A novel camera based approach for automatic expiry date detection and recognition on food packages

Liyun Gong[1], Miao Yu[1], Wenting Duan[1], Xujiong Ye[1],
Kjartan Gudmundsson[2] and Mark Swainson[2]

[1] Scool of Computer Science,University of Lincoln, Lincoln, LN6 7TS, UK,
{lgong,myu,wduan,xye}@lincoln.ac.uk
[2] National Centre for Food Manufacturing,University of Lincoln, Lincoln, LN6 7TS, UK,
{kgudmundsson,mswainson}@lincoln.ac.uk

Abstract. There is abundant of information on food packages, which include the food name, the expiry date and the ingredients. These information, especially the expiry date needs to be coded correctly before the products can be released into the market/supply chains. Failure of printing the correct expiry date can lead to both the health issues to the public and financial issues for recalling product back and even reimbursement. In this paper, we develop an automatic system that can achieve the expiry date region detection and recognition in an efficient and effective way. A deep neural network (DNN) based approach is firstly applied to find the expiry date region on the food package. The date characters are then extracted and recognized through the image processing and machine learning methods from the expiry date region. The system is the first camera based automatic system for recognizing expiry date on food packages. And the results tested on different types of food packages show that the system can achieve good performance on both detection and recognition of the expiry date.

Keywords: fully convolutional network, stroke width transform, Maximally stable extremal regions, tesseract OCR

1 Introduction

In the European Union food production is the largest manufacturing sector where it accounts for 13.3% of the total EU-28 manufacturing sector with a reported turnover of 945 billion [1]. Whilst food availability is a primary concern in developing nations and food quality(value) is a focal point in more affluent societies, food safety is a requirement that is common across all food supply chains. Food safety in the sector is typically underpinned by food science and technology and assured by a combination of operational control systems and procedures including Good Manufacturing Practice (GMP) and Hazard Analysis Critical Control Point (HACCP) [2].

The food product information printed on the food package is a vital for the food safety. Pre-packaged food product information which are incorrectly labelled, especially the expiry date results in product recalls as the fault/issue could cause a food safety incident such as food poisoning due to the consumption of product which is past its actual safe Use-by date. These recalls are usually at very high financial and reputational cost to food manufacturers.

The reasons/root causes for issues or mistakes resulting in label faults on food packaging are many and varied. They include human error and equipment faults. For example, a label printer on a production line can break down and the line carries on running. The faulty packaging therefore needs to be identified and the production line stopped. A common current process line Use-by check approach is to use a human operator to read and verify the packaging label. This check is conducted by either manually picking a pack from the line for inspection or verifying it through an image captured of the pack. However, these methods create mundane and repetitive tasks and therefore place the operator in an error-prone working environment.

Another common approach to control date codes is to use Optical Character Verification (OCV). This involves a supervisory system holding the correct expiry date string and transferring it to both the printer and the vision system. The latter will then verify its read and actions are taken depending on the result. However, OCV systems rely on consistency in expiry date format, packaging and camera view angle. This consistency tends to be hard to achieve in the food and drink manufacturing environment and therefore there is a great need for a more robust solution.

In this work, we have developed an automatic system based on a camera, which can efficiently and effectively recognize the expiry date information printed on different types of food packages. Food packages with wrong expiry date printed on will be picked up. This system will enable far greater control over the accuracy and legibility of critical 'Use-by'/'Best Before' dates and also key traceability information in food and drink manufacturing operations, resulting in significantly increased food safety and compliance with related legislation.

To develop such a system, the first step is to identify expiry date regions as the region of interest (ROI) on a recorded food package image. And the expiry date recognition task is then performed within the ROI instead of the whole image. In this way, the computational costs can be saved to a large extent. One straightforward method to determine ROI is applying the text detection method for detecting text regions as ROI on a food package. For text detection, a number of traditional image processing based techniques have been applied, examples include Stroke Width Transform (SWT) based appraoch [3] and Maximally Stable Extremal Regions (MSER) based approach [4]. With the deep learning techniques having become mainstream in the image processing, computer vision and machine learning communities, different types of deep neural networks have been applied for the text detection ([5,6]) with better results being obtained.

