Title: OCR for Product Labels: A Machine Learning and Machine Vision Approach

A picture containing text, screenshot, font, diagram

Description automatically generated

<https://github.com/openfoodfacts/off-nutrition-table-extractor>

<https://stacks.stanford.edu/file/druid:bf950qp8995/Grubert_Gao.pdf>

1. Introduction

In this paper, we propose a machine learning and machine vision approach for performing Optical Character Recognition (OCR) on product labels. The goal is to automatically extract relevant information from various types of product labels, enabling efficient data retrieval and analysis. Unlike traditional OCR methods, our approach leverages machine learning techniques to improve accuracy and adaptability to different label formats.

2. Methodology

2.1 Data Collection and Preprocessing

We collected a diverse dataset of product labels encompassing different industries and label designs. The dataset was manually annotated with ground truth information for training and evaluation purposes. Text preprocessing techniques were applied, including lowercase transformation, punctuation removal, and stop word removal, to standardize the text data.

2.2 Machine Learning Models

We employed deep learning models for OCR tasks, specifically Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). CNNs were used for feature extraction from label images, while RNNs were employed for sequence modeling and character recognition. The models were trained on the labeled dataset using appropriate loss functions and optimization techniques.

… CRNN.

2.3 Machine Vision Techniques

To enhance the OCR process, machine vision techniques were incorporated. Image preprocessing techniques such as image normalization, denoising, and contrast enhancement were applied to improve the quality of label images. Additionally, advanced image analysis algorithms, including edge detection and segmentation, were used to isolate and extract individual characters or text regions from the labels.

3. Experimental Results

3.1 OCR Accuracy Evaluation

The trained OCR model was evaluated on a separate test dataset comprising diverse product labels. Accuracy metrics such as character-level accuracy, word-level accuracy, and label-level accuracy were calculated to assess the performance of the OCR system. The results demonstrated high accuracy rates, indicating the effectiveness of the proposed approach.

3.2 Comparison with Traditional OCR Methods

…

4. Applications and Limitations

The proposed OCR system for product labels has various applications across industries. It can be utilized in inventory management systems, retail operations, quality control processes, and regulatory compliance. However, the system does have some limitations. Certain label designs, low-quality printing, and complex backgrounds can affect OCR accuracy. Ongoing research is focused on addressing these challenges and improving the system's performance.

5. Conclusion

This paper presents a machine learning and machine vision approach for OCR on product labels. The combination of deep learning models, image preprocessing, and advanced vision techniques enables accurate and efficient extraction of information from diverse label formats. The proposed system offers significant advantages over traditional OCR methods and has wide-ranging applications in various industries. Future work involves further refinement of the system to address limitations and explore additional optimization techniques.

Research links:  
<https://pytorch.org/docs/stable/index.html>

<https://pillow.readthedocs.io/en/stable/index.html>

<https://github.com/pytorch/pytorch>

<https://docs.opencv.org/4.x/>

<https://tesseract-ocr.github.io/tessdoc/Home.html>

<https://www.hindawi.com/journals/complexity/2019/1671340/>