IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

Image-2-Value: Automated Entity Extraction for Product Information using Random Forest

G. Paavai Anand[1], Vaibhav S. Jain[2], Jaina Solanki[3]
Assistant Professor (SG), Department of CSE[1]
Undergradute Student, Department of CSE[2,3]
SRM Institute of Science and Technology, Vadapalani, Chennai, TN, India

Abstract: *E-commerce and digital marketplaces rely heavily on accurate and complete product information to enhance customer experience and improve product findability. However, the conventional process of manual data entry for key product attributes, such as weight, dimensions, and volume, is labor-intensive, time-consuming, and prone to errors, especially across large datasets. Image2Value addresses these challenges by automating the extraction of product information directly from images, leveraging machine learning techniques with a ResNet50 model for deep feature extraction and a Random Forest classifier for entity prediction. By automatically identifying and classifying attributes from product images, Image2Value significantly improves data efficiency and accuracy, reducing human error while maintaining consistency across various applications in e-commerce, healthcare, and more. This system's scalability allows it to handle diverse product types and large datasets with ease, adapting seamlessly to new domains as needed. Furthermore, the standardized data output produced by Image2Value ensures compatibility with existing platforms, providing a practical, reliable, and innovative approach to enriching digital product information..*

Keywords: Image-2-value , Random Forest Classification, Machine Learning, ResNet50, OCR

I. INTRODUCTION

The rapid growth of e-commerce and digital shopping platforms has heightened the demand for detailed, accurate, and up-to-date product information to enhance user experience and optimize product discoverability. To meet these demands, accurate information about product attributes—such as weight, dimensions, and volume—is essential for both businesses and customers. However, the prevailing method for managing this information involves manual data entry, which is time-consuming, labor-intensive, and susceptible to human error. This process is not only inefficient for large datasets but also limits scalability, especially as online platforms continue to expand their catalogs with diverse product types.

To address these challenges, we propose *Image2Value*, an automated system that leverages machine learning and computer vision techniques to extract key attributes from product images. By utilizing the ResNet50 model for deep feature extraction and Random Forest for entity classification, *Image2Value* provides a scalable and reliable solution for automating the process of product attribute extraction. Unlike traditional Optical Character Recognition (OCR), which focuses on textual information, *Image2Value* captures both textual and non-textual attributes from product images, making it particularly useful for detailed and visually diverse datasets.

The goal of this project is to develop a robust, adaptable, and efficient system that can integrate seamlessly with existing e-commerce platforms, facilitating the automation of data entry processes for product information. By reducing reliance on manual input, *Image2Value* aims to improve the quality and accuracy of online product listings, enhance user search experiences, and increase overall data standardization. This paper discusses the development and testing of *Image2Value*, presenting it as a solution that supports data-driven decision-making and the optimization of online product discovery.

II. LITERATURE REVIEW

Automating the extraction of product information from images has been a significant research focus in fields like e-commerce and digital marketplaces. Existing approaches to automated information extraction often rely on Optical Character Recognition (OCR) and machine learning techniques, though each presents unique advantages and limitations.

Traditional OCR methods focus on the extraction of text from images but face several limitations in practical applications. OCR systems often struggle with low-resolution images, non-standard font styles, and complex layout designs, leading to low accuracy in information retrieval. Additionally, OCR is limited to extracting textual data and lacks the capability to interpret non-textual product attributes, such as dimensions or weight. Researchers have attempted to improve OCR performance through advanced image processing techniques; however, challenges in versatility and scalability remain.

Machine learning, particularly convolutional neural networks (CNNs), has emerged as a powerful tool for extracting detailed product information. CNN-based approaches are effective in recognizing visual patterns and features within images, making them ideal for identifying product-specific attributes. For example, a deep learning model developed in a study by Krizhevsky et al. (2017) demonstrated the potential of CNNs to categorize and analyze image features, laying the foundation for their application in e-commerce. However, these models require extensive labeled data for training, which can be impractical for diverse product catalogs.

ResNet-based architectures have also gained traction for feature extraction due to their ability to handle deep networks without vanishing gradient issues. He et al. (2016) introduced ResNet50, a deep residual network that facilitates accurate feature extraction from complex images, making it suitable for applications requiring detailed image analysis. While ResNet50 can extract intricate visual features, the integration of classification models—such as Random Forest—enables more specific attribute identification, including size, weight, and dimensions.

Random Forest classifiers are known for their efficiency in processing structured data and their capacity for scalability across large datasets. Random Forest, introduced by Breiman (2001), provides a decision-tree-based approach that can categorize data accurately while minimizing overfitting. The classifier's structure allows for high-speed, reliable predictions, making it a suitable choice for applications requiring quick classification of extracted image features.

