Classification of defective and non-defective images of leather using GoogleNet

Abstract

Modern leather industries are focused on producing high quality leather products for sustaining the market competitiveness. However, various leather defects are introduced during various stages of manufacturing process such as material handling, tanning and dyeing. Manual inspection of leather surfaces is subjective and inconsistent in nature; hence machine vision systems have been widely adopted for the automated inspection of leather defects. It is necessary to develop suitable image processing algorithms to localize leather defects such as folding marks, growth marks, grain off, loose grain, and pinhole due to the ambiguous texture pattern and tiny nature in the localized regions of the leather.

Introduction

This project explores the use of the GoogleNet model, a powerful convolutional neural network (CNN) architecture, for the classification of defective and non-defective images of leather. By leveraging transfer learning, we adapt the pre-trained GoogleNet model to our specific task, saving time and computational resources.

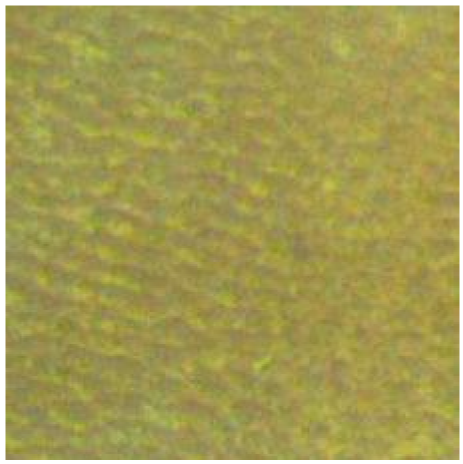
Our objective is to develop a robust and accurate classification model for distinguishing between defective and non-defective leather images. We discuss the methodology, including data collection and preprocessing, model architecture, and training procedure.

The outcomes of this study provide practical solutions for the leather industry, automating defect detection and enhancing quality control measures.

Dataset

The dataset used in this study consists of images of leather pieces categorized into six folders: "Folding marks," "Grain off," "Growth marks," "Loose grains," "Pinhole," and "Not defective." Each folder contains 600 colored images.

The "Folding marks" (1) folder contains images of leather pieces with visible folding marks, which are creases or wrinkles on the surface. These marks can affect the appearance and quality of the leather.



1.(Folding marks)

The "Grain off" (2) folder includes images of leather pieces where the grain layer has been removed or damaged. Grain refers to the natural texture and pattern on the surface of the leather.

2.(Grain off)

Images in the "Growth marks" (3) folder depict leather pieces with visible growth marks, which are irregular patterns caused by the natural growth of the animal's skin. These marks can affect the aesthetics of the leather.

3.(Growth marks)

The "Loose grains" (4) folder contains images of leather pieces with loose or detached grain layers. This can result in an uneven or compromised surface texture.

4.(Loose grains)

Images in the "Pinhole"(5) folder showcase leather pieces with small holes or punctures. These imperfections can impact the durability and visual appeal of the leather.

5.(Pinhole)

Lastly, the "Not defective" (6) folder contains images of leather pieces that are considered to be in a non-defective condition. These images serve as a reference for the model to learn what constitutes a normal, defect-free leather piece.

6.(Not defective)

The dataset provides a diverse range of leather defects and non-defective samples, enabling the development and evaluation of a robust classification model using the GoogleNet architecture.

Methodology

1 Preprocessing

1. Merging Defective Images: The images from the "Folding marks," "Grain off," "Growth marks," "Loose grains," and "Pinhole" folders were combined into a single

dataset. This step ensures that all defective images are treated equally during training.
2. Shuffling: The merged dataset of defective images was shuffled randomly. Shuffling the data helps to reduce any inherent bias in the order of examples present in the dataset and prevents the model from learning patterns related to the order rather than the actual features.
3. Selecting 800 Defective Images: From the shuffled dataset of defective images, 800 images were selected. This ensures a balanced representation of different defect types and avoids an imbalance in the training data.
4. Adding Non-defective Images: 600 non-defective images were added to the dataset. These images represent the "Not defective" category and serve as a reference for the model to learn what constitutes a normal, defect-free leather piece.
5. Labeling: The merged dataset of defective and non-defective images was labeled as follows: 0 for defective images and 1 for non-defective images. This labeling scheme allows the model to differentiate between the two classes during training and prediction.

By merging, shuffling, and balancing the dataset, biasness in the data is minimized, ensuring that the model learns to classify defective and non-defective leather images accurately.

2. Model Training

1. Loading the Pre-trained GoogleNet Model: The pre-trained GoogleNet model, also known as Inception-v3, was loaded. This model has learned important features from a large dataset and can be used as a base for our classification task.
2. Freezing Layers: All layers of the pre-trained GoogleNet model, except for the top layers, were frozen. By freezing these layers, we prevent their weights from being updated during training, preserving the learned features.
3. Adding GlobalAveragePooling2D Layer: A GlobalAveragePooling2D layer was added after the frozen layers. This layer reduces the spatial dimensions of the feature maps and extracts the most important features from each channel.
4. Adding Dense Layer: A Dense layer with a sigmoid activation function was added on top of the GlobalAveragePooling2D layer. This layer serves as the final output layer.
5. Model Compilation and Fitting: Using Adam optimizer and binary cross entropy loss , model is fitted on 10 epochs.

3. Prediction

Here we have applied the model's prediction function sigmoid to the prepared input data. This will generate the predicted outputs based on the learned patterns and parameters of the model.

