Non-Invasive Detection of Expired Medicines Using Artificial Intelligence and Image Analysis

Prarthita Kuri, *Anindita* Das Bhattacharjee*, *Rounak* Saha

Department of Computer Science & Engineering, Institute of Engineering and Management (IEM), University of Engineering and Management Kolkata, IEM Centre of Excellence for InnovAI, Kolkata, West Bengal, India

Abstract. This study investigates non-invasive techniques for determining whether drugs are safe, expired, or deteriorating using packaging images. This method uses the extraction of visual and text characteristics driven by artificial intelligence (AI) as opposed to conventional methods for manipulation, chemical analysis or barcode scanning. Optical Character Recognition (OCR) is used to identify expiry-related text, to designate image features using the vector neural network (CNN) and to designate relationships to the graph neural network (GNN). In addition to self-clicking test photos, an image data set for pharmaceutical tablet packs issued. The model obtained accuracy (0.91), accuracy (0.91), recall (0.89) and F1 score (0.90). The results show that the packaging characteristics have sufficient patterns to accurately assess the quality of drugs and allow safer and more scalable quality control without destroying seals.

Keywords: Convolutional neural network, Graph convolutional network, Graph neural network, Medical expiry, Optical character recognition, Pharmaceutical safety.

*Corresponding Author: anindita.dasbhattacharjee@iem.edu.in

1 Introduction

The use of outdated medications is an increasing concern, especially in underprivileged and rural areas, according to recent studies. While 85.2% of participants in a study conducted in rural South India, were aware that expired medications could be dangerous, illiteracy was a significant obstacle to verifying expiration dates [1]. In India, 78.6% of

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

respondents reported that they regularly toss their drugs in the trash and store them incorrectly [2]. According to a Brazilian study, 75% of participants disposed of improperly [3]. 93.75% of homes in Western India kept medications in storage, while urban areas had a larger percentage of over-the-counter medications (16.76%) than rural ones (11.82%).

Reduced effectiveness and serious health risks are only two of the issues that expired medications may present. Although many medications retain 70–80% potency for one to two years after expiration, the exact degree of potency loss with time is uncertain [4]. Reduced effectiveness and perhaps harmful degradation products can result from instability [5]. According to stability studies, the expiration date guarantees that the product retains at least 90% of its original capacity if properly stored [4, 5]. The pharmaceutical sector can save considerable amounts of money annually by recycling obsolete drugs, which is both economically and environmentally beneficial [5]. However, using expired medications in emergency situations presents moral dilemmas regarding how to weigh such dangers against the necessity of medical care [6].

Conventional microbiological methods need specialised equipment, skilled workers, and a lot of time and money [7, 8]. These techniques frequently lack the speed and sensitivity required for making food safety decisions in real time [9]. Although portable biosensors and quick diagnostic technologies provide flexibility and convenience of use, further work is still needed to improve detection limits and lower prices [9]. Numerous techniques, each with unique advantages and disadvantages, have been used for wood product assessment, including optical examination, canine detection, and CT scans [10].

Researchers propose integrated strategies that combine multi-step detection systems for wood product inspection [10], improved detection techniques, and quantitative risk assessments for food safety in order to get around these restrictions [9].

1.1 Problem statement

Non-invasive, automated, real-time, and economical detection methods are required. Pharmacists and end users cannot scale or employ traditional approaches.

1.2 Propose approach

This study suggests using artificial intelligence (AI) in conjunction with picture analysis methods to find visual indicators linked to medication expiration. On pills, capsules, or packages, the emphasis is on a non-invasive, non-destructive detection framework that makes use of image-based features such colour variation, surface texture alterations, morphological distortions, and text degradation. These visual indicators are analysed using machine learning or deep learning models to determine whether medications are expired or not, without the use of chemicals or physical intervention.

