Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 6, pp. 5722–5733, DOI 10.12694/scpe.v25i6.4967

INTEGRATION OF LLM IN EXPIRATION DATE SCANNING FOR VISUALLY IMPAIRED PEOPLE

THEODOR-RADU GRUMEZA[*], BOGDAN BOZGA[†], AND GRIGORE-LIVIU STANILOIU[‡]

Abstract. In this study, the authors explore an approach to detect expiration dates of food products using a live feed stream and the integration with Large Language Models in order to improve accessibility for visually impaired people. The main objective is to enhance their capacity to engage in common tasks like grocery shopping autonomously. The novelty of this research lies in employing Meta LLAMA 2, a large language model, and experimenting with both traditional and a new OCR solution to find the expiration date using image processing. This approach offers audio information about whether the product has expired or when it will expire, helping in shopping and product recognition for visually challenged customers. The proposed solution consists of optical character recognition, mainly the EasyOCR library, fine-tuned on cropped images containing only the expiration dates and a validation phase that filters and checks the extracted data.

Key words: Expiration date scanning, Image processing, Visually impaired people, Large Language Models

1. Introduction. Our research is centred on integrating Large Language Models such as Meta's LLAMA 2 [16] in expiration date scanning applications to improve accessibility for visually impaired people. According to the World Health Organization (WHO), 285 million individuals worldwide experience visual impairments, with 39 million being blind and 246 million having low vision [3].

The identification of the date on which the product is due, which is printed on each perishable product like cosmetics and pharmaceuticals, but especially on food products, is used to specify the date until that product is safe to use or consume.

Automatic expiration date detection [1] is challenging due to the non-existence of a standard and the different fonts, materials, and positions on which they are printed. Due to this inconsistency, a lot of food may be consumed after the due date, which can lead to food poisoning. Finding the due date of the product becomes genuinely concerning in the case of visually impaired people who may find it hard or even impossible to read and understand the expiration date on the products they have in the fridge or pantry.

Some solutions exist, but those are not widely available, easy to use, or work only in particular conditions. The classic OCRs [14] have appeared as a need to recognise characters in papers or other structured formats. These algorithms still have one significant limitation: they are trained to read text from clear and structured images, most commonly synthetically or documents. At the same time, the accent when developing new OCRs or upgrading the already developed ones was heavily towards adding new languages [6] and not detecting specific information in images from the real world. Specifically, product information, such as expiration dates, may be highly obscure and distorted, with a significant variation of lighting or even dotted formats.

To evaluate the effectiveness of detecting expiration dates on products, we compared a standard Optical Character Recognition (OCR) solution with our custom solution that was specifically fine-tuned for this task. Our aim was to determine which method more accurately identifies and reads expiration dates from various product images.

We considered two widely-used OCR systems: EasyOCR and Tesseract. Tesseract is a well-established OCR engine extensively used and researched, as detailed in [15]. Despite its long history and widespread

[*]West University of Timișoara, Faculty of Mathematics and Informatics, Timișoara 300223, Romania (theodor.grumeza@e-uvt.ro)

[†]West University of Timișoara, Faculty of Mathematics and Informatics, Timișoara 300223, Romania (bogdan.bozga01@e-uvt.ro)

[‡]West University of Timișoara, Faculty of Mathematics and Informatics, Timișoara 300223, Romania (grigore.staniloiu02@e-uvt.ro)

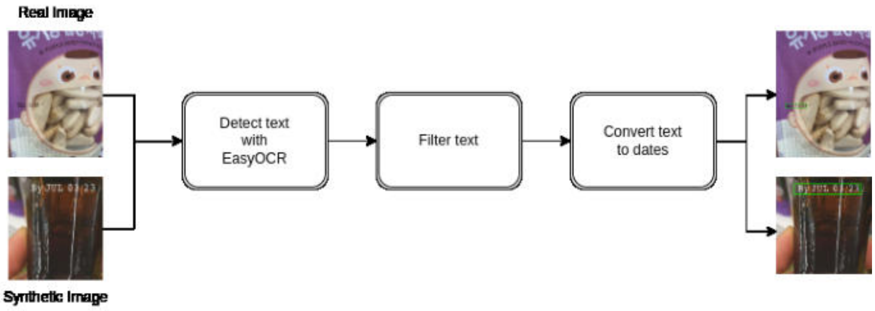

Fig. 1.1: Detection workflow

use, our preliminary tests with Tesseract yielded unclear results for this specific use case. It struggled with recognizing expiration dates on product packaging, including variations in font styles, sizes, and backgrounds.

