discrimination. By analysing color attributes through histograms, the method refines defect identification and simplifies the process compared to traditional grayscale techniques. The approach, implemented in MATLAB, shows improved accuracy and can be adapted for real-time automated inspection using graphical processing units.

Liong ,S. T., Zheng, D., Huang, Y. C., Gan, Y. S. (2020) This paper presents an automatic leather defect detection system using deep learning, which includes classification to identify defective samples and instance segmentation to locate defects precisely. The system, tested on a dataset of 250 defective and 125 non-defective samples, achieved about 95% accuracy in classification and 99.84% Intersection over Union (IoU) for defect localization. It effectively detects black lines and wrinkles, showcasing strong performance with relatively few training samples.**Amorim ,W. P., Pistori, H., Pereira, M. C., Jacinto, M. A. C. (2010)** This paper investigates attribute reduction for leather defect classification by comparing five discriminant analysis techniques—FisherFace, CLDA, DLDA, YLDA, and KLDA. The study evaluates these techniques with four classifiers: C4.5, kNN, Naïve Bayes, and Support Vector Machines. Experimental results demonstrate the effectiveness of various discriminant analysis methods in improving defect detection performance.

Chen, Z., Deng, J., Zhu, Q., Wang, H., Chen, Y. (2022) This paper offers a comprehensive review of machine vision-based methods for inspecting leather surface defects, highlighting the importance of automatic detection, location, and recognition in intelligent manufacturing. It evaluates various edge and threshold detectors, as well as the performance of the classical machine learning method SVM for defect identification. The review addresses key challenges and future trends, providing valuable insights for developing advanced solutions in leather defect inspection.**Aslam, M., Khan, T. M., Naqvi, S. S., Holmes, G., Naffa, R. (2020)** This paper addresses the challenges in automated leather defect classification, including the variability of defect morphology and the lack of available data. It introduces a new dataset of annotated wet-blue leather images (WBLID) and discusses methods for transferring information from different domains. The study proposes an ensemble network, EfficientNet-B3+ ResNext-101, which achieves superior performance in defect classification, outperforming existing methods in AUC and F1-score.

Chen, Z., Zhu, Q., Zhou, X., Deng, J., Song, W. (2024) This paper examines the use of the YOLO real-time detection model for quickly identifying and localizing leather surface defects to enhance industrial efficiency. It tests YOLO models from YOLOv1 to YOLOv8 on images with eight types of defects, achieving up to 52.3% mean average precision (mAP) and 68.7% recall for multi-defect detection, and up to 85.1% mAP and 90.9% precision for single-class detection. The study provides effective intelligent solutions for leather defect detection and sets a foundation for further advancements in this field. **Jawahar, M., Vani, K. (2019)** This research tackles the problem of automating leather defect classification to minimize the inconsistencies of manual inspections. It utilizes a machine vision system combined with an innovative multilevel thresholding technique to differentiate between defective and non-defective leather areas, and extracts texture features for analysis. Evaluated with a dataset of 90 images, the system achieved 90% accuracy with a neural network classifier, proving its efficiency in automatic leather defect detection.

Liong, S. T., Gan, Y. S., Huang, Y. C., Liu, K. H., Yau, W. C. (2019) The goal of this study is to create an automated system for classifying tick-bite defects on calf leather. The system integrates both handcrafted feature extraction techniques and data-driven approaches using artificial neural networks. It employs multiple classifiers, including decision trees and Support Vector Machines, to accurately detect defects, achieving an 84% classification accuracy with a dataset of around 2500 leather patches. **Khanal, S. R., Silva, J., Magalhaes, L., Soares, J., Gonzalez, D. G., Castilla, Y. C., Ferreira, M. J. (2022)** The objective of this research is to develop an automated leather defect detection system using machine vision techniques to replace manual inspection. The system utilizes a conveyor platform, camera, and deep learning-based semantic segmentation to achieve a 94% Intersection over Union (IoU) for accurate defect detection on the MVTEC leather dataset.