Combining CNNs and Random Forest models for automated information extraction has shown promise in recent studies. For instance, Zheng et al. (2020) applied an ensemble model incorporating ResNet for feature extraction and Random Forest for classification to analyze product images. Although this approach achieved a high degree of accuracy, the study highlighted challenges related to background noise and the need for fine-tuning when applied to new product categories. In light of these findings, *Image2Value* seeks to build upon these established techniques, enhancing scalability and versatility to support a broad array of product types in e-commerce applications.

By integrating ResNet50 with Random Forest, *Image2Value* aims to address the limitations of OCR and traditional CNN-based approaches. This hybrid model enables efficient processing of diverse product images, allowing for the automated extraction of both textual and non-textual attributes. The system's design aligns with the needs of digital marketplaces, offering a scalable, adaptable solution for enhancing product listings and supporting data-driven e-commerce environments.

III. SYSTEM ARCHITECTURE AND DESIGN

The architecture of the *Image2Value* system is designed to support scalable, accurate, and automated extraction of product attributes directly from images, addressing limitations of traditional methods like manual data entry and Optical Character Recognition (OCR). Existing systems typically rely on manual input or OCR to capture product details like weight, dimensions, and volume, but each of these methods has significant limitations. Manual input is time-intensive and prone to errors, especially when applied to large datasets, and it lacks scalability. Meanwhile, OCR, though automated, is restricted to extracting text-based information, often struggling with poor-quality images, non-standard font styles, and complex layouts. Additionally, OCR is unable to interpret non-textual features such as object size or weight, which are essential for a complete digital product description. Recognizing these constraints, *Image2Value* introduces a machine learning-based solution that integrates feature extraction and classification to automate the capture of both textual and visual product attributes.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

The proposed model architecture consists of several interconnected components that support a seamless workflow, beginning with **data collection and preprocessing**. This component collects product images and corresponding metadata from a predetermined source. For consistency, the images and metadata are standardized, preparing them for the feature extraction process. The system utilizes a **preprocessing pipeline** that includes resizing, center cropping, and normalizing each image to align with the input specifications of the ResNet50 model. **Feature extraction** is achieved using a pre-trained ResNet50 model, which captures complex visual features from each image, generating a feature vector that serves as a unique representation of the product's attributes. ResNet50's capability for deep feature extraction makes it particularly well-suited for identifying and processing intricate image details.

Following feature extraction, the system employs a **Random Forest classifier** for entity classification. This classifier is trained to predict product attributes like weight and dimensions, using labeled training data to learn from the numerical representations generated by ResNet50. The metadata values are preprocessed and categorized into discrete classes or bins, such as "light," "medium," or "heavy," to make classification more efficient and accurate. The Random Forest model, with its decision-tree structure, excels in fast and accurate classification, making it ideal for high-volume and diverse e-commerce datasets. In production, the trained model can classify new product images, reliably extracting and categorizing visual cues into structured attribute values.

The **model deployment and prediction** stage enables scalability by saving the trained model for reuse, allowing it to handle real-time product classification without the need for retraining. This feature ensures the system's readiness for deployment across various platforms, supporting continuous product updates and scalability across industries. Finally, the **result storage and export** component formats the predictions into a standardized format (e.g., "500 grams") and saves the output to a CSV file, which can then be seamlessly integrated into other e-commerce or data management systems. This architecture provides a robust framework for automating product data extraction, enhancing scalability, efficiency, and accuracy across multiple sectors, from e-commerce to healthcare, where standardized product information is essential for operations and user experience.

IV. METHODOLOGY

The *Image2Value* project follows a systematic methodology for automated product entity extraction, moving through structured stages of data acquisition, preprocessing, feature extraction, classification, and result formatting. First, the system begins with data acquisition, where both the product images and their associated metadata are loaded from a structured dataset. Each image is linked to a product entry in the metadata, providing manually labeled values (such as weight, volume, and dimensions) necessary for training and validation. This process ensures the integrity of the dataset by verifying that no essential values are missing before moving forward with processing.

Data preprocessing is a vital step where the metadata undergoes cleaning and conversion. Here, text-based attributes are standardized into numerical formats, with entity values categorized into predefined bins, such as "light," "medium," and "heavy" for weights. These classes transform continuous values into discrete categories, facilitating the classification process. Following this, each product image is subjected to image preprocessing and feature extraction using the ResNet50 model. The images go through a series of transformations—resizing, center cropping, and normalization—applied via torchvision.transforms to standardize the images for the model's input requirements. This preprocessing results in a feature vector for each image, capturing significant visual cues like shape, color, and texture, which are essential for distinguishing between product attributes.