On the other hand, EasyOCR is a more recent OCR system designed to handle a wide range of text recognition tasks, including reading text from complex real-world images. As referenced in [17], EasyOCR provides enhanced capabilities in recognizing text in diverse conditions, such as varying lighting, orientations, and noisy backgrounds, common in product images.

Given these considerations, EasyOCR was selected for our comparison. Its ability to more accurately read and interpret text from real-world images made it suitable for detecting expiration dates, which often appear in small, variable fonts on uneven surfaces. This choice was instrumental in establishing a baseline performance against which we could measure the effectiveness of our fine-tuned solution. By leveraging the strengths of EasyOCR, we aimed to achieve a more reliable detection and extraction of expiration date information from product images.

The literature review in the second section of this research covers integrating a Large Language Model with expiration date scanning using Optical Character Recognition [9].

The third section refers to the implied methods and the system specifications. It highlights some open problems that could be useful in implementing the approach for combining the LLM with OCR scanning and the phases while implementing the system. The fourth section offers the research results, emphasising our model's accuracy and a comparison with other solutions. The last section is dedicated to conclusions and future work.

2. Related Work. Several specialized solutions have been developed to address the problem of detecting expiration dates on food products. For example, Florea and Rebedea proposed an approach titled "Expiry Date Recognition Using Deep Neural Networks" [18], which employs deep neural networks trained on labelled datasets comprising both real and synthetic images of expiration dates. Their method achieved a general accuracy of approximately 63% when utilizing only real images. However, a significant drawback of their solution is the complexity involved in replicating the implementation and deploying it across various mobile devices or smart glasses platforms.

Previous works have explored various approaches, such as using neural networks with specific feature extraction techniques [11] and direct OCR application at different angles. These studies have highlighted the challenges in accurately identifying expiry dates, given their varying formats, locations, and printing styles on product packaging. Image pre-processing methods, like the Hough transform and adaptive thresholding, have been investigated to improve OCR accuracy, demonstrating varied levels of success. This review can delve into these methods, comparing their results and discussing potential areas for further research and improvement.

The study "Recognition of Expiration Dates Written on Food Packages with Open Source OCR" [7] takes the same approach to the problem, but it uses Tesseract-OCR. The authors developed a smartphone application that recognizes expiration dates on food packages by utilizing Tesseract's capabilities to scan and interpret these dates. The solution uses the OCR without any fine-tuning, demonstrating its capabilities and limitations. The detector struggled in specials when the data were in dot format or on shiny surfaces that reflected light, such as cans. The study's results aren't clear since the general focus wasn't on developing the application and explaining how the solution works (which is real work) but on incorporating it into a smart fridge.

The most notable solution on which the complex implementation and which is used to compare the current solution is the paper "A generalized framework for recognition of expiration dates on product packages using fully convolutional networks" [13]. This solution develops a more complex approach by incorporating fully convolutional networks (FCNs) to enhance the accuracy and performance of expiration date recognition.

The authors of this study designed a complex framework that integrates image pre-processing, segmentation, and recognition steps. They used deep learning to improve the detection and interpretation of expiration dates from diverse packaging. The two authors created a vast dataset with both real and synthetic images to implement and test their solution. This dataset was also used to create and test this solution.

The quality and safety of food products are crucial for human health, with proper labelling of expiry dates playing a vital role in preventing the consumption of unsafe food. Food wastage, mainly due to expiration, has become a significant global concern, as highlighted by the Food and Agriculture Organization (FAO), which reports that approximately 1.3 billion tons of food are wasted annually. This concern has led to the development of various methods for expiry date detection and recognition of food products. Several approaches have been proposed, including Scazzoli et al.'s [12] method using the Hough transform and adaptive Gaussian thresholding for OCR accuracy, Ribeiro et al.'s [10] end-to-end deep neural network architecture for text detection, and Zaafouri et al.'s [20] automated vision approach utilizing a multilayer neural network and S-Gabor filters for image enhancement. Gong et al. [4] also introduced a unified deep neural network combining convolutional and recurrent neural networks for automatic date recognition in food packages.