1.3 Identified research gaps

Numerous techniques for recognizing expired drugs have been investigated in recent studies. In order to identify the expiration dates of medications [11], created a smartphone application that uses QR codes. A Raman spectroscopy-based method for identifying expired medications was presented, providing a quick and eco-friendly substitute for conventional techniques [12]. In order to determine the amount of epinephrine that is still available in expired and discoloured autoinjectors, [13] looked at smartphone image colorimetry. They discovered a good link between colour indices and epinephrine content. It is possible to distinguish between expired and unexpired antibiotic and antimalarial medications with high identification rates by utilising a pocket-sized NIR spectrometer in conjunction with multivariate algorithms [14]. Although this research offers encouraging methods, they still depend on pre-existing labelling systems or call for specialised equipment, suggesting that a non-invasive, low-cost, real-time method utilising readily available imaging instruments is still elusive.

There are currently no non-invasive, inexpensive, and real-time techniques for identifying expired medications using widely accessible imaging devices like webcams or smartphone cameras. Automated detection based on visual markers that may be readily observed on tablets or capsules, like colour changes or surface degradation, has not been extensively studied. Furthermore, the actual application of current techniques is limited due to their frequent inability to generalise well across a variety of pharmacological formulas, changing illumination situations, and packaging types.

2 Literature Survey

Numerous recent studies have investigated technological approaches for identifying expired drugs; nevertheless, the majority of these investigations encounter issues with scalability or practicality. A smartphone app that uses QR codes to detect expiration was created, however it needs intact digital labelling [12]. Although it requires costly equipment, a non-invasive drug analysis method was proposed based on Raman spectroscopy [13]. A relationship was established between colour degradation and efficacy by estimating the epinephrine potency in expired autoinjectors using smartphone-based colorimetry [14]. Multivariate models and a pocket-sized NIR spectrometer were used to show great accuracy in identifying expired medications [15]. These techniques, however, frequently rely on specific equipment or regulated environments. There is currently no method that uses commonly available equipment, such as webcam cameras or smartphones, to offer real-time, inexpensive, and non-invasive detection. This emphasizes the necessity of an AI-powered method that can recognize visual cues from drug photos, like colour shifts or surface deterioration.

3 Methodology

This study suggests a non-invasive machine learning system that combines textual metadata taken from packaging with image-based visual analysis to identify expired or deteriorated medications. The methodology as depicted in Fig.1, highlighting the integration of CNN and OCR components.

Fig.1 illustrates the end-to-end workflow of the proposed machine learning model for non-invasive detection of expired or degraded medicines. A CNN backbone is used to extract deep visual information from an input image of the medication strip or tablet at the start of the process. In parallel, OCR is used to process the text portion of the strip, specifically the parts that carry the manufacturing and expiration dates, in order to extract date information. An expiry flag is then used to combine these textual and visual elements to show whether the expiration date has passed. A self-attention merging layer receives this combination of structured and unstructured data, allowing the model to adaptively prioritize pertinent characteristics. To categorize the sample into one of three groups - Safe, Warning, or Expired, his hybrid pipeline uses both text and appearance metadata to improve prediction accuracy.

