4.1. System Architecture Overview

The proposed expiry date detection system integrates state-of-the-art deep learning models for both vision and language, combined with rule-based pattern matching, to robustly extract expiry information from food product images.

The architecture (see fig. 4.1) comprises three primary modules:

(a) image preprocessing and feature extraction via convolutional neural networks (cnn),

(b) text extraction using optical character recognition (ocr), and

(c) hybrid date entity detection utilizing regular expressions and transformer-based named entity recognition (ner).

4.2 Image Enhancement and Dimensionality Reduction

To guarantee uniformity in image representation, each input image undergoes preprocessing steps such as resizing and normalization. Images are resized to a resolution of 224×224 pixels through bicubic interpolation, ensuring compatibility with the cnn backbone.

A ResNet-18 model, pre-trained on ImageNet and shortened by excluding its final classification layer, is utilized for extracting features. This CNN module generates a 512-dimensional feature vector, which captures important layout and content details from the product label. Although not directly involved in string-based date extraction, these features provide a foundation for potential downstream tasks, such as expiry presence filtering and multi-modal learning.

4.3 Text Extraction using EasyOCR

To recognize text, the system utilizes the easyocr engine, which has been proven to be reliable in handling various fonts and backgrounds commonly found on product packaging. The ocr module analyzes each preprocessed image to identify individual text segments, which are then combined into a single, continuous string representation. This unified string acts as the input for subsequent modules that determine the expiry date. All experiments were conducted using English as the detection language, with gpu acceleration turned off to simulate deployment on standard hardware.

4.4 Expiration Date Identification

The most important part of the pipeline is the expiry date detection sub-system, which uses a combination of rule-based and deep learning techniques to ensure the highest level of reliability. The procedure is as follows:

4.4.1. Regular expression-based pattern matching

Initially, the extracted ocr text is analyzed using a set of regular expressions (regex), specifically designed to handle a wide variety of date formats commonly found in real-world product packaging.

These patterns encompass a range of formats, including 'yyyy-mm-dd', 'dd/mm/yyyy', '22 april 2024', and variations that involve the use of ordinal suffixes and alternative delimiters. All string matches are combined as initial date possibilities.

4.4.2. Transformer-Based Named Entity Recognition

In situations where pattern-based matching falls short, such as when dealing with nonstandard language cues or obscured date delimiters, a fallback semantic extraction is carried out using a transformer-based ner pipeline. Specifically, the system utilizes the "dslim/bert-base-ner" checkpoint, a bert model that has been fine-tuned for the task of token classification. The ner pipeline identifies and extracts spans within the ocr output, specifically focusing on entities labeled as 'date'. To prevent unnecessary duplications, all potential dates are checked for duplicates and whitespace is normalized.

4.5 Annotated Visual Output

To improve model transparency and facilitate user verification, the system overlays the final detected expiry date onto the original image. OpenCV's text interface is used to render detected strings, which are displayed at specific anchor points for easy understanding. Matplotlib is used to visualize data, offering an easy-to-use interface for both qualitative analysis and gathering feedback from end-users.

4.6 Evaluation procedure

To measure the accuracy of the pipeline, the outputs are compared to a dataset that contains manually labeled expiry dates for each image. Evaluation metrics encompass precision, recall, f1-score, and overall accuracy, which are calculated as follows:

For quantitative assessment, the pipeline outputs are evaluated against a ground-truth dataset consisting of manually annotated expiry dates for each image. Evaluation metrics include Precision, Recall, F1-Score, and overall Accuracy, computed as follows:

1. Precision: The proportion of system-predicted dates that are correct.
2. Recall: The proportion of ground-truth dates successfully identified by the system.

3. F1-Score: The harmonic mean of precision and recall.
4. Accuracy: The ratio of images for which the predicted expiry date matches any of the annotated ground truths.

The evaluation loop iterates through the entire dataset, invoking the detection pipeline and aggregating results using these metrics.

4.7 complete pipeline flow

To summarize, the complete inference process involves the following sequential steps:

1. Input acquisition: product label image is provided as input.

2. CNN preprocesses the image to extract visual features for further analysis.

3. The easyocr module effortlessly extracts all visible text and combines it into a single string.

4. Expiry date detection: text is analyzed using regular expressions and, if required, bert-based natural language processing for comprehensive date extraction.

5. annotation and visualization: detected dates are overlaid on the image for transparency.

6. evaluation of performance: the outputs of the pipeline are compared with the manual annotations to calculate precise metrics.