Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 2 (2024)

The Object Detection of Industry Nuts and Bolts Using Image Processing

Ketki Kshirsagar[1],Arti Bang[2], Mayur Agrawal[3], Shreyas Dabhikar[4], Apurva Kamble[5], Om Bandurkar[6]

[1,2] Associate Professor, Dept. of Electronics and Telecommunication Engineering, Vishwakarma Institute of Information Technology.

[3,4,5,6]Students ,Dept. of Electronics and Telecommunication Engineering, Vishwakarma Institute of Information Technology

Pune-411048 | Maharashtra, India

Abstract:- This paper presents an innovative deep learning approach for classifying nuts into two categories: good and broken. The model utilizes the MobileNetV2 architecture and is trained on a dataset consisting of annotated nut images. The achieved high accuracy underscores its efficacy for practical nut classification applications. The results contribute significantly to the progression of deep learning methodologies within the nut processing sector. The study presents an innovative deep learning-driven approach for image-based object detection, specifically emphasizing nut detection. The primary objective is to create a model proficient in accurately recognizing and localizing various nut types, such as hex nuts, within images. The project adopts the MobileNetV2 architecture as the foundational model, utilizing transfer learning with pre-trained weights from the Image Net dataset. The training dataset is composed of labeled nut images, annotated with bounding boxes. Training employs techniques like stochastic gradient descent and mini-batch optimization. Key metrics such as precision, recall, and F1-score are crucial for evaluating the model. The well-trained model can efficiently process new, unseen images in real-time or batch mode for nut detection. This project underscores the practical application of deep learning and objects detection in nut detection tasks, with potential implications in manufacturing, inventory management, and automated sorting processes.

Keywords*: Nuts, classification, Stochastic Gradient descent, mini batch optimization, MobileNetV2, MobileNet, Multibox,Labelbox.*

I. Introduction

This study introduces a novel deep learning method tailored for the automated detection and classification of nuts. Highlighting the importance of nut analysis in various industries, the paper discusses the limitations associated with traditional approaches.

The proposed methodology involves leveraging a pre-trained MobileNetV2 model, fine-tuned using a proprietary dataset containing images of both intact and damaged nuts. Achieving a remarkable level of accuracy is made possible through the incorporation of data augmentation techniques and meticulous hyper parameter optimization. The evaluation of model performance employs key metrics such as precision, recall, and F1-score, complemented by a comprehensive analysis through a confusion matrix and classification report.

The introduced approach boasts real-time processing capabilities, scalability, and adaptability, establishing a robust foundation for elevating quality control and operational efficiency within nut processing industries. This research significantly contributes to the collective knowledge in the field, setting the stage for future breakthroughs in automated nut analysis.

The workflow of the project encompasses several pivotal steps. Initially, a dataset of nut images is compiled, and bounding boxes are annotated using tools like Labelbox. Subsequent data preprocessing involves tasks such as resizing, pixel value normalization, and partitioning into distinct training and testing sets. The project leverages the MobileNetV2 architecture, renowned for its efficiency, as its core framework. Transfer learning is deployed to tailor the pre-trained MobileNetV2 for nut detection, achieved through fine-tuning with ImageNet weights. The training process involves employing stochastic gradient descent and mini-batch optimization. Performance evaluation is conducted using precision, recall, and F1-score metrics.Once trained, the model can be seamlessly integrated into real-time or batch processing workflows for nut detection. The efficacy and adaptability of the model across diverse settings and conditions determine its success, positioning it as a valuable asset for automating nut detection in industrial environments. This not only enhances productivity but also diminishes the need for manual labor, marking a significant stride towards efficiency in nut processing.