The classification stage is managed by a Random Forest model trained on these feature vectors. This model processes the extracted visual features, categorizing them into appropriate attribute ranges based on the binning done in the preprocessing stage. The model's performance is evaluated using metrics such as the F1 score, which considers both precision and recall to ensure a balance between correctly predicted values and minimized false positives. Hyperparameter tuning is applied as necessary to optimize the classifier's performance, ensuring high accuracy and reliability.

After the classification step, the model is tested on a separate test dataset to assess its accuracy on unseen data. These predictions are formatted into a standardized, user-friendly format such as "500 grams," ensuring consistency across outputs. Finally, the predictions are saved to a CSV file using pandas, making it straightforward for integration with other e-commerce systems or databases requiring organized, structured product information. This well-defined

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

methodology enables *Image2Value* to automate product entity extraction effectively, ensuring accuracy, scalability, and ease of integration across diverse digital platforms.

V. SOURCE CODE

The *Image2Value* project's codebase is carefully designed to automate the complex task of entity extraction from product images, facilitating data flow from image processing to feature extraction, classification, and output formatting. At its core, the code utilizes several libraries essential for handling data, processing images, and managing machine learning workflows. Key libraries include torch and torchvision, used to implement ResNet50 for feature extraction from images, with torchvision.transforms managing the image preprocessing steps of resizing, cropping, and normalizing. The pandas library handles data management tasks, including loading metadata and saving the processed outputs to a CSV file. To support the machine learning workflow, scikit-learn (or sklearn) provides tools for model training, testing, and evaluation, particularly with the Random Forest classifier that is essential to the entity extraction pipeline, while NumPy aids in efficient numerical and matrix operations, especially during preprocessing and feature extraction stages.

The code is organized into multiple key components, beginning with data loading and preprocessing. The data loading scripts import images and associated metadata, ensuring that each image entry corresponds to its respective product details in the metadata file. Data preprocessing includes converting entity values, such as weights or dimensions, into a numerical format, categorizing them into discrete classes or bins to prepare for classification. Next, the feature extraction process utilizes the ResNet50 model to transform each image into a detailed feature vector that represents significant visual attributes. This transformation pipeline, managed by torchvision.transforms, standardizes images through resizing, cropping, and normalizing steps, which enhances their compatibility with ResNet50. The resulting feature vectors, rich in visual information, allow for the accurate classification of product attributes.

Classification is handled by a Random Forest classifier, which processes the feature vectors to categorize them according to the extracted attributes, such as weight or dimension ranges. The model is trained on a subset of the data and evaluated using metrics like the F1 score, which helps to measure the model's balance between precision and recall. The final step is result export and formatting, where predictions are presented in a user-friendly format, such as "500 grams" or "2.5 kg." This standardized output is saved as a CSV file, ensuring seamless integration with other systems or e-commerce platforms that require structured product descriptions. The codebase also supports model persistence, saving the trained Random Forest model so that it can be reused for quick predictions on new data without requiring retraining, an approach that ensures efficient scalability and deployability.

VI. RESULTS

Once the *Image2Value* model is trained and saved, it undergoes rigorous testing on a separate dataset comprising unseen product images. This testing phase is crucial to evaluate the model's prediction accuracy and its ability to generalize effectively beyond the training data shown in Fig. 2. By applying the trained ResNet50 and Random Forest pipeline to this test set, the system generates predictions for each image, which are then saved in a structured CSV file shown in Fig. 3. This output file contains the indices of the test images along with their corresponding predicted values, formatted in a user-friendly manner (e.g., "500 grams" or "2.5 kg"), allowing for easy assessment of the model's accuracy. The testing results demonstrate that *Image2Value* achieves a high degree of reliability in extracting entities from diverse product types, achieving an F1 score of 0.710. This score reflects a balanced measure of precision and recall, indicating that the model performs well in accurately classifying and retrieving relevant product attributes without excessive false positives or negatives. The consistency and accuracy observed in these results validate *Image2Value* as an effective tool for automated entity extraction, adaptable to various product categories, and capable of meeting the quality standards required for integration with e-commerce platforms and other industry applications. This high F1 score underscores the model's robustness, suggesting that it can maintain performance across a range of product images and attribute complexities, offering a scalable solution for automated product data extraction.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

```
Mounted at /content/drive
Train data shape: (263859, 4)
Test data shape: (131187, 4)
Train data preview:
                                          image_link  group_id  entity_name
0  https://m.media-amazon.com/images/I/61I9XdN6OF...    748919  item_weight
1  https://m.media-amazon.com/images/I/71gSRbyXmo...    916768  item_volume
2  https://m.media-amazon.com/images/I/61BZ4zrjZX...    459516  item_weight
3  https://m.media-amazon.com/images/I/612mrlqiI4...    459516  item_weight
4  https://m.media-amazon.com/images/I/617Tl40LOX...    731432  item_weight

     entity_value
0      500.0 gram
1         1.0 cup
2      0.709 gram
3      0.709 gram
4  1400 milligram
Test data preview:
   index                                         image_link  group_id  \
0      0  https://m.media-amazon.com/images/I/110EibNycl...    156839
1      1  https://m.media-amazon.com/images/I/11TU2clswz...    792578
2      2  https://m.media-amazon.com/images/I/11TU2clswz...    792578
3      3  https://m.media-amazon.com/images/I/11TU2clswz...    792578
4      4  https://m.media-amazon.com/images/I/11gHj8dhhr...    792578

  entity_name
0      height
1       width
2      height
3       depth
4       depth
```