Chen, Z., Xu, D., Deng, J., Chen, Y., Li, C. (2023) The objective of this study is to evaluate and compare 26 classical deep learning models for recognizing various types of leather surface defects. By using ultra-high definition imaging and diverse datasets, the research aims to identify the most effective models for defect classification, with a focus on improving accuracy and addressing challenges in detecting defects with varying shapes, sizes, and colors.**Jawahar, M., Babu, N. C., Vani, K. L. J. A., Anbarasi, L. J., Geetha, S. (2021)** The objective of this study is to develop and evaluate a computer vision system for leather surface defect inspection using a novel Fast Convergence Particle Swarm Optimization (FCPSO) algorithm. The system aims to enhance defect segmentation and classification by combining handcrafted texture features with various supervised classifiers, and to assess the effectiveness of the FCPSO algorithm compared to other optimization techniques.

Gan, Y. S., Chee, S. S., Huang, Y. C., Liong, S. T., Yau, W. C. (2021) The objective of this study is to develop a fully-automated leather defect detection system using image processing techniques, specifically gray level histogram analysis. The method aims to improve efficiency and consistency in leather quality control by extracting and selecting histogram-based features, applying dimensionality reduction, and classifying defects with high accuracy using various classifiers.**Chudzik, S. (2020)** The objective of this study is to develop and evaluate a novel method for detecting internal and external structural defects in natural, tanned hides using thermal excitation and infrared imaging. The approach aims to identify flaws in leather by analyzing surface temperature distributions during heating and cooling, with the goal of advancing industrial devices for quality control in luxury upholstered furniture.

Iqbal, S., Khan, T. M., Naqvi, S. S., Holmes, G. (2023) This research aims to improve industrial quality control for wet-blue leather by developing an automated defect detection system. The study introduces MLR-Net, a deep learning model, to analyze and classify leather surfaces based on visual features. MLR-Net shows high performance, with strong sensitivity, specificity, accuracy, and IoU, making it effective for accurately identifying and segmenting defects in leather images.

IV. METHODOLOGY

The methodology for the project ”Deep Learning-Driven Machine Vision for Precise Leather Defect Detection” involves several key steps, from data preparation and model development to training, evaluation, and visualization. Below is a detailed explanation of each step:

A. *Data Structure*

The Leather Defect image dataset is designed to facilitate the identification and classification of various defects found in leather materials. Understanding the nature and impact of each class is crucial for developing an effective machine learning model that can accurately classify leather defects.The dataset consists of images categorized into distinct classes, each representing a specific type of defect or quality of leather.

B. *Data Preprocessing*

Image Scaling and Augmentation To prepare the images for model training, the ImageDataGenerator class from Keras is utilized. The images are rescaled by a factor of 1/255 to normalize pixel values from a range of 0-255 to 0-1. This normalization is a common practice in CNN training as it aids in model convergence. **Training and Validation Split** A validation split of 20% is applied to the dataset. This allows 80% of the data to be used for training while reserving 20% for validation, which helps in assessing model performance and mitigating overfitting. **Generator Configuration** The train_generator and validation_generator are configured to load images in batches of 32, resizing each image to dimensions of (227, 227) pixels. This uniform image size is crucial for input into the neural networks and optimizes memory usage during processing. The class mode is set to categorical, which is appropriate for multi-class classification tasks.

C. *Data Selection and Visualization*

Batch Sampling for Inspection One batch of images and labels is retrieved from the generator to inspect their shapes. For example, a shape of (32, 227, 227, 3) indicates that each batch contains 32 images, each resized to (227, 227) pixels with 3 color channels (RGB).

Visualization of Sample Images To facilitate a quick visual inspection, nine sample images are plotted alongside their class labels. This step is critical for verifying class distribution, assessing the effects of data augmentation, and identifying any irregularities within the dataset.

Label Shape Confirmation The shape of the labels is checked to ensure they conform to the multi-class structure, resulting in a shape of (32, 6), indicating a batch of 32 images with 6 possible categories.

D. *Data Selection and Visualization*

Convolutional Neural Network (CNN)

- **Convolutional Layers:** Three layers with an increasing number of filters (32, 64, and 128), each using a 3x3 kernel and ReLU activation to capture spatial hierarchies in the images.
- **Pooling Layers:** Max pooling layers follow each convolutional layer to downsample feature maps, reducing computational load while preserving essential features.
- **Dense Layers:** After flattening the feature maps, a fully connected layer with 128 units and ReLU activation is included, followed by an output layer with softmax activation for multi-class classification.

ResNet Model

The ResNet architecture is particularly effective for image classification tasks due to its ability to learn residual mappings, which helps in training deeper models. The use of pretrained weights allows the model to leverage knowledge learned from large datasets, improving performance on the leather defect classification task.