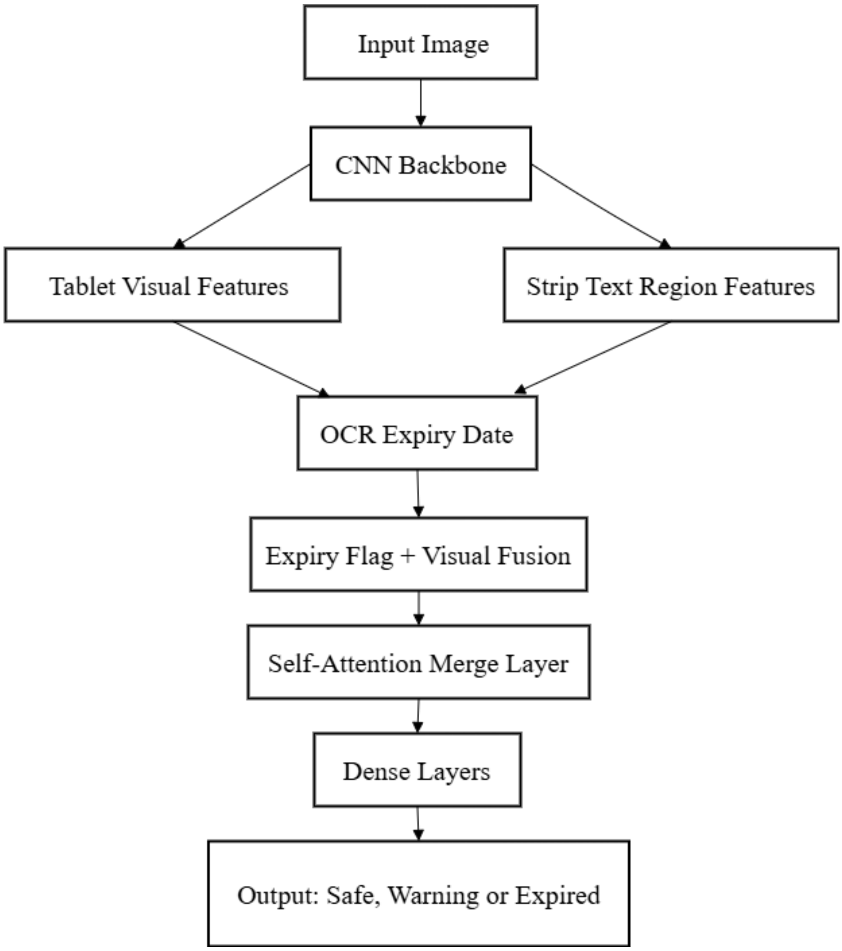

Fig. 1. Process of innovative machine learning model architecture workflow

3.1 Datasets

The "Medicine Tablet Pack Image Dataset" provided the images of tablet strips and capsules. These pictures show several labelled medication packs that have been arranged according to their expiration dates. Additional self-clicked examples were added to the dataset to test it in various deterioration scenarios and lighting conditions.

To evaluate the model's performance in the actual world, custom photos were taken and examined in addition to the dataset. These are labelled below for analysis and are displayed in Fig.2. Three real-world medicine images were analysed to test model accuracy. Fig. 2(a) (Calpol 500, 225×224) infer dull tablets and faded packaging, lacking the usual gloss. The model labelled it as Expired, consistent with its known expired status. Fig. 2(b) (Montair FX 1600-1600) shows a well-preserved silver strip with clear labelling and intact foil. The model is classified as Good and verified by the validity of the printed expiry date. Figure 2(c) (iron supplement, 9641599) shows surface coloration, yellowing of coatings and yellowing of edges. The model accurately predicts it as expired, allowing visual degradation to occur. These examples validate the model's ability to assess the quality of medicine using only visual indications, even if the expiry date is not clearly printed.

(a) Paracip-500 (b) Montair FX (c) Iron supplement

Fig. 2. Sample input images from the original data set (to be validated) and new real-world images of medical strips (to demonstrate the real-world performance). (a) Paracip-500 (944) 999, expiry March 2024; image captured May 2023. (b) Valid Montair FX (16001600), expiry July 20, 2025; photo taken July 1, 2025. This image is used to validate the model's ability to distinguish between valid and expired medication strips. c) Expiry date of the iron supplement (225224), July 20, 2025; image taken July 1, 2025. This image demonstrates the model's ability to detect expired medicines from real-world, self-clicked photographs, collected with controlled lighting and standardized angles to ensure consistency. These test images were validated by cross-referencing with the product's batch number and expiration date for reproducibility.

3.2 Feature extraction using CNN

A pre-trained ResNet-18 model was utilized to extract significant visual elements from medical photos. To obtain 512-dimensional deep feature vectors from the penultimate convolutional block, the fully connected (FC) layer was eliminated. Convolution in CNN is defined in mathematically as:

$$Y_{i,j}^{k} = \sigma\left(\sum_{m=1}^{M} \sum_{u=1}^{h} \sum_{v=1}^{w} X_{i+u,j+v}^{m} * W_{u,v}^{m,k} + b^{k} \right) \tag{1}$$

where $Y_{i,j}^{k}$ is the output feature at position (i,j) for the k^{th} filter, W is the kernel, b^k is the bias, and σ is the ReLU activation function.

3.3 OCR-based textual feature extraction

Using the EasyOCR library, OCR was used to retrieve embedded text from packaging, specifically batch numbers and expiration dates. An expiry flag of 1 was inserted if the expiry year was determined to be < 2024; otherwise, 0. This produced a combined feature vector of size 513 as shown in Eq. (2):

$$\mathbf{x} = [\mathbf{v}\,; f] \tag{2}$$

where $\mathbf{v} \in R^{512}$ represents CNN features and $f \in \{0,1\}$ denotes the expiry flag.