Despite their contributions, these methods often process entire images to extract expiry dates, leading to extensive computation and inefficiency, especially in real-time applications. Many existing systems also rely on traditional barcodes like EAN-13 and UPC, which lack expiry date information, requiring manual input and monitoring. To address these challenges, the current study proposes a streamlined method that uses the Single Shot Detector MobileNet (SSD MobileNet) for object detection and Attention OCR for text recognition. This approach focuses on detecting the region of interest, significantly reducing redundant processing and enhancing accuracy. SSD MobileNet offers faster processing and lower latency than Faster RCNN, while the Attention OCR model improves text recognition through attention mechanisms. This combination provides a promising solution to reduce food wastage by facilitating more effective monitoring and managing food product expiry dates.

The article of Liyun Gong et al. [5] introduces an innovative camera-based system designed to automatically detect and recognise expiry dates on food packages. This system addresses a critical need for accurate labelling in the food manufacturing sector. Mislabeling expiry dates can lead to serious health issues and significant financial losses due to product recalls. This system aims to improve the reliability and efficiency of expiry date verification, which is currently prone to human error and equipment faults.

The proposed method utilizes a fully convolutional network (FCN), a type of deep neural network (DNN), to detect the expiry date region on food packages. This approach differs from traditional methods that rely on Optical Character Verification (OCV) systems, which can be inconsistent due to variations in expiry date formats, packaging, and camera angles. The FCN is fine-tuned on a dataset of food packages to identify expiry date regions specifically, minimizing the influence of other text on the package.

After detecting the region of interest (ROI), the system extracts the date characters using image processing techniques like the Maximally Stable Extremal Regions (MSER) algorithm. This extraction is followed by character recognition using the Tesseract OCR engine, which classifies the date characters. The process is designed to reduce computational costs by focusing only on the detected ROI rather than the entire image.

The experimental results show that the system achieves a high detection rate of 98% for expiry date regions across various types of food packages in different image formats. However, blurred characters can affect the recognition performance, indicating an area for future improvement. Overall, the system offers a robust solution for enhancing manufacturing food safety and quality assurance by automating the expiry date detection and recognition process.

3. Proposed Methodology. Our research is centred on integrating Meta's LLAMA 2 in expiration date scanning applications to improve accessibility for the visually impaired. This approach wants to offer audio

Table 3.1: Date Patterns for Filters

Pattern	**Example**
`\d{1,2}/\d{1,2}/\d{2,4}`	MM/DD/YYYY or MM/DD/YY
`\d{1,2}-\d{1,2}-\d{2,4}`	MM-DD-YYYY or MM-DD-YY
`\d{2,4}/\d{1,2}/\d{1,2}`	YYYY/MM/DD or YY/MM/DD
`\d{2,4}-\d{1,2}-\d{1,2}`	YYYY-MM-DD or YY-MM-DD
`\d{1,2}\s+\w+\s+\d{2,4}`	”12 Jan 2023”, ”5 April 22”
`\w+\s+\d{1,2},?\s+\d{2,4}`	”January 12, 2023”, ”Apr 5, 22”
`\d{1,2}\s+\w{3}\s+\d{2,4}`	”12 Jan 2023”, ”5 Apr 22”
`\w{3}\s+\d{1,2},?\s+\d{2,4}`	”Jan 12, 2023”, ”Apr 5, 22”
`\d{2,4}\.\d{1,2}\.\d{1,2}`	YYYY.MM.DD
`\d{1,2}\.\d{1,2}\.\d{2,4}`	DD.MM.YYYY or DD.MM.YY
`\d{2,4}/\d{1,2}`	YYYY/MM or YY/MM
`\d{1,2}/\d{2,4}`	MM/YYYY or MM/YY
`\d{1,2}\s+\w{3}\.?\s+\d{2,4}`	”12 Jan. 2023”, ”5 Apr 22”
`\w{3}\.?\s+\d{1,2},?\s+\d{2,4}`	”Jan. 12, 2023”, ”Apr. 5, 22”
`\d{4}\d{2}\d{2}'`	”YYYYMMDD”
`\d{2}\d{2}\d{4}`	”DDMMYYYY”

information about whether the product has expired or when it will expire, helping in shopping and product recognition for visually challenged customers.

Without needing to create a new, more specialized Neural Network that can be used only for this use case, we want to be able to extend it to other similar use cases. To fulfil this, we use a fine-tuned EasyOCR to enhance its capabilities for our specific application. EasyOCR is an open-source OCR tool that is known for combining Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to detect and recognize text [2].

The fine-tuning process was conducted on the Google Colab platform, utilizing the NVIDIA T4 GPU. This setup provided a high-performance training environment where no particular settings were needed.