However, if the food package contains too much other text information in addition to the expiry date (that is the usual situation on the food package), the

obtained ROI will still be tremendous. Instead of the text region detection for the ROI identification, in our work we apply a deep neural network approach for directly identifying the expiry date region as ROI. The fully convolutional network (FCN) in [5], which is originally developed for text detection, is fine tuned by our dateset for expiry date region detection. By adopting such an approach, only the expiry date region can be extracted while other texts on the food package are excluded. The most precise ROI is directly obtained and computational costs can then be further reduced by performing the recognition on only the ROI of the expiry date region.

Based on the extracted ROI, the date characters blobs in the ROI can be directly extracted. Related shape features are then extracted for classification by an efficient nearest neighbor method. In our experiment, we have tested our system for both expiry date region detection and classification on different types of food packages in different captured image formats (colour/grayscale), with good results being obtained.

2   Method

In this section, we present the methodology for expiry date recognition on the food package, which is divided into two parts: expiry region identification and recognition. The block diagram of the proposed methodology is shown in Fig. 1. Details of every block are presented as follows.
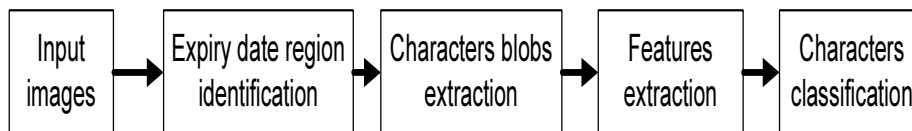

Fig. 1. The block diagram of the proposed system.

2.1   Date code region identification

For effectively identifying the expiry date region on the food package which contains different types of pictures/texts contents with different colors, a deep neural network based approach is applied. The deep neural network structure is a fully convolutional network (FCN) as described in [5], which was originally developed for detecting texts. The network is fine tuned on our food package dataset for detecting the date expiry region.

The FCN structure is shown in Fig. 2, which is decomposed into three parts: feature extractor stem, feature-merging branch and output layer.

The stem part is a PVANet [7], with interleaving convolution and pooling layers. Four levels of feature maps, denoted as f_i are extracted from the original
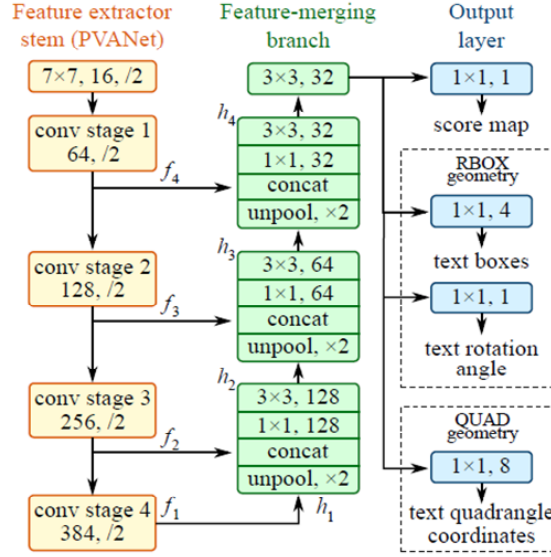
Fig. 2. The FCN structure for the expiry date identification.

input image, whose sizes are $\frac{1}{32}$, $\frac{1}{16}$, $\frac{1}{8}$ and $\frac{1}{4}$ of the original input image. Features from different scale levels meet the requirements of detecting text regions with different sizes.

In the feature-merging branch, features are merged in the following strategy:

$$g_i = \begin{cases} unpool(h_i) & if \ \ i \leq 3 \\ conv_{3 \times 3}(h_i) & if \ \ i = 4 \end{cases}$$
$$h_i = \begin{cases} f_i & if \ \ i = 1 \\ conv_{3 \times 3}(conv_{1 \times 1}([g_{i-1}; f_i])) & if \ \ i = 4 \end{cases} \tag{1}$$

where g_i is the merge based as in [5] and h_i is the merged feature map. And the operator $[;]$. represents concatenation along the channel axis. In each merging stage, the feature map from the last stage is first fed to an unpooling layer to double its size, and then concatenated with the current feature map. A $conv_{1 \times 1}$ bottleneck cuts down the number of channels to reduce computation, followed by a $conv_{3 \times 3}$ that fuses the information to finally produce the output of this merging stage. Following the last merging stage, a $conv_{3 \times 3}$ layer produces the final feature map of the merging branch and feed it to the output layer.