- **Input Layer:** Accepts images resized to (227, 227, 3) pixels.
- **Convolutional Layers:** A series of convolutional layers with varying filter sizes and strides, designed to extract features at different levels of abstraction.
- **Residual Blocks:** Each residual block includes skip connections that allow gradients to flow through the network more easily during backpropagation. This helps in training deeper networks without suffering from vanishing gradients.
- **Global Average Pooling:** Reduces the dimensionality of the feature maps before passing them to the dense layers.
- **Dense Layers:** Hidden Layer-A fully connected layer with 128 units and ReLU activation.Output Layer-A dense layer with softmax activation for multi-class classification.
- **Fine-tune:** Fine-tuning is a technique in machine learning that adapts a pre-trained model, which has been trained on a large dataset, to a new, often smaller dataset for a specific task. By freezing certain layers and allowing others to adjust, fine-tuning retains valuable learned features while enabling the model to learn from new data. This method is particularly useful when labeled data is limited, as it enhances performance and accelerates convergence, making it a popular strategy in transfer learning.
- **Early stopping:** Early stopping is a regularization technique used in machine learning to prevent overfitting or underfitting by monitoring validation loss during training. If the validation loss does not improve for a specified number of epochs, known as the ”patience” parameter, training is halted. This method not only enhances model generalization but also conserves computational resources by avoiding unnecessary training. By saving the model state with the lowest validation loss, early stopping helps

maintain predictive power while balancing model complexity.

VGG Model

The VGG architecture is designed to be deep and uniform, using small convolutional filters (3x3) throughout. This design choice allows the model to capture fine details in images while maintaining computational efficiency. The depth of the model enables it to learn complex patterns necessary for accurate classification.

- **Input Layer:** Accepts images resized to (227, 227, 3) pixels.
- **Convolutional Layers:** A stack of convolutional layers (e.g., 3x3 kernels) with increasing depth, typically starting with 64 filters and doubling the number of filters after each block.
- **Max Pooling Layers:** Following groups of convolutional layers, max pooling layers are used to downsample the feature maps.
- **Flattening Layer:** Converts the pooled feature maps into a 1D vector.
- **Dense Layers:** Hidden Layer-A fully connected layer with 256 units and ReLU activation. Output Layer-A dense layer with softmax activation for the six defect categories.

AlexNet Model

The architecture consists of eight layers, including five convolutional layers followed by three fully connected layers. Key features of AlexNet include the use of Rectified Linear Unit (ReLU) activation functions, which help mitigate the vanishing gradient problem, and dropout layers that reduce overfitting by randomly deactivating a portion of the neurons during training. Additionally, AlexNet effectively utilizes data augmentation techniques to enhance the training dataset and leverages GPU acceleration for faster training. Its innovative design and success have laid the foundation for subsequent deep learning architectures, making it a cornerstone in the evolution of image classification tasks.

E. Model Compilation

Each model is compiled using the Adam optimizer, known for its adaptive learning rate capabilities, which can enhance convergence speed and overall accuracy. The loss function employed is categorical crossentropy, suitable for multi-class classification tasks where each instance is assigned a single label. Model performance is assessed using accuracy metrics for both training and validation datasets.

F. Model Training

The model training process is a critical phase in the development of the Leather Defect Classification system. It involves the iterative optimization of the model parameters to minimize the loss function and improve classification accuracy. The following steps outline the training process in detail:

Training Configuration

- **Epochs:** Each model is trained for a total of 10 epochs. An epoch is one complete pass through the entire training dataset. The choice of 30 epochs is based on preliminary experiments that indicate sufficient training without overfitting.
- **Batch Size:** The batch size is set to 32, meaning that during each iteration, 32 images are processed before the model's weights are updated. This batch size balances the efficiency of training and the stability of gradient updates.
- **Validation:** At the end of each epoch, the model's performance is evaluated on the validation dataset. This evaluation provides insights into how well the model is generalizing to unseen data.
- **Loss Function and Metrics:** The model uses categorical crossentropy as the loss function, which is suitable for multi-class classification problems. The accuracy metric is tracked during training to monitor the model's performance.