Regarding the integration of LLM in the context of expiration date identification and information retrieval, some open problems need to be addressed:

- How can we improve the robustness of expiration date detection in varying environmental conditions, such as poor lighting, low-resolution images, or images with occlusions and distortions?
- What strategies can be implemented to handle diverse formats and languages of expiration dates, considering the global variations in date formats and the use of non-standard characters or symbols?
- How can OCR systems be trained to differentiate between expiration dates and other date-related information on product packaging, such as manufacturing dates or batch numbers, to avoid misinterpretation?
- How can multimodal approaches combining OCR, natural language processing, and visual context recognition be developed to enhance the understanding of product labels, including expiration dates, ingredients, and usage instructions?

In order to train the model, we use both real and synthetic images. After fine-tuning the workflow as shown in Figure 1.1, the first step is the detection of texts in the test images. The texts that we extracted go to a filtering step. We use date patterns to obtain only the relevant formats as shown in Table 3.1. Afterwards, we used the parser from the dateutil library to convert the dates extracted using the patterns into datetime format. This conversion removes parts that passed the filters but aren't valid dates. Then, another verification that checks if the dates obtained were between 10 years ago and 30 years from the current date is applied.

Because we want to be sure that the information on the product respects the general date formats, we will use the LLM to double-check the obtained due date and also give audio feedback regarding whether the product has expired or when it will expire.

Fig. 3.1: Real images

3.1. Dataset. The dataset created by Ahmed Zaafouri et al. [21] was used for training and has 1102 real images of products and expiration dates that capture real-world variety, as seen in 3.1. Another 11860 synthetic images designed to extend the number of pictures with the model are trained. There were images in the dataset with food products that had packaging imperfections. Because of this, we had the idea to simulate the distortions of the products to enhance the dataset and make the images look as good as the ones taken in the shopping store, as seen in Figure 3.2.

Real images are important for training the OCR model to recognize the target text in real life because the packaging fonts and light conditions differ from one case to another. The synthetic images enhance the data, variety, and number of available images for training.

Some images contain more than one expiration date, for example, the first image from Figure 3.2. The total number of dates from non-synthetic images is 1244 and 16674 systematical ones 4.1.

3.2. Performance Metrics. The performance of the models was evaluated using several metrics described in a further paragraph. These metrics help us quantify the effectiveness of the OCR system in different scenarios, providing insights into areas that may need improvement.

Correct Detection Rate.

$$\text{Correct}(\%) = \frac{\text{Number of Correct Detections}}{\text{Total Number of Images}} \times 100 \tag{3.1}$$

This equation calculates the percentage of images where the OCR system correctly identifies the expiration date. In our research, this metric is crucial for understanding the model's base accuracy. A higher percentage indicates that the model accurately detects and recognises expiration dates across the dataset.

Partial Detection Rate.

$$\text{Partial}(\%) = \frac{\text{Number of Partial Detections}}{\text{Total Number of Images}} \times 100 \tag{3.2}$$

This metric captures cases where the detected date is incorrect but falls within a reasonable margin of error (less than 365 days from the actual date). This is important in our research because it shows the model's ability

Fig. 3.2: Synthetic images

to approximate dates even if not perfectly accurate, which can be useful in contexts where a close estimation is acceptable.

Missing Detection Rate.

$$\text{Missing}(\%) = \frac{\text{Number of Images with No Valid Date Detected}}{\text{Total Number of Images}} \times 100 \tag{3.3}$$

The missing detection rate measures the percentage of images where the OCR system fails to detect a valid date. This helps us understand the model's limitations in terms of its sensitivity and ability to recognize expiration dates under different conditions. A lower missing rate is desirable, indicating the robustness of the model.

Wrong Detection Rate.

$$\text{Wrong}(\%) = \frac{\text{Number of Incorrect Detections (More than 365 Days Apart)}}{\text{Total Number of Images}} \times 100 \tag{3.4}$$

This metric indicates the percentage of images where the detected date was significantly incorrect (more than 365 days off). It helps identify scenarios where the model makes substantial errors, possibly due to confusing background text or poor image quality. Analyzing these cases can guide improvements in the OCR model or pre-processing steps.

3.3. General methods for recognition of text features. The Convolutional Recurrent Neural Network (CRNN) combines Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to extract and recognize text features. Understanding the equations used in CRNN helps fine-tune and interpret how the model processes the images.