The final output layer contains several $conv_{1 \times 1}$ operations to project 32 channels of feature maps into 1 channel of score map F_s, which gives the likelihood that a pixel belong to the expiry date region as well as a multi-channel geometry map F_g, which could be either rotated box (RBOX) or quadrangle (QUAD) representing different geometries. RBOX geometry map contains a 4-channel map

representing 4 distances from every pixel location to the top, right, bottom, left boundaries of a rectangle enclosing the candidate expiry date region, as well as a 1-channel map representing the angle of the related rectangle. QUAD geometry map is a 8-channel map, which contains the coordinate shift from four corner vertices of a quadrangle (representing candidate expiry date region) to every pixel position.

FCN training and testing For obtaining the network parameters, firstly, a loss function is defined as:

$$L = L_s + \lambda L_g \tag{2}$$

where L_s and L_g represent losses for score and geometry maps respectively, while λ is a balancing parameter.

The term L_s is defined as:

$$L_s = -\beta Y^* log\hat{Y} - (1 - \beta)(1 - Y^*)log(1 - \hat{Y}) \tag{3}$$

where $\hat{Y}$ and Y^* represent the predicted and groundtruth score maps respectively. β is a balancing parameter. While the L_g is defined as scale-invariant IoU loss for RBOX geometry map and scale-normalized smoothed-L1 for the QUAD one as [5]. Based on the defined loss function, the network is trained end-to-end using ADAM optimizer until performance stops improving.

To determine the final expiry date region, first a threshold is set to find positions at which score map values are larger than it. The geometries associated with those positions on the geometry map will then be merged by the locality-aware Non-Maximum Suppression (NMS) to determine the final expiry date region, which can achieve lower computational costs compared with the basic NMS algorithm. Under the assumption that the geometries from nearby pixels tend to be highly correlated, the locality-aware NMS is proposed to merge the geometries row by row. And while merging geometries in the same row, the geometry currently encountered will be merged with the last merged one. In this way, the computational costs could be reduced from $O(n^2)$ of the original NMS to $O(n)$, where n is the number of candidate geometries. Fig. 3 shows the results of different parts the expiry date region detection procedure.

2.2 Expiry date recognition

Expiry date will then be recognized based on the identified region by Tesseract OCR [10]. The Maximally Stable External Regions (MSER) algorithm will firstly be applied, to make a binarization of the extracted date code region with characters being differentiated from the background (Fig. 4 (b)). Component connected analysis [9] is then made to find blobs representing different characters, with small noisy blobs being filtered out (Fig. 4 (c)).

As in [8], for each candidate blob, the boundary will be extracted as in Fig. 4 (d) (here the Canny edge extraction operator is applied) and the corresponding
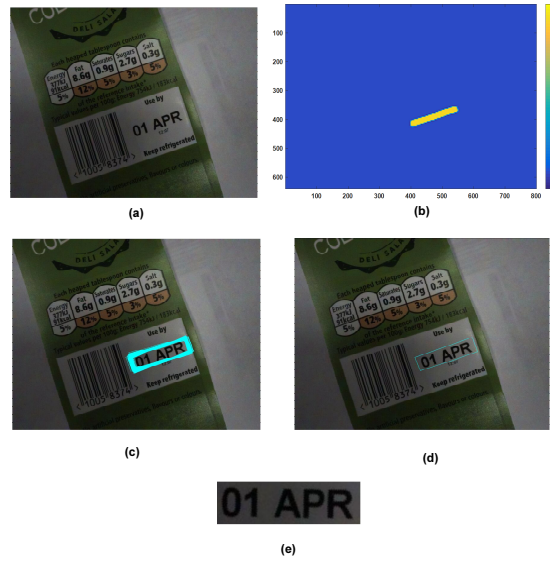
Fig. 3. The expiry date region detection procedure. (a) Original image (b) Score map output by the FCN (c) Candidate expiry date region results before NMS (d) Expiry date region results after NMS (e) Expiry date patch output