V. RESULTS AND FNDINGS

The performance of four different convolutional neural network (CNN) architectures—CNN, ResNet, VGG, and AlexNet was assessed based on their training and validation accuracies. The results are as follows:

CNN Model:

- **Epoch 1:**The model started with a training accuracy of 36.57% and a validation accuracy of 47.78% The training loss was 1.8463, and the validation loss was 1.1197.
- **Epoch 30:**The model achieved a training accuracy of 98.60% and a validation accuracy of 80.46%. The training loss decreased to 0.0406, and the validation loss to 1.9593 The CNN model achieved an impressive training accuracy of 0.98, indicating that it has learned the training data exceptionally well

ResNet:

- **Epoch 1:**The model started with a training accuracy of 39.76% and a validation accuracy of 24.17% The training loss was 1.5195, and the validation loss was 1.5506.
- **Epoch 30:**The model achieved a training accuracy of 34.43% and a validation accuracy of 33.21%. The training loss decreased to 1.5240, and the validation loss to 1.5262.
  The ResNet model exhibited low training and validation accuracies of 0.34 and 0.33, respectively. This performance indicates that the model is struggling to learn from the dataset, which could be due to various factors such as insufficient training epochs, inappropriate learning rate, or issues with the dataset itself. So further proceeded the model by tunnig to fit and did a early stopping to get good accuracy without any overfitting issuse.After doing this process for 20 epoch the training acuuracy shows 92.06% and the validation accuracy of 73.61%. The training loss decreased to 0.2317 , and the validation loss to 1.2053.

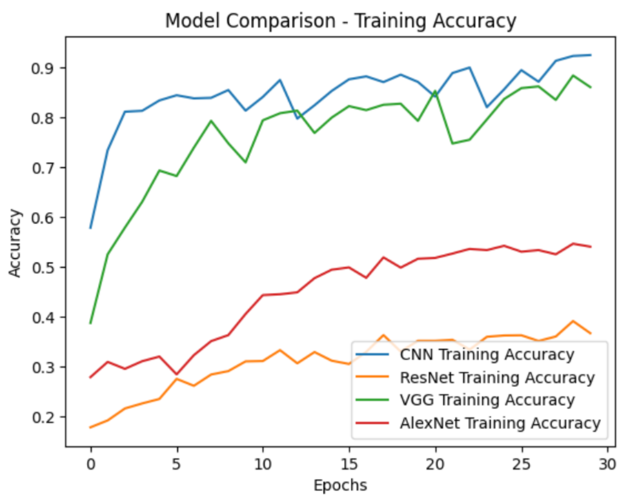
VGG Model:

- **Epoch 1:** The model started with a training accuracy of 25.27% and a validation accuracy of 32.22%. The training loss was 1.7658, and the validation loss was 1.5468.

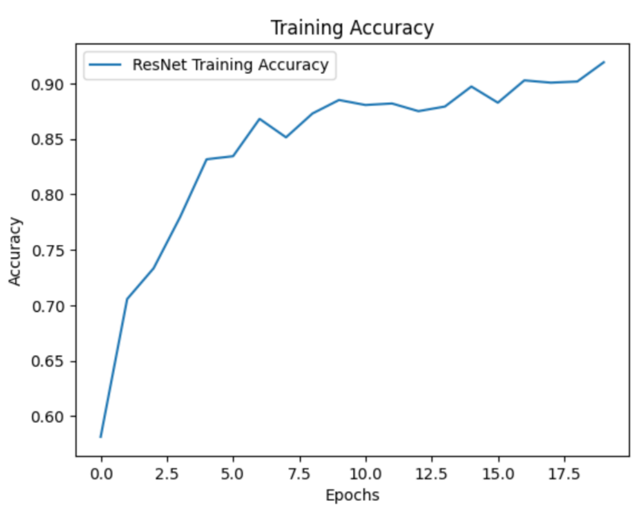
- **Epoch 30:** The model achieved a training accuracy of 82.12% and a validation accuracy of 70.58%. The training loss decreased to 0.3959, and the validation loss to 1.2835 The VGG model achieved a training accuracy of 0.83 and a validation accuracy of 0.70. While the training accuracy is relatively high.

AlexNet Model:

- **Epoch 1:** The model started with a training accuracy of 35.44% and a validation accuracy of 24.31%. The training loss was 1.4386, and the validation loss was 1.7223.
- **Epoch 30:** The model achieved a training accuracy of 68.66% and a validation accuracy of 62.08%. The training loss decreased to 0.7014 , and the validation loss to 1.5511 The AlexNet model recorded a training accuracy of 0.68 and a validation accuracy of 0.62. This performance indicates that the model is learning some features from the training data but is not performing optimally on the validation set.