Feature Extraction (CNN). The convolution operation used in the CNN layer:

$$(I * K)(i, j) = \sum_m \sum_n I(m, n) \cdot K(i - m, j - n) \tag{3.5}$$

This equation represents the convolution operation in the CNN part of CRNN, which extracts features from the input images. In our research, this step is crucial for identifying patterns within the image indicative of text, such as edges or shapes corresponding to characters. By adjusting the kernel K, we can refine what features are extracted, potentially improving the model's accuracy in detecting expiration dates.

Sequence Modeling (RNN). Using Long Short-Term Memory (LSTM) cells to model the sequence:

$$h_t = o_t * \tanh(C_t) \tag{3.6}$$

where:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{3.7}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3.8}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3.9}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{3.10}$$

These equations describe the LSTM units used in the RNN part of CRNN. They are responsible for capturing the sequential dependencies in the detected text features, allowing the model to understand the context and order of characters in the expiration dates. In our research, this is important for ensuring that the text is recognized in the correct order, especially in cases where characters are closely spaced or overlapping.

Connectionist Temporal Classification (CTC) Loss. CTC is used to align input and output sequences:

$$\text{CTC Loss} = -\log(p(y|x)) \tag{3.11}$$

where:

$$p(\pi|x) = \prod_{t=1}^{T} y_{\pi_t}^{(t)} \tag{3.12}$$

CTC loss is used during training to allow the model to predict sequences without requiring pre-segmented inputs. In our research, this is particularly useful for handling varying lengths of expiration dates and different text alignments within images. By using CTC, we can train the model more flexibly, improving its generalization to different formats of expiration dates.

3.4. Enhanced methods for character-level detection and localisation. Character Region Awareness for Text Detection (CRAFT) is used for character-level detection and image localisation. It allows us to precisely detect individual characters, making it a complementary approach to CRNN for recognizing text.

Text Region Localization. CRAFT focuses on character-level detection using: Region Score: Represents the probability of a character at each location. Affinity Score: Represents the link between neighboring characters.

In our research, CRAFT's ability to localize individual characters improves the detection of expiration dates, especially in cluttered backgrounds or images with irregular text arrangements. Using the region and affinity scores, we can identify isolated characters and connected components, enhancing the model's performance on complex images.

Loss Function for CRAFT. The training loss for CRAFT is a combination of region and affinity losses:

$$L = L_{region} + \lambda \cdot L_{affinity} \tag{3.13}$$

This equation is used during the training of the CRAFT model. The loss function combines the errors in detecting individual characters (region loss) and the links between characters (affinity loss). By minimizing this loss, we fine-tune the model to accurately detect and connect characters, which is critical in our research for recognizing expiration dates in various layouts and fonts.

3.5. Comparison and Analysis. The performance of the models is compared based on accuracy and time efficiency. The comparison is done on the entire image and only on the section that contains the due date. These comparisons helped us determine the most effective OCR approach for our specific use case.

Table 4.1: Default Performance

Detector	Correct (%)	Partial (%)	Missing (%)	Wrong (%)
Default E	61.4%	7.3%	30.2%	1.1%
DBnet E	61.96%	6.25%	30.36%	1.43%
Default C	62.86%	5.36%	30.36%	1.43%
DBnet C	63.21%	5.0%	30.36%	1.43%

Accuracy Comparison. To observe the differences between detection rates for entire images versus cropped images:

$$\Delta\text{Accuracy} = \text{Accuracy}_{\text{cropped}} \times \text{Accuracy}_{\text{entire}} \tag{3.14}$$

This equation allows us to quantify the impact of using cropped images versus entire images for training and testing. By comparing the accuracy, we can understand how focusing on specific regions (e.g., where the expiration date is located) can improve the model's performance. This informs our decision on whether pre-processing steps like cropping are beneficial in real-world applications.

Average Time per Image. To evaluate the efficiency of the models:

$$\text{Average Time} = \frac{\text{Total Time Taken}}{\text{Number of Images}} \tag{3.15}$$

This equation measures the computational efficiency of the OCR system. In our research, this is important for assessing the feasibility of using the model in real-time applications or large-scale processing. A lower average time per image indicates a more efficient model, which is crucial for scenarios requiring quick results.

3.6. Iteration and Model Tuning. The effect of different numbers of training iterations on performance:

$$\text{Performance} \propto \text{Number of Iterations} \tag{3.16}$$

This relationship indicates that the model's performance typically improves with more training iterations up to a point of convergence. In our research, we use this principle to determine the optimal number of iterations for training, balancing between underfitting and overfitting. We can select the best model for accurate expiration date detection by monitoring the performance over iterations.