shape features, such as topological and polynomial approximation can be extracted for characters classification. In this work, a simple but effective nearest neighbour (NN) approach is applied for the classification. The features of every blob will be compared with prototypes representing different characters. A blob will be classified as the character for which the related distance is the smallest.

3 Experimental results

The proposed system is trained and tested on different types of food packages, with representative examples being shown in Fig. 5. We have collected 800 images from stores, among which 70% (560 images) are used for training and 30% (240 images) are used for testing.

3.1 Expiry date region detection evaluation

The FCN as mentioned in the previous section is fine tuned for identifying the expiry date region on the food package. We have manually masked the ground truth expiry date regions in the training dataset for tuning the FCN, to transfer it from a text detection network to a date code region detection one. We train the network in the GPU-supported tensorflow environment, with two PASCAL GPUs. The input image patches are resized to 512×512. Mini-batch training is
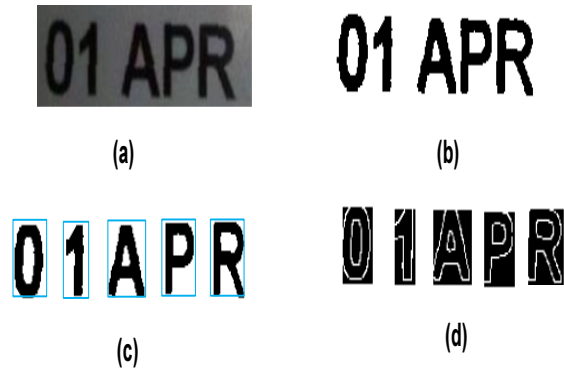
Fig. 4. Extracted expiry date region and related processing. (a) Original date code region (b) Binarization for differentiating character regions from the background (c) Character blobs extraction (d) boundary extraction

Fig. 5. Examples of food package pictures containing date code information.

(a)

(b)

Fig. 6. FCN detection results before/after the fine tuning. (a). Text region detection results before the fine-tuning (b). Date code regions detection results by the FCN after the fine tuning

Fig. 7. FCN detection results of the date code on different packages.

	Char	X	Y	Width	Height
1	0	10	9	26	39
2	1	40	10	17	37
3	A	80	10	34	37
4	P	116	10	27	36
5	R	148	10	33	36

	Char	X	Y	Width	Height
1	2	10	8	25	37
2	4	37	10	27	34
3	N	80	9	31	37
4	O	116	9	32	37
5	V	151	10	32	35

	Char	X	Y	Width	Height
1	2	11	8	25	36
2	2	39	8	24	36
3	N	81	8	30	36
4	O	116	8	32	36
5	V	151	8	31	36

(a)

	Char	X	Y	Width	Height
1	2	12	12	25	35
2	0	39	10	25	38
3	N	81	11	31	36
4	9	117	13	32	35
5	"	155	22	24	24

	Char	X	Y	Width	Height
1	1	11	11	17	35
2	0	37	11	25	36
3	N	79	12	32	35
4	O	115	11	33	37
5	'	151	13	13	21
	!	163	13	20	34