3.4 Classification using gaussian naive bayes

A Gaussian Naive Bayes(GNB) model was used to tackle the classification task. For a given class C_k, each feature dimension is assumed to have a normal distribution, and the posterior probability is computed as in Eq. (3):

$$P(C_k \mid \mathbf{x}) \propto P(C_k) \prod_{i=0}^{n} = \frac{1}{\sqrt{2\pi\sigma_{k,i}^{2}}} \exp\left[-\frac{(x_i - \mu_{k,i})2}{2\sigma_{k,i}^{2}} \right] \tag{3}$$

where $\mu_{k,i}$, $\sigma_{k,i}$: mean and standard deviation of feature i for class C_k, $\mathbf{x}$: input vector, and n=513: total features. The predicted class is determined using Eq. (4):

$$\hat{y} = \arg\max_k P(C_k \mid \mathbf{x}) \tag{4}$$

3.5 Evaluation metrics

To evaluate performance, standard classification metrics were used, as defined in Eq. (5), Eq. (6), Eq. (7), and Eq. (8):

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (5)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (6)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (7)$$

$$\text{F1-Score: F1} = 2 * \frac{\text{Precision}\cdot\text{Recall}}{\text{Precision+Recall}} \quad (8)$$

where TP, TN, FP, and FN are true or false positives or negatives.

3.6 Graph-based enhancement using GCN

To further improve categorization, a GCN was constructed to describe sample similarity. Individual medication samples were represented by nodes, and pairs in the feature space with Euclidean distances less than 1.0 were connected by edges. GCN propagates information using Eq. (9):

$$H^{(l+1)} = \sigma\,(\widetilde{D}^{-1/2} * \tilde{A}\,\widetilde{D}^{-1/2} H^{(l)} W^{(l)}) \quad (9)$$

where $\widetilde{A} = A+I$: adjacency matrix with self-loops, $\widetilde{D}$: degree matrix, $H^{(0)} = X$: input features, $W^{(l)}$: learnable weights, and σ: activation function. (e.g., ReLU).

The final output is generated via the log-softmax function as shown in Eq. (10):

$$\hat{y} = \log(\frac{\exp z_i}{\sum_j \exp z_j}) \quad (10)$$

Training is performed by minimizing the Negative Log-Likelihood Loss (NLL), as given in Eq. (11):

$$\mathcal{L}_{\text{NLL}} = -\sum_{i\in y} \log P(y_i|x_i) \quad (11)$$

Each sample was finally categorized into Good (label 0), Degraded (label 1) and Expired (label 2).

4 Results

The model was trained and assessed using a larger dataset of more than 500 medication strip photos that were categorized as Good, Degraded, and Expired. Using OCR-based expiry metadata and visual characteristics, the Naive Bayes classifier demonstrated its efficacy in non-invasive medication quality prediction with 0.91 accuracy, 0.91 precision, 0.89 recall, and a F1-score of 0.90.

The GNN node graph in Fig. 3 illustrates the GCN's capacity to represent semantic links between samples. In this case, each node stands for a sample of medication and is coloured according to its actual label: Red for expired, Orange for degraded, and Green for good. The circular clustering pattern shows how the GCN groups related data into dense local neighbourhoods whereas solitary central nodes might be a sign of low-confidence predictions or outliers that need human assessment. This makes it easier to comprehend the model in real-world situations when degradation may not always be readily apparent.