4.Thresholding

A threshold of 0.5 is used to divide the predictions into two categories. Here values above 0.5 are classified as one category (e.g., non-defective) and values below 0.5 are classified as the other category (e.g., defective).

Performance Analysis

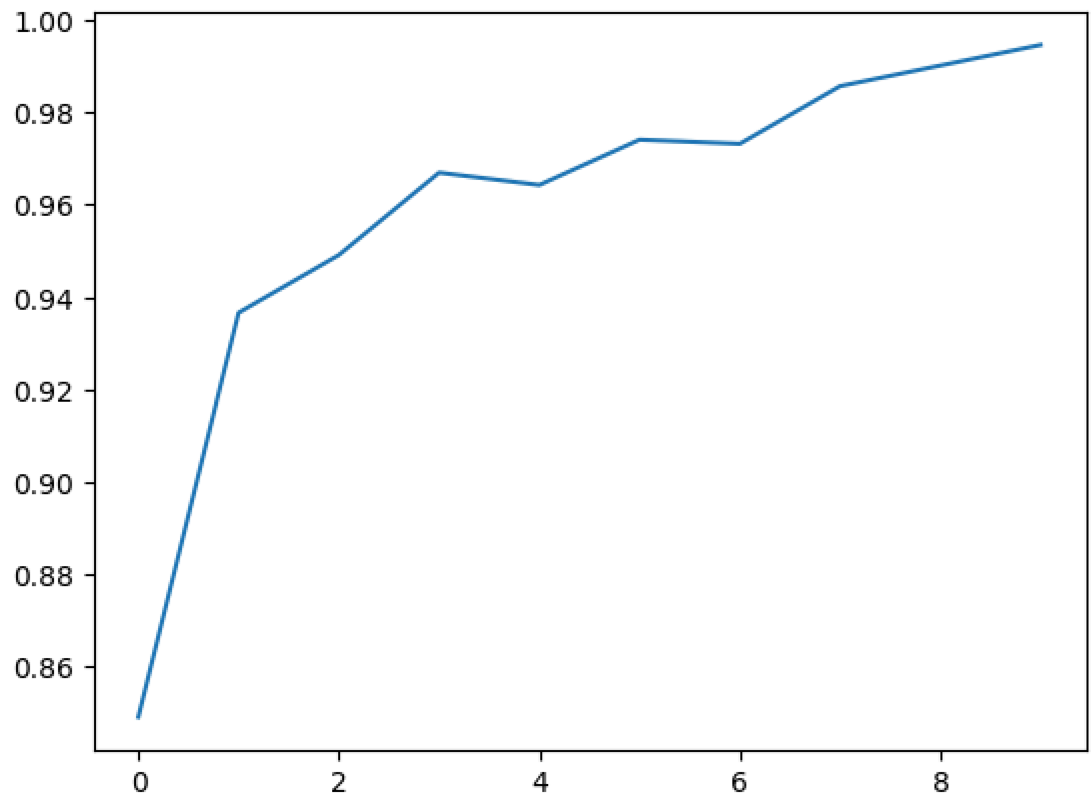
In the evaluation of our model's performance on the test set, we utilized key metrics such as accuracy, precision, and recall. The following results were obtained:

- Accuracy: 0.9857
- Precision: 0.9655
- Recall: 1.0

The accuracy score of 0.9857 indicates that our model performs exceptionally well in classifying both defective and non-defective leather images, with a high rate of correct predictions overall. This demonstrates the effectiveness of our model in accurately categorizing the images.

Precision, with a score of 0.9655, reflects the proportion of correctly predicted positive samples (non-defective) out of all samples predicted as positive. This indicates that when our model predicts an image as non-defective, it is correct approximately 96.55% of the time. Our model exhibits a high level of precision, ensuring a low rate of false positives.

Recall, also known as sensitivity, measures the proportion of correctly predicted positive samples out of all actual positive samples. In our case, the recall score of 1.0 signifies that our model is able to identify all the non-defective images correctly. This indicates a low rate of false negatives and highlights the model's ability to capture all the positives in the dataset.



History of accuracy in each epoch

Insights Gained

1. Accuracy: The model achieved an accuracy of 0.9857, indicating that it is able to correctly classify the leather images with a high level of accuracy.
2. Precision: The precision score of 0.9655 suggests that when the model predicts an image as non-defective, it is correct approximately 96.55% of the time. This indicates a low rate of false positives.
3. Recall: The recall score of 1.0 indicates that the model is able to identify all the non-defective images correctly. This suggests a low rate of false negatives.

These insights demonstrate that the model performs exceptionally well in classifying both defective and non-defective leather images. It exhibits high accuracy, precision, and recall, which are crucial metrics for evaluating the model's performance.

Conclusion

This model demonstrates promising results for the task of classifying leather images as defective or non-defective.

During the training phase, the model achieved an impressive accuracy of 98.57%, indicating its ability to accurately classify leather images. The precision score of 96.55% suggests a low rate of false positives, while the recall score of 100% indicates a low rate of false negatives. These metrics highlight the model's effectiveness in correctly classifying both defective and non-defective leather images.

To make predictions on new, unseen data, a methodology was outlined. This involved preprocessing the input data, loading the trained model, preparing the input for prediction, applying the model's prediction function, performing post-processing, and interpreting the predicted outputs.

Overall, the developed model provides a solid foundation for accurately classifying leather images. Its performance demonstrates potential value in the leather industry or related domains.