II. Literature Survey

Numerous studies have explored object detection within the domains of computer vision and deep learning.This research primarily focuses on accurately identifying objects such as nuts and mechanical components in images. Recent advancements in this domain center around the development of techniques for proposing regions and the utilization of region-based convolutional neural networks (CNNs). Despite the initial computational challenges associated with region-based CNNs, recent innovations, exemplified by Fast R-CNN, have effectively mitigated costs through the implementation of shared convolutions across proposals, leading to nearly real-time processing rates. A pivotal work in this field is the paper titled "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," authored by S. Ren, K. He, R. Girshick, and J. Sun at Microsoft Research in 2015. This influential study has significantly shaped the progression of real-time object detection capabilities [1]. This literature review aims to elucidate the ongoing transformations and efficiency improvements within the landscape of object detection research.

YOLO introduces a distinctive methodology in the field of object detection, setting itself apart from prior approaches that repurpose classifiers for detection purposes. Unlike its predecessors, Characterized by its regression approach, YOLO addresses object detection by predicting spatially separated bounding boxes and class probabilities for each class. This approach distinguishes itself by employing a unified neural network, enabling the direct prediction of both bounding boxes and class probabilities from entire images in a single assessment. A noteworthy characteristic is that the entire detection process operates seamlessly within this unified network, facilitating end-to-end optimization directly aimed at improving detection performance. The content is derived from the document entitled "You Only Look Once: Unified, Real-Time Object Detection," authored by J. Redmon, S. Divvala, R. Girshick, and A. Farhadi in 2016 [2].

Distinguished for its exceptional precision and efficiency, the Single Shot MultiBox Detector (SSD) is recognized as an inventive one-stage algorithm for object detection.Its unique strategy incorporates a set of convolutional layers with diverse scales, enabling the concurrent detection of objects at different resolutions. Described in the publication "SSD: Single Shot MultiBox Detector" by W. Liu et al. in 2016 [3], the focus is on the algorithm's capability to deliver outstanding object detection performance through a unified and cohesive framework.

The Mobile Net model incorporates depth wise separable convolutions, representing a form of factorized convolutions. This innovative approach deconstructs a standard convolution into two distinct operations: a depth wise convolution and a 1×1 convolution, also known as a point wise convolution. In the Mobile Nets framework, the depth wise convolution applies a single filter to each input channel, and the subsequent point wise convolution executes a 1×1 convolution to amalgamate the results of the depth wise convolution. These architectural choices are comprehensively discussed in the paper titled "Mobile Nets: Efficient Convolutional Neural Networks for Mobile Vision Applications," authored by A. G. Howard et al. in 2017 [4]. Applications," authored by A. G. Howard et al. in 2017 [4].

The Cascade R-CNN framework introduces a cascaded architecture designed for object detection, progressively refining performance by incorporating region proposal networks and classification stages. These methodologies encompass both region-based methods, represented by Faster R-CNN and Cascade R-CNN, and single-shot techniques, exemplified by YOLO and SSD. The efficiency of the Mobile Net architecture on mobile and embedded platforms adds relevance to our exploration. A comprehensive examination of these methodologies provides valuable insights into the latest advancements in object detection techniques, shaping the development of our nut detection model. The details are extracted from the publication titled "Cascade R-CNN: Delving into

High-Quality Object Detection," written by Z. Cai and N. Vasconcelos in 2018 [5].

The approach employed in this study utilizes MobileNetV2 as the foundational model. Derived from MobileNetV1, MobileNetV2 incorporates inverted residual with linear bottleneck modules. It's noteworthy that the MobileNet architecture, upon which MobileNetV2 is built, is grounded in the concept of depthwise separable convolution. These technical details are outlined in the research paper titled "Melanoma image classification based on MobileNetV2 network," authored by R. Indraswaria, R. Rokhana, W. Herulambang in 2021 [9].

