

**Youa Raj Chettri**  
**Dheeraj Kumar Prasad**  
**Rahul Shah**

Review paper  
DOI – 10.24874/QF.25.132



## **COMPREHENSIVE REVIEW OF CONVOLUTIONAL NEURAL NETWORK (CNN) MODELS FOR DRUG DETECTION USING IMAGE PROCESSING**

**Abstract:** *AI and image processing have revolutionized the way that differentiating drugs occur in the health and pharmaceutical sectors. This review paper focuses on the use of AI, particularly deep learning technology, in the identification of drugs from images using a Convolutional Neural Network. these models can also identify pills, blister packs, and even chemical structures according to their forms, colors, and some other characteristics. We suggest how researchers and experts have applied AI to construct some systems that assist in fulfilling duties ranging from identifying falsified drugs, reading drug labels, estimating drug interactions, as well as arranging pharmacy stocks and shelves. Others include integrating the CNNs with conventional methods or approaches like SVM and the k-NN to boost the performance. The research examines different challenges related to medical imaging AI, including unclear drug appearances, together with small training data, and a requirement for transparency when AI systems operate. The paper discusses potential future applications, including developing quick mobile models and utilizing artificial intelligence with different forms of data for drug safety improvements. The review provides students, researchers, and developers understanding of contemporary drug identification methods using artificial intelligence in combination with image processing and upcoming advancements.*

**Keywords:** *Drug Detection, Image Processing, Artificial Intelligence (AI), Deep Learning, Convolutional Neural Networks (CNNs), Image Recognition*

### **1. Introduction**

Pharmaceuticals or drugs are substances used to treat diseases or illnesses. In other words, they can generally be categorized as prescription drugs, over-the-counter (OTC) drugs, controlled substances, herbal remedies, and illicit drugs. There are thousands of drugs in differing sizes,

packaging, and imprints, including active ingredients and inactive formulation; accurate identification is critical to patient safety, correct treatment, and regulatory compliance (Ting et al., 2020).

The general use of drugs and substances requires proper detection within the health care sector in order to avoid the deterioration of the effects of the drugs. Medication errors

---

<sup>1</sup> Corresponding author: Youa Raj Chettri  
Email: [yuvrazz2001@gmail.com](mailto:yuvrazz2001@gmail.com)

like wrong drug name, wrong dosage, failure to recognize fakes could lead to adverse health impacts, more hospitalizations, and may cause deaths (Rehman et al., 2024). By and large, medication errors are documented to rank high among potentially emergent adverse events in hospitals across the world. Also, the introduction of fake drugs in both developed and developing nations continues to have complexities in their health systems.

To tackle these problems, image-based drug detection came to the fore as the most effective solution. Using features such as the colour, shape, texture, and packaging labels of drugs, it is possible to come up with an automated process that is used in the identification or validation of the products. The main benefit of this approach has to do with the assurance of identification of drugs, particularly in pharmacies, hospitals, or customs and remote healthcare facilities that require a physical examination of drugs to be certain regarding the substances they contain.

CNNs, at the base of the image-based drug detection, are a unique kind of deep learning model aimed at recognizing and analyzing images. CNNs are good at recognizing complex patterns in images since they do not require manual inputting of the features to look for (Kourounis et al., 2023). These include but are not limited to medical imaging, face recognition, and self-driving, and have placed them in a position to adapt and apply in drug detection.

This review paper is focused on reviewing CNN-based models regarding their usage in drug detection involving image processing. We explore various application areas, commonly used CNN architectures, available datasets, evaluation metrics, and current research trends. Furthermore, the review highlights the limitations of existing approaches and outlines potential directions for future work to enhance accuracy, robustness, and real-world applicability (Li et al., 2024).

### **1.1. Drug Detection and Why It Is Important**

The phrase drug detection means the identification of the appropriate medication to administer to the patient. This is important in order not to administer the wrong tablet, wrong quantity, or some non-genuine drugs. Some reproductions of drugs are dangerous, and this issue of people's irresponsible use of the substances is alarming. Proper detection provides ways of preventing all these problems with emphasis on safety (Sakib & Chowdhury, 2024).

### **1.2. Image Processing Helps in Drug Detection**

Image processing enables the system to recognize the images of the pills, packaging, or structure of chemical compounds. It helps in identifying the shape, color, and markings on the drug, facilitating the ease in identifying the medicine that pharmacists or medical practitioners may prescribe.

### **1.3. CNNs and Why They Are Useful**

CNNs refer to systems that assist computer systems in interpreting images or pictures, in this case. There are two types of architectural configurations that are found in a CNN, and these are filters and layers. The filters, in turn, look for the simplest features, such as edges or lines in the picture. These layers aid the CNN to learn more and more details of the image as it goes through the layers. They can learn from an example: while they are able to distinguish shapes, for instance, or different textures. It turns out that CNNs are excellent when it comes to distinguishing between very similar objects, such as pills and the type of drugs in this case (Zhang et al., 2022).