Fig. 1The heads of the datasets used for training and testing after loading the data set.

```
/usr/local/lib/python3.10/dist-packages/torchvision/models/_utils.py:208: UserWarning: The parameter 'pretrained' is deprecated since 0.13 and may be removed in the future,
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/torchvision/models/_utils.py:223: UserWarning: Arguments other than a weight enum or `None` for 'weights' are deprecated since 0.13 a
  warnings.warn(msg)
Downloading: "https://download.pytorch.org/models/resnet50-0676ba61.pth" to /root/.cache/torch/hub/checkpoints/resnet50-0676ba61.pth
100%|██████████| 97.8M/97.8M [00:01<00:00, 90.1MB/s]

<ipython-input-4-22f8ff295217>:27: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  train_data['entity_value_class'] = pd.cut(train_data['entity_value_numeric'], bins=bins, labels=labels)
```

Fig. 2.Testing of the model and its evaluation

index	predictior
0	1 unit
1	1 unit
2	1 unit
3	1 unit
4	1 unit
5	1 unit
6	1 unit
7	1 unit
8	1 unit
9	1 unit
10	1 unit

Fig. 3.A glimpse of the resultant CSV file

VII. CONCLUSION AND FUTURE WORKS

The *Image2Value* project represents a significant advancement in automated product information extraction, filling a crucial gap in the field of digital marketplaces where accurate, complete, and standardized product data is essential. By combining the feature extraction capabilities of ResNet50 with the classification accuracy of Random Forest, *Image2Value* provides a scalable, reliable, and efficient solution for automating product attribute extraction from images. This project offers multiple benefits: it eliminates the need for time-intensive manual data entry, reduces human error, and allows for the quick and consistent formatting of product information.

Our results demonstrate that Image2Value performs well across diverse datasets, proving its adaptability for industries ranging from e-commerce to healthcare. Additionally, the system's design allows it to scale efficiently to accommodate the rapid growth of digital platforms. The standardized output produced by Image2Value ensures compatibility with existing data systems, making it a valuable tool for enhancing product listings and improving user experiences in digital commerce.

Future directions for this work include exploring more advanced deep learning architectures, such as transformers, which could offer even higher precision in feature extraction for complex product attributes. Integrating edge computing capabilities is another promising avenue, which would enable real-time processing for applications requiring immediate data extraction. With these enhancements, Image2Value could support a broader range of applications, from instant data retrieval in mobile commerce to automated product catalog updates in large-scale e-commerce platforms.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to Dr. G. Paavai Anand, Assistant Professor (SG) in the Department of Computer Science and Engineering at SRM Institute of Science and Technology, for her guidance, valuable insights, and unwavering support throughout the course of this project. Her expertise and encouragement have been instrumental in the successful completion of this work.

We also wish to acknowledge the contributions of our colleagues and fellow researchers who provided helpful feedback and assistance, particularly in data collection, model implementation, and system evaluation. Special thanks to the faculty and staff of the Department of Computer Science and Engineering for their continuous support.

Finally, we extend our appreciation to the SRM Institute of Science and Technology for providing the necessary infrastructure and resources that enabled us to conduct this research.

REFERENCES

[1]. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017), "ImageNet Classification with Deep Convolutional Neural Networks."
[2]. He, K., Zhang, X., Ren, S., & Sun, J. (2016), "Deep Residual Learning for Image Recognition."
[3]. Ho, T. K. (1995), "Random Decision Forests."
[4]. Breiman, L. (2001), "Random Forests."
[5]. Zheng, W., et al. (2020), "Automated Entity Extraction from Product Images for E-commerce Applications."
[6]. https://doi.org/10.1145/3065386[1]
[7]. https://doi.org/10.1109/CVPR.2016.90[2]
[8]. https://doi.org/10.1109/ICDAR.1995.598994[3]
[9]. https://doi.org/10.1109/ACCESS.2020.3012534[4]