4. Experimental Results. We have done experiments to check and compare how this type of solution performs in different situations by comparing the performance of various stages of the fine-tuning with the due date written in different formats.

The validation set of images was constant, and we took into account the following metrics:

- *Correct* - represents the percentage of images in which the expiration date was identified correctly;
- *Partial* - represents the percentage of images in which the identified date wasn't correct, but the detected date was less than one year (365 days) apart from the real one;
- *Missing* - is the percentage of images with no valid date after the filtering step;
- *Wrong* - the number of images with a valid date was detected, but it was further apart than 365 days from the real one.

4.1. Default results. To check the OCR's default capabilities, we first run the detection without fine-tuning the model. Two detectors were used, the first being the default one with the model weights for the English language and the second using DBnet detector [8].

The results are presented in Table 4.1 where we evaluate both the full images, indicated by rows labelled 'E', and the cropped sections containing expiration dates, labelled 'C'.

Table 4.2: Performance for Cropped Images

Iterations	Correct(%)	Partial(%)	Missing(%)	Wrong(%)
10k	61.4%	7.3%	30.2%	1.1%
30k	61.96%	6.25%	30.36%	1.43%
50k	62.86%	5.36%	30.36%	1.43%
70k	63.21%	5.0%	30.36%	1.43%
90k	63.21%	4.82%	30.54%	1.43%
100k	63.21%	4.64%	30.71%	1.43%
Best Acc.	63.21%	5.89%	29.82%	1.07%

4.2. CRNN Architecture fine-tuned. The Convolutional Recurrent Neural Network (CRNN) backbone architecture was used, which is a text recognition network specifically designed to recognize sequences of characters in images accurately. The CRNN architecture combines Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to extract features effectively and capture the sequential nature of text.

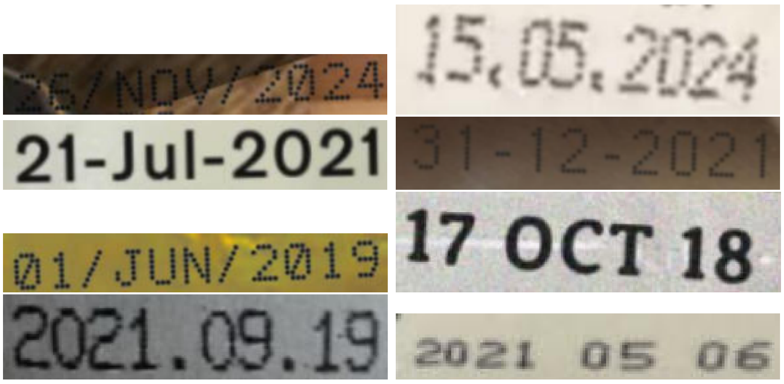

Fig. 4.1: Cropped dates

Fine-tuned using only cropped images. The model was trained on cropped parts of the images containing only the expiration dates as presented in Figure 4.1. This had the purpose of checking its accuracy when only the wanted text is present and no other unwanted information, as shown in Table 4.2. The total number of images tested is 666, all of which are real; the average time per image is around 0.022 seconds, including the detection and filtration phases.

The last row of Table 4.2, entitled ”Best Accuracy”, represents the result for the model that achieved the highest accuracy during the training phase. Overall, all the results are promising due to the low percentage of partial and wrong detections.

In a real-life use case for the missing detection, another image will be taken, and considering the low time needed for each image, this approach represents a valid solution.

Fine-tuning using the entire images. In this case, we trained on the entire image that contains the expiration dates, other text and noise. The purpose was to compare the accuracy in the case when no before-hand cropping was done. As expected, the results are not as good as in the case with only the wanted text, as can be seen in Table 4.3.

The same images were used to evaluate the previous model, but this time, the entire image was analyzed rather than just a cropped section. The average processing time per image is approximately 0.945 seconds.

The issue arises when dealing with large images, as the detector has difficulty distinguishing the expiration date from other text, occasionally merging them. This leads to more text being detected, which means more

Fig. 4.2: Partial Detection

Fig. 4.3: Wrong detection

Table 4.3: Performance for Uncropped Images

Iterations	Correct(%)	Partial(%)	Missing(%)	Wrong(%)
10k	48.49%	4.52%	0%	46.99%
30k	49.12%	4.47%	0.1%	46.31%
50k	49.68%	4.39%	0.15%	45.78%
70k	50.21%	4.29%	0.19%	45.31%
90k	50.62%	4.18%	0.22%	44.98%
100k	51.02%	4.07%	0.25%	44.66%

text might pass through the filter and sometimes get incorrectly converted into dates, negatively affecting the statistics. For instance, numbers from barcodes might be misinterpreted as valid dates, as seen in the first image, the second row of Figure 4.3. In this scenario, the LLM is employed to further eliminate unnecessary text, enhancing the accuracy of expiration date detection on the product.