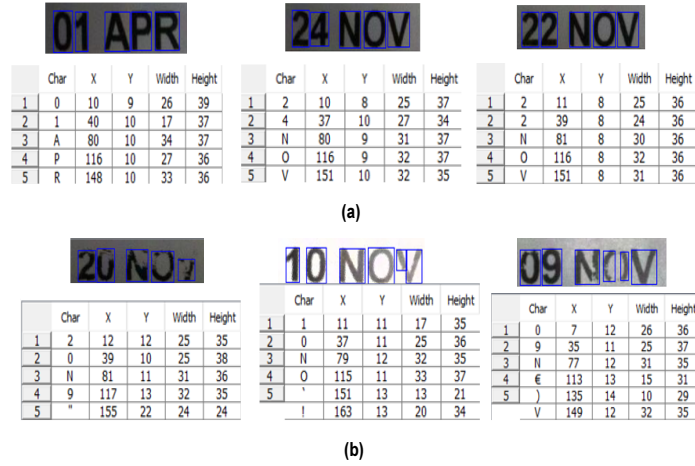|	Char	X	Y	Width	Height
1	0	7	12	26	36
2	9	35	11	25	37
3	N	77	12	31	35
4	€	113	13	15	31
5	)	135	14	10	29
	V	149	12	32	35

(b)

Fig. 8. Texts classification results by Tesseract OCR. (a). Classification results on clear date characters (b). Classification results on blurred characters

applied with the batch size is 14 per GPU and the learning rate is set to be 0.0001 in the ADAM optimizer.

Fig. 6 shows the comparison results between FCN before and after fine tuning. We can see that after fine-tuning, the original text detection network transfers from a text detection network to a expiry date detection one. The transferred network is tested on captured images of different food packages, with both colour and grayscale formats. Related results are presented in Fig. 7, we can see that the date code regions on different food packages can be successfully identified. For a qualitative analysis, we test the developed FCN on the aforementioned testing dataset containing 240 images. The testing results show that 236 out of 240 images, the date region is correctly identified with 4 out of 240 are missing. A total detection rate of 98 % is obtained.

3.2   Expiry date recognition evaluation

Based on the detected expiry date region, the characters within it are extracted and classified. We have applied the Tesseract OCR [10] for classification, which has implemented the characters extraction, feature extraction and classification steps as mentioned in the section 2.2. Some initial results are presented that in Fig. 8. It is shown that the characters on the extracted expiry date region can be successfully recognized; however, some classification mistakes may happen when characters are blurred.

4   Conclusions

In this work, we have developed a novel food package expiry date recognition system based on the camera. A FCN deep neural network approach is applied to detect the expiry date region. Based on the detection results, the date character blobs will be extracted. Related features will be extracted and classified to particular characters. Such a system will potentially advance the assurance of food quality and safety. The experimental results show that the proposed method can achieve very good performance in identifying the expiry date region and classifying characters correctly when they are clear. However, if the characters are blurred, misclassifications can be made and that will be the future researches which will be investigated.

References

1. Eurostat   (2014).   http://ec.europa.eu/eurostat/statistics-explained/index.php/Manufacturing_statistics_-_NACE_Rev._2.
2. WHO FAO (2009). http://www.fao.org/docrep/012/a1552e/a1552e00.htm.
3. Epshtein, B., Ofek, E., Wexler, Y.: Detecting text in natural scenes with stroke width transform. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA (2010).
4. Chen, H., Tsai, S., Schroth, G., Chen, D., Grzeszczuk, R., Girod, B.: Robust text detection in natural images with edge-enhanced Maximally Stable Extremal Regions. In: 18th IEEE International Conference on Image Processing (ICIP), Brussels, Belgium (2011).
5. Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W., Liang, J.: EAST: An Efficient and Accurate Scene Text Detector. In: International conference on Computer Vision and Pattern Recognition, Honolulu, Hawaii,USA (2017).
6. Zhang, Z., Shen, W., Yao, C., Bai, X.: Symmetry-based text line detection in natural scenes. In: International conference on Computer Vision and Pattern Recognition, Boston, Massachusetts,USA (2015).
7. Kim, K., Hong, S., Roh, B., Cheon, Y., Park, M.: PVANET: Deep but lightweight neural networks for realtime object detection. In: arXiv preprint arXiv:1608.08021 (2016).
8. Smith, R.: An Overview of the Tesseract OCR Engine. In: Ninth International Conference on Document Analysis and Recognition,Parana, Brazil (2007).
9. Gonzalez, R.: Digital image processing. Third Edition, Pearson Education (2008).
10. Tesseract-ocr. https://github.com/tesseract-ocr/tesseract.