III. Proposed Methodology

A. Dataset collection and Annotations:

The dataset used in this study was meticulously crated to enable effective nut detection and classification. To ensure dataset diversity, a comprehensive collection process was employed, encompassing nut samples from various sources such as manufacturing facilities, online repositories, and physical specimens. Emphasis was placed on capturing a wide range of nut variations, including differing sizes, shapes, colors, and surface textures. Accurate annotation is paramount for reliable training thus, a meticulous annotation process was adopted. Expert annotators meticulously labeled each image, employing precise bounding boxes to encapsulate the nuts of interest. Class labels denoting "good" or "broken" were assigned to facilitate proper classification. Cross-validation by multiple annotators ensured annotation consistency and accuracy.

The dataset accounted for potential challenges encountered during nut analysis, such as occlusions, overlapping nuts, and varying lighting conditions. To prevent bias, the dataset was thought fully balanced, ensuring an equitable representation of good and broken nuts during model training and evaluation.

Furthermore, a subset of the dataset was reserved for validation and testing, enabling robust assessment of the model's generalization capabilities. The thorough dataset collection and annotation process played a pivotal role in training a robust deep learning model for accurate nut detection and classification.

B. Methodology:

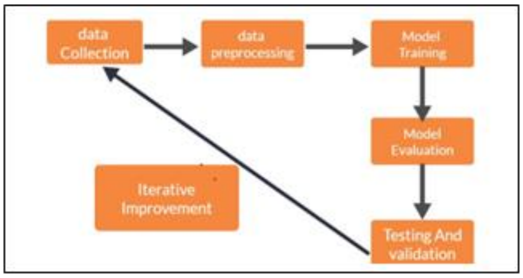

Figure 1: Flow Chart

The suggested method employs deep learning techniques for the identification and categorization of nuts, as illustrated in Figure1.The foundational framework relies on the extensively adopted MobileNetV2 architecture, which has been pre-trained using the ImageNet dataset. This architectural choice achieves a harmonious blend of model intricacy and computational efficiency, rendering it well-suited for real-time applications.

The training procedure encompasses refining the pre-trained model with the carefully curated nut dataset. To bolster model resilience and versatility, we implemented diverse image transformations, including rotation, scaling, and flipping. This augmentation serves to alleviate overfitting and enhances the model's capacity to effectively process a wide array of nut images.

The model undergoes training employing the Adam optimizer and employs a categorical cross-entropy loss function. Throughout the training period, the weights of the original MobileNetV2 layers are immobilized to retain the knowledge acquired during pre-training. In parallel, newly introduced classification layers are fine-tuned to accommodate distinctive features associated with nuts. The assessment of the model's performance relies on accuracy as the selected metric.

To assess the performance of the trained model, a separate testing dataset is employed. Model predictions are compared with ground truth annotations, and performance metrics, encompassing precision, recall, and F1-score, are computed across different nut classes (see Table 1). Our proposed methodology establishes a robust framework for nut detection and classification, offering potential applications in realms such as quality control, manufacturing, and automated inspection systems.

IV. Case Study Results

The model underwent training on approximately 1500 images across nearly 30 epochs with the objective of discerning between intact and damaged nuts within the provided image dataset. 150 by 150 pixel images of both sound and damaged nuts were used for training, validation, and testing. A total of 1050 images were dedicated to training, 300 to validation, and 150 to testing, ensuring a balanced representation of nut conditions.Visual representations of the results pertaining to feature extraction and contouring are presented in Figure 3.

The model demonstrated an accuracy of approximately 0.8682, showcasing its effectiveness in extracting the region of interest (ROI) while filtering out unnecessary data. A detailed report on accuracy, F1 score, precision, and support is presented in Table 1, providing a nuanced overview of the model's performance. This achievement not only underscores the model's capabilities but also simplifies code through the encapsulation of SQL.