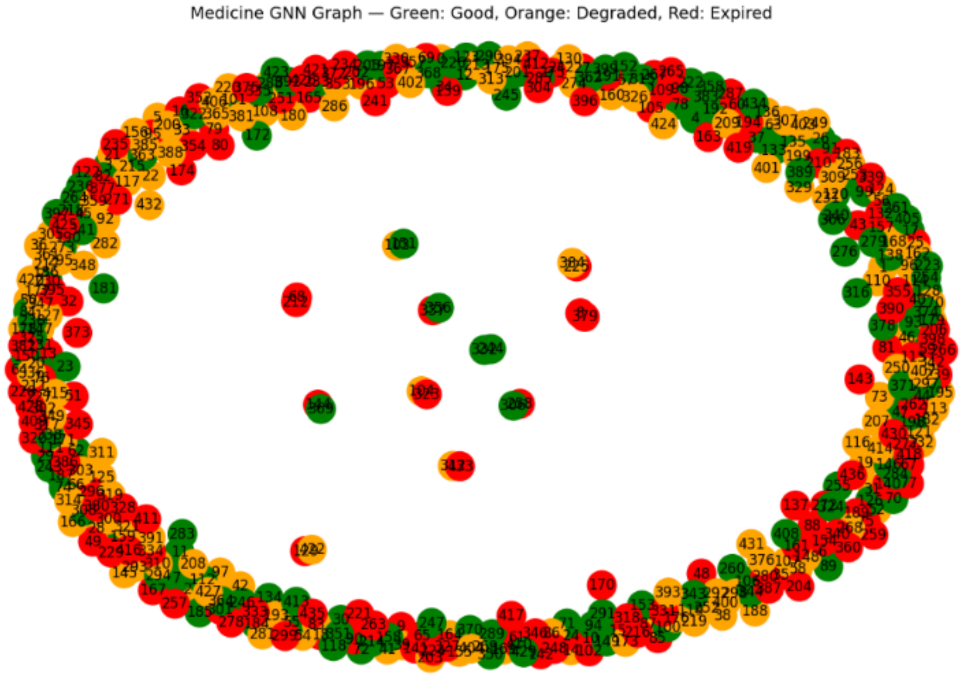

Fig. 3. GNN Node Graph of Medicine Samples

Fig.4 shows the distribution and interrelations among medicine classes: Good (green), Degraded (orange), and Expired (red). Each central node (labelled 132, 152, 153) represents the number of samples in its respective class, derived from the dataset used. These nodes are linked to the blue nodes around them, representing the individual sample clusters of each category. The red curves depict the semantic and visual characteristics of the different categories and represent transitional degradation (for example, from good to decay or decay to expired).

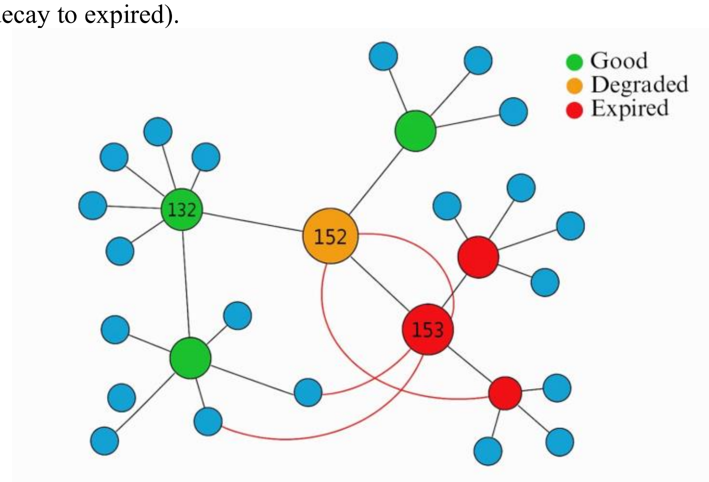

Fig. 4. Complex Network Graph Representation of Medicine Class Distribution

 https://doi.org/10.1051/bioconf/202520401021

As illustrated in Fig.5, t-SNE shows that the Good, Degraded, and Expired samples are partially clustered; overlaps indicate that combining textual and visual information is crucial for improved categorization.