4.3. CRAFT Architecture Fine-Tuned. We also experimented with the CRAFT backbone architecture, a text detection network specifically designed to identify and localize text regions in images. CRAFT is adept at detecting both regular and irregular text arrangements, including curved and rotated text [19].

Training and testing were conducted on the same image sets used for the CRNN section.

The primary difference between CRNN and CRAFT lies in their text detection approach. While CRNN is designed to detect entire words in an image, CRAFT focuses on localizing and detecting individual characters and symbols.

A significant drawback of CRAFT is that it is more time-consuming to train and slower than CRNN. Due to time constraints imposed by Google Colab, we could only complete a relatively small number of iterations.

For a model with 5000 iterations, testing on entire images yielded results of 49.1% correct, 4.37% partial, 41.27% missing, and 5.27% wrong, with an average processing time of 0.92 seconds per image. Using the same model on cropped parts of the images containing only the date, the results were 51.79% correct, 9.64% partial, 19.64% missing, and 18.93% wrong. This comparison indicates that the CRAFT model performs better on uncropped images, highlighting the importance of contextual information.

Figure 4.2 provides examples of expiration dates that were incorrectly recognized, typically due to one character being misinterpreted. Figure 4.3 illustrates cases where dates were detected incorrectly, often because other numbers or text, such as barcodes, were misinterpreted as valid dates.

5. Conclusion and Future work.

5.1. Conclusion. In this study, we proposed a new approach to assist visually impaired individuals in identifying expiration dates on food products through the use of live feed streams and the integration of Large Language Models, specifically leveraging Meta LLAMA 2. By incorporating image processing techniques with optical character recognition (OCR), primarily utilizing the EasyOCR library enhanced with CRNN and CRAFT architectures, we aimed to provide an accessible and scalable solution. The results indicate that while our solution does not introduce a novel neural network architecture, it demonstrates the capacity to be readily updated and refined, making it a practical alternative to more complex models. Despite the current solution requiring further optimization to compete with the high accuracy rates of existing models, such as the work by Ahmet Cagatay Seker and Sang Chul Ahn [13], it establishes a solid groundwork for future advancements. The CRAFT architecture, in particular, shows considerable promise in terms of effectiveness in this use case, suggesting its potential for further development. Overall, our application meets its intended goal of enhancing accessibility for visually impaired individuals, enabling them to perform grocery shopping tasks with greater autonomy.

5.2. Future Work. Building on the promising initial results of the CRAFT architecture, future research will focus on several key areas to refine and extend the capabilities of the proposed solution. First, we plan to perform extensive model optimization and fine-tuning using a larger, more diverse dataset, which will enhance the model's accuracy and robustness across various expiration date formats and packaging types. Integrating the system with real-time processing capabilities will also be a priority, allowing for seamless, live detection and feedback to users in dynamic environments like grocery stores. Moreover, expanding the system's language support to recognize expiration dates in multiple languages is critical for its applicability in different regions, catering to a global user base. To increase the reliability of expiration date detection, advanced validation mechanisms incorporating contextual information from product labels will be developed, reducing false positives. We also aim to enhance the user interface, focusing on delivering more intuitive audio and haptic feedback options to ensure a user-friendly experience for visually impaired individuals. Finally, extensive field testing and deployment in real-world environments will be conducted to gather comprehensive user feedback, enabling iterative improvements to both the system's performance and user experience. These future directions will collectively contribute to evolving the solution into a highly accurate, reliable, and accessible tool, significantly benefiting visually impaired individuals' independence and quality of life.

REFERENCES

[1] A. Abdulraheem and I. Y. Jung, *Effective digital technology enabling automatic recognition of special-type marking of expiry dates*, Sustainability, 15 (2023), p. 12915.

[2] J. AI, *Easyocr*, 2023. Accessed: 2024-07-25.

[3] H. Buch, T. Vinding, and N. V. Nielsen, *Prevalence and causes of visual impairment according to world health organization and united states criteria in an aged, urban scandinavian population: the copenhagen city eye study*, Ophthalmology, 108 (2001), pp. 2347–2357.