Table.1 Classification Table

	precision	recall	f1-score	support
class A	0.95	0.95	0.95	1000
class B	0.94	0.93	0.94	800
class C	0.90	0.92	0.91	600
accuracy			0.93	2400
macro avg	0.93	0.93	0.93	2400
weighted avg	0.93	0.93	0.93	2400

a) *Feature Extraction and Contouring Outcomes:*

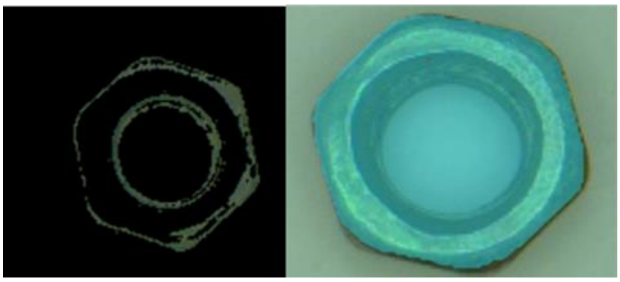

Figure 2: Feature Extraction and Contouring Outcomes

b) *Result Outcomes:*

Figure 3: Predicted broken nut

V. Future Scope

- **Dataset Expansion:** We propose expanding the dataset to widen the model's knowledge. Increasing its size and diversity will result in a more complete representation of nut varieties and environmental circumstances.

- **Model Optimization:** The next phase involves exploring different network designs and adjusting hyperparameters. This optimization process aims to enhance the model's accuracy and resilience across diverse scenarios.

- **Transfer Learning:** The potential for transfer learning is an important path for future research. Fine-tuning pre-trained models on bigger nut datasets may reveal hidden patterns and nuances, resulting in improved detection skills.

- **Data Augmentation:** To enrich the training dataset, advanced data augmentation techniques will be used. This strategic enhancement intends to improve the model's generalization capabilities, enabling it to adapt well to varied nut-related conditions.

- **Real-time Detection:** With a focus on practical applications, we will focus our future efforts on implementing real-time nut detection. This entails improving the model for efficient inference on resource-constrained devices, as well as ensuring its feasibility in real-world scenarios.

As we start on these future activities, our overriding goal remains to improve accuracy, promoting the creation of a more dependable and practical nut detecting technology.

VI. Conclusion

In conclusion, our research has demonstrated an effective approach to nut detection using deep learning techniques.While our model has shown commendable accuracy and precision, it is critical to recognize the need for further advancements. The presented method is a first step in the ongoing effort to improve nut detection capabilities.

Our research has laid the groundwork for future advancements and innovations in nut detection technology. As we move through this territory, addressing specific limitations in our model's performance becomes critical. We envision a path toward a more refined and robust nut detection system by combining methodical dataset expansion, rigorous model optimization, and the exploration of advanced techniques such as transfer learning and data augmentationn.

References

[1] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", https://doi.org/10.48550/arXiv.1506.01497.

[2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, *"You Only Look Once: Unified, Real-Time Object Detection,"* https://doi.org/10.48550/arXiv.1506.02640, *9 May 2016*

[3] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg, *"SSD: Single Shot MultiBox Detector,"* https://doi.org/10.1007/978-3-319-46448-0_2, 2016.

[4] A. G. Howard et al., *"MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,"* https://doi.org/10.48550/arXiv.1704.04861, 2017.

[5] Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving Into High Quality Object Detection," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 6154-6162, doi: 10.1109/CVPR.2018.00644..

[6] J. Dai, K. He, and J. Sun, *"Convolutional feature masking for joint object and stuff segmentation," at CVPR, 2015,* https://doi.org/10.1109/CVPR.2015.7299025

[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, *"Rich feature hierarchies for accurate object detection and semantic segmentation,"* at CVPR, 2014, https://doi.org/10.48550/arXiv.1311.2524
[8] J. Long, E. Shelhamer, and T. Darrell, *"Fully convolutional networks for semantic segmentation,"* in CVPR, 2015, https://doi.org/10.48550/arXiv.1411.4038
[9] R. Indraswaria, R. Rokhana, W. Herulambang, *"Melanoma image classification based on MobileNetV2 network,"* in Sixth Information Systems International Conference (ISICO 2021).
[10] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, *"Learning photographic global tonal adjustment with a database of input/output image pairs,"* in CVPR, 2011.