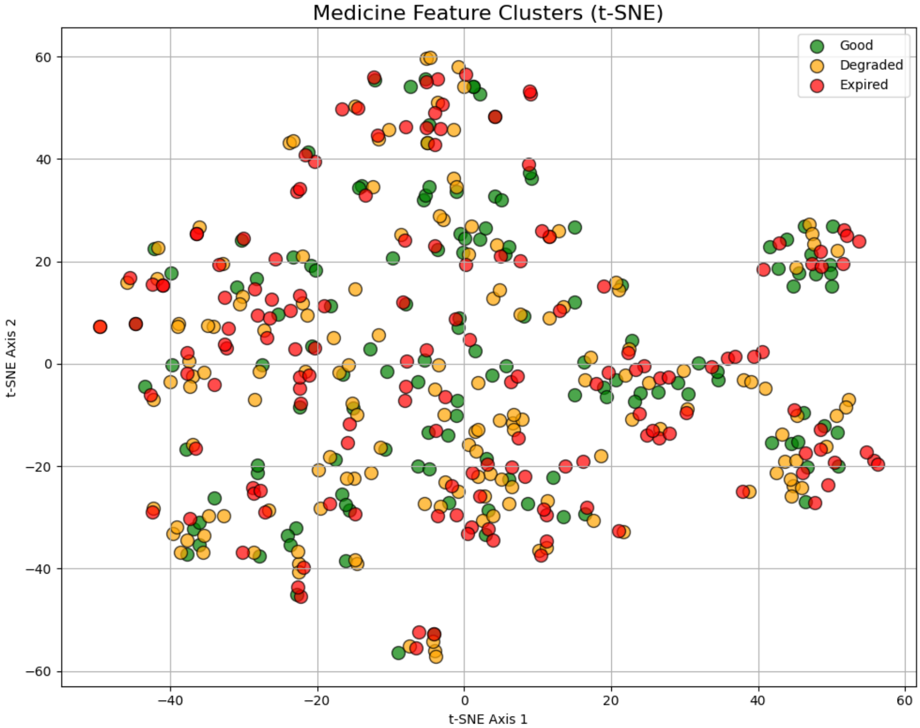

Fig. 5. t-SNE Cluster Plot

Overall, the results show that a reliable, comprehensible, and scalable system for identifying expired or deteriorated medication is created by combining CNN-based visual signals, OCR-derived textual metadata, and GNN-based structural learning.

5 Discussion

By combining CNN-based visual feature extraction, OCR-derived textual metadata, and GCN-based relational learning, the suggested architecture shows that a non-invasive, AI-driven method of expired medication identification is feasible. With an accuracy of 0.91, precision of 0.91, recall of 0.89, and an F1-score of 0.90, the model demonstrated that packaging photos include enough discriminative features to evaluate drug quality without manipulation or laboratory assistance.

GNB was chosen as the classifier because of its interpretability, robustness with high-dimensional data, and simplicity. GNB is appropriate for situations where it is challenging to gather big annotated pharmaceutical datasets because it performs effectively even with relatively limited datasets. Other well-known models like Support Vector Machines (SVM), Random Forests, and deep learning-based end-to-end classifiers, which may produce better results but at the expense of interpretability and computational efficiency, were not benchmarked against in this work. Future analyses would assist in determining the best balance between transparency, scalability, and accuracy.

6 Future Work

To increase generalizability, future research can concentrate on growing the dataset with a variety of medicinal goods under various packaging, illumination, and degradation conditions. The best performance-interpretability trade-offs can be found by comparing the existing Gaussian Naive Bayes model to more sophisticated classifiers like SVM, Random Forest, and deep learning-based architectures. Real-time detection in pharmacies and homes can be made possible by integration with mobile applications and inexpensive edge devices, especially in rural or resource-constrained environments. Furthermore, adding multimodal data such as chemical sensors, spectrum analysis, and blockchain-based supply chain metadata could improve resistance to changes and counterfeit drugs, and open up a more comprehensive and reliable pharmaceutical quality assurance system.

7 Conclusion

This research uses artificial intelligence to analyze packaging images and provide a non-invasive way of detecting altered or expired drugs. By combining deep learning techniques such as visual function extraction from CNN and expiry detection from OCR with GNN for relational learning, we provide a practical method of pharmaceutical quality management. The model correctly identified the three categories of good drugs, degraded drugs, and expired drugs. Self-click photos, real expiry patterns and GCN-Naive Bayes ensembles make the system scalable and readable. The effectiveness of automated drug quality verification has been demonstrated in this study and shows the potential for integration into drug and health care workflows, especially in situations where resources are limited or unavailable.