[4] L. Gong, M. Thota, M. Yu, et al., *A novel unified deep neural networks methodology for use by date recognition in retail food package image*, SIViP, 15 (2021), pp. 449–457.

[5] L. Gong, M. Yu, W. Duan, X. Ye, K. Gudmundsson, and M. Swainson, *A novel camera based approach for automatic expiry date detection and recognition on food packages*, in Artificial Intelligence Applications and Innovations: 14th IFIP WG 12.5 International Conference, AIAI 2018, Rhodes, Greece, May 25–27, 2018, Proceedings 14, Springer, 2018, pp. 133–142.

[6] O. Ignat, J. Maillard, V. Chaudhary, and F. Guzmán, *Ocr improves machine translation for low-resource languages*, arXiv preprint arXiv:2202.13274, (2022).

[7] R. H. W. Kento Hosozawa and K. Mizutani, *Recognition of expiration dates written on food packages with open source ocr*, International Journal of Computer Theory and Engineering, (2018).

[8] M. Liao, Z. Zou, Z. Wan, C. Yao, and X. Bai, *Real-time scene text detection with differentiable binarization and adaptive scale fusion*, 2022.

[9] G. Monteiro, L. Camelo, G. Aquino, R. d. A. Fernandes, R. Gomes, A. Printes, I. Torné, H. Silva, J. Oliveira, and C. Figueiredo, *A comprehensive framework for industrial sticker information recognition using advanced ocr and object detection techniques*, Applied Sciences, 13 (2023), p. 7320.

[10] F. D. S. Ribeiro et al., *An end-to-end deep neural architecture for optical character verification and recognition in retail food packaging*, in 2018 25th IEEE International Conference on Image Processing (ICIP), IEEE, 2018, pp. 2376–2380.

[11] D. Scazzoli, G. Bartezzaghi, D. Uysal, M. Magarini, M. Melacini, and M. Marcon, *Usage of hough transform for expiry date extraction via optical character recognition*, in 2019 Advances in Science and Engineering Technology International Conferences (ASET), 2019, pp. 1–6.

[12] D. Scazzoli, G. Bartezzaghi, D. Uysal, M. Magarini, M. Melacini, and M. Marcon, *Usage of hough transform for expiry date extraction via optical character recognition*, in 2019 Advances in Science and Engineering Technology International Conferences (ASET), IEEE, 2019, pp. 1–6.

[13] A. C. Seker and S. C. Ahn, *A generalized framework for recognition of expiration dates on product packages using fully convolutional networks*, Expert Systems with Applications, 203 (2022), p. 117310.

[14] K. M. Shanthini, P. Chitra, S. Abirami, G. Aninthitha, and P. Abarna, *Recommendation of product value by extracting expiry date using deep neural network*, in 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, 2021, pp. 1–7.

[15] R. Smith, *An overview of the tesseract ocr engine*, in Ninth international conference on document analysis and recognition (ICDAR 2007), vol. 2, IEEE, 2007, pp. 629–633.

[16] H. Touvron, L. Martin, K. Stone, P. Albert, A. Almahairi, Y. Babaei, N. Bashlykov, S. Batra, P. Bhargava, S. Bhosale, et al., *Llama 2: Open foundation and fine-tuned chat models*, arXiv preprint arXiv:2307.09288, (2023).

[17] D. Vedhaviyassh, R. Sudhan, G. Saranya, M. Safa, and D. Arun, *Comparative analysis of easyocr and tesseractocr for automatic license plate recognition using deep learning algorithm*, in 2022 6th International Conference on Electronics, Communication and Aerospace Technology, IEEE, 2022, pp. 966–971.

[18] T. R. Vlad Florea, *Expiry date recognition using deep neural networks*, International Journal of User-System Interaction, (2020).

[19] C. A. R. N. C. Youngmin Baek, Bado Lee, *Character region awareness for text detection*, Computer Vision and Pattern Recognition, (2019).

[20] A. Zaafouri et al., *A vision approach for expiry date recognition using stretched gabor features*, Int. Arab J. Inf. Technol., 12 (2015), pp. 448–455.

[21] A. Zaafouri, M. Sayadi, F. Fnaiech, O. al. Jarrah, and W. Wei, *A new method for expiration code detection and recognition using gabor features based collaborative representation*, Advanced Engineering Informatics, 29 (2015), pp. 1072–1082.