After fine-tunning the ResNet model



VI. CONCLUSION

The evaluation of the CNN, ResNet, VGG, and AlexNet models reveals significant differences in their performance on the classification task. The CNN model stands out with the highest training accuracy but faces challenges with overfitting. The VGG model shows promise with a balanced performance, while both the ResNet and AlexNet models require further optimization to improve their learning and generalization capabilities. These findings underscore the importance of model selection, hyperparameter tuning, and the implementation of strategies to enhance model performance in deep learning applications. Based on the findings, the VGG model is recommended as the best-performing model for defect classification. It demonstrates a good balance between training and validation performance, indicating its ability to generalize better than the other models evaluated.To further improve the VGG model's performance, consider implementing techniques such as data augmentation, hyperparameter tuning, and additional training epochs. This will help enhance its accuracy and robustness in real-world applications.

VII. REFERENCES

Moganam, P.K., Seelan, D.A.S. (2020). Perceptron neural network-based machine learning approaches for leather defect detection and classification. Instrumentation Mesure Métrologie, Vol. 19, No. 6, pp. 421-429. https://doi.org/10.18280/i2m.190603

M. Praveen Kumara and S. Denis Ashok(2020).A Multi-Level Colour Thresholding Based Segmentation Approach for Improved Identification of the Defective Region in Leather Surfaces. DOI: https://doi.org/10.4186/ej.2020.24.2.101.

Liong, S. T., Zheng, D., Huang, Y. C., & Gan, Y. S. (2020). Leather defect classification and segmentation using deep learning architecture. International Journal of Computer Integrated Manufacturing, 33(10-11), 1105-1117.

Amorim, W. P., Pistori, H., Pereira, M. C., & Jacinto, M. A. C. (2010, August). Attributes reduction applied to leather defects classification. In 2010 23rd SIBGRAPI conference on Graphics, patterns and images (pp. 353-359). IEEE.

Chen, Z., Deng, J., Zhu, Q., Wang, H., & Chen, Y. (2022). A systematic review of machine-vision-based leather surface defect inspection. Electronics, 11(15), 2383.

Aslam, M., Khan, T. M., Naqvi, S. S., Holmes, G., & Naffa, R. (2020). Ensemble convolutional neural networks with knowledge transfer for leather defect classification in industrial settings. IEEE Access, 8,198600-198614.

Chen, Z., Zhu, Q., Zhou, X., Deng, J., & Song, W. (2024). Experimental Study on YOLO-Based Leather Surface Defect Detection. IEEE Access.

Jawahar, M., & Vani, K. (2019). Machine vision inspection system for detection of leather surface defects. Journal of the American Leather Chemists Association, 114(1).

Liong, S. T., Gan, Y. S., Huang, Y. C., Liu, K. H., & Yau, W. C. (2019). Integrated neural network and machine vision approach for leather defect classification. arXiv preprint arXiv:1905.11731.

Khanal, S. R., Silva, J., Magalhães, L., Soares, J., Gonzalez, D. G., Castilla, Y. C., & Ferreira, M. J. (2022). Leather defect detection using semantic segmentation: A hardware platform and software prototype. Procedia Computer Science, 204, 573-580.

Chen, Z., Xu, D., Deng, J., Chen, Y., & Li, C. (2023). Comparative study on deep-learning-based leather surface defect identification. Measurement Science and Technology, 35(1), 015402.

Jawahar, M., Babu, N. C., Vani, K. L. J. A., Anbarasi, L. J., & Geetha, S. (2021). Vision based inspection system for leather surface defect detection using fast convergence particle swarm optimization ensemble classifier approach. Multimedia Tools and Applications, 80, 4203-4235.

Gan, Y. S., Chee, S. S., Huang, Y. C., Liong, S. T., & Yau, W. C. (2021). Automated leather defect inspection using statistical approach on image intensity. Journal of Ambient Intelligence and Humanized Computing, 12(10), 9269-9285.

Chudzik, S. (2020). Quality inspection of natural leather using non-destructive testing technique. Quantitative InfraRed Thermography Journal, 17(4), 249-267.

Iqbal, S., Khan, T. M., Naqvi, S. S., & Holmes, G. (2023). MLR-Net: A multi-layer residual convolutional neural network for leather defect segmentation. Engineering Applications of Artificial Intelligence, 126, 107007.