We would like to thank the Institute of Engineering and Management (IEM), Kolkata, and IEM Centre of Excellence for InnovAI, for their institutional support and invaluable infrastructure. The authors disclose that they have no sponsorship or conflicts of interest. Availability of Data: https://www.kaggle.com/datasets/nitesh31mishra/medicine-tablet-pack-image-dataset

References

1. M. Rajalakshmi, P. Datchanamourtty, P. Vasigar, M. Rajalakshmi, P. Vasigar, Insights into medicine expiry awareness among the population of rural South India: a mixed-methods approach. Cureus **16**, e53715 (2024). https://doi.org/10.7759/cureus.70314

2. P. Sivasankaran, E.B. Mohammed, N. Ganesan, R. Durai, Storage and safe disposal of unwanted/unused and expired medicines: a descriptive cross-sectional survey among Indian rural population. J. Young Pharm. **11**, 97 (2019). https://doi.org/10.5530/jyp.2019.11.20

3. M.R. Fernandes, R.C.D. Figueiredo, L.G.R.D. Silva, R.S. Rocha, A.O. Baldoni, Storage and disposal of expired medicines in home pharmacies: emerging public health problems. Einstein (São Paulo) **18**, eAO5066 (2020). https://doi.org/10.31744/einstein_journal/2020AO5066

4. G.S. Sarla, Efficacy and disposal of drugs after the expiry date. Egypt. J. Intern. Med. **31**, 431–434 (2019). https://doi.org/10.4103/ejim.ejim_110_19

5. S.C. Basha, K.R. Babu, M. Madhu, Y.P. Kumar, C. Gopinath, Recycling of drugs from expired drug products: comprehensive review. J. Glob. Trends Pharm. Sci. **6**, 2596–2599 (2015).
https://doi.org/10.3390/ecmc-1-A005

6. E.I. Ette, Conscience, the law, and donation of expired drugs. Ann. Pharmacother. **38**, 1310–1313 (2004).
https://doi.org/10.1345/aph.1D375

7. I.A. Quintela, T. Vasse, C.S. Lin, V.C. Wu, Advances, applications, and limitations of portable and rapid detection technologies for routinely encountered foodborne pathogens. Front. Microbiol. **13**, 1054782 (2022).
https://doi.org/10.3389/fmicb.2022.1054782

8. I.H. Cho, S. Ku, Current technical approaches for the early detection of foodborne pathogens: challenges and opportunities. Int. J. Mol. Sci. **18**, 2078 (2017).
https://doi.org/10.3390/ijms18102078

9. R.L. Buchanan, C.M. Deroever, Limits in assessing microbiological food safety. J. Food Prot. **56**, 725–729 (1993).
https://doi.org/10.4315/0362-028X-56.8.725

10. N. Mirza, B. Ganguly, Utilization of medicines available at home by general population of rural and urban set up of Western India. *J. Clin. Diagn. Res.* **10**, FC05 (2016).
https://doi.org/10.7860/JCDR/2016/20600.8298

11. H. Banjar, R. Alrowithi, S. Alhadrami, E. Magrabi, R. Munshi, M. Alrige, An intelligent system for proper management and disposal of unused and expired medications. *Int. J. Environ. Res. Public Health* **19**, 2875 (2022).
https://dio.org/10.3390/ijerph19052875

12. M. Ramalingam, R. Puviarasi, N.D.A.B. Zakaria, E. Chinnavan, Developing mobile application for medicine expiry date detection. Int. J. Pure Appl. Math. **119**, 3895–3901 (2018).

13. S. Zhang, H. Chen, R. Li, Z. Yu, F. Lu, Raman spectroscopy and mapping technique for the identification of expired drugs. Spectrochim. Acta A Mol. Biomol. Spectrosc. **224**, 117407 (2020).
https://doi.org/10.1016/j.saa.2019.117407

14. A. Saleheen, B.M. Campbell, R.A. Prosser, C.A. Baker, Estimation of available epinephrine dose in expired and discolored autoinjectors via quantitative smartphone imaging. Anal. Bioanal. Chem. **412**, 2785–2793 (2020).
https://dio.org/10.1007/s00216-020-02505-y

15. J. Opoku-Ansah, E.S. Yahaya, C.L.Y. Amuah, R. Nyorkeh, A. Adom-Konadu, P.O.W. Adueming, E. Teye, A feasibility study on the use of a pocket-sized NIR spectrometer and multivariate algorithm to distinguish expired drugs from unexpired ones. Anal. Methods **14**, 2405–2414 (2022).
https://dio.org/10.1039/d2ay00541g