#### 1.4. CNNs Are Better Than Other Methods

In comparison to the traditional methods, CNNs possess several advantages that make them more suitable for use in the analysis of data.

- Learn automatically from images.
- It is also easy to notice the smallest difference in terms of the color of the pills and the imprint on them.
- Provide better work done in less time, as well as in noisy and blurred situations.

This makes CNNs suitable for use in the detection of drugs (Asfand-e-yar et al., 2024).

**Table 1.** Comparison of Different Drug Detection Methods

| Feature                      | Learning        | Speed    | Accuracy         |
|------------------------------|-----------------|----------|------------------|
| CNNs                         | Automatic       | Fast     | High             |
| Traditional Image Processing | Manual          | Slow     | Low to Moderate  |
| Machine Learning             | Low to Moderate | Moderate | Moderate to High |
| Spectroscopy                 | Does not learn  | Fast     | High             |
| Chromatography               | Does not learn  | Slow     | High             |

## 2. Related Work

The following are the investigations of the different scholarly articles on drug detection using image processing. The findings of drug research achieved in this study are summarized in Table 2 below.

**Md. Nazmul Sakib and Mahfuzulhoq Chowdhury (2024).** This paper presents a machine learning system built with Faster R-CNN and YOLOv5 detection, along with Tesseract OCR text extraction, which detected medicine boxes on pharmacy shelves effectively and proved Faster R-CNN better than YOLOv5 for pharmaceutical inventory management.

**Yue Li et al. (2024).** The authors argue that image-based molecular representation learning serves as an innovative approach for creating drugs that surpasses NLP and GNN systems through visual data that better expresses pharmaceutical features to enable AI-based drug discovery.

**Ashfaq Ur Rehman et al. (2024).** This paper demonstrates how artificial intelligence, particularly machine learning models including deep learning, random forest, and support vector machines, enhances drug discovery by enabling quicker and less expensive drug development through disease identification and virtual screening platforms, and toxicity prediction and drug repositioning applications.

**Muhammad Asfand-e-yar et al. (2024).** The authors applied Multimodal Convolutional Neural Network (MCNN-DDI), proposed as a method to forecast drug-to-drug interactions through chemical structures, along with enzyme pathways and targets, reaching 90% accuracy with 94.78% AU, PR better than conventional DDI prediction approaches.

**Georgios Kourounis et al. (2023).** The authors of this paper emphasize that convolutional neural networks (CNNs) drive medical image analysis progress because they automate assessments of medical pictures across radiology, histopathology, and surgery, which leads to improved healthcare services.

**Khalil Al-Hussaeni et al. (2023).** In this paper, researchers built a CNN-based pill recognition system using CNN+SVM, CNN+kNN, and ResNet-50 models, which demonstrated that the CNN+kNN architecture reached 90.8% accuracy above all other existing pill image retrieval systems.

**Chengcheng Zhang et al. (2022).** In this paper, the researchers demonstrated that CNN-DDI predicted drug-drug interactions by generating multiple drug feature combinations, including categories, targets, pathways, and enzymes, which yielded

superior DDI prediction results than single features performed alone.

**Hsien-Wei Ting et al. (2020).** In this paper, the goal was to develop a deep learned drug identification model using YOLO that would accurately recognize blister packaged drugs in medicine units, and with an F1 score of more than 95%, it is a more effective tool for the prevention of an error medication in healthcare.

**Laxmi R. Deshpande and Praveen S. Challagidad (2019).** All the authors in this paper have claimed that an image processing-based application was developed to identify the drugs from the images of

drugs with an aim of providing drug information such as usage, side effects, and expiry details to ensure patients' awareness and reduce medication errors.

**Ranjitha et al. (2018).** In this paper, other authors have noted that a system of color and shape recognition of the pills using Raspberry Pi and OpenCV was established in an effort to assist in radical identification of pills in hospitals, noting that accurate identification of drugs is crucial to prevent instances whereby elderly patients end up taking the wrong drugs.

**Table 2.** Major contributions of research in the field of drug detection.

| Authors (Year)             | Title  | Method                    | Outcome                         | Limitation                 |
|----------------------------|--|---------------------------|---------------------------------|----------------------------|
| Sakib & Chowdhury (2024)   | A Machine Learning-Based Approach for the Detection of Drugs in Drug Self  | Faster R-CNN, YOLOv5, OCR | Accurate medicine detection     | Real-world testing missing |
| Yue Li et al. (2024)       | Image-Based Molecular Representation Learning for Drug Development   | Image-based Learning      | Better molecular representation | Few practical examples     |
| Rehman et al. (2024)       | Role of Artificial Intelligence in Revolutionizing Drug Discovery  | AI, ML, DL                | Faster drug discovery           | Data & ethics challenges   |
| Asfand-e-yar et al. (2024) | Multimodal CNN-DDI: Using Multimodal CNN for Drug to Drug Interaction Associated Events                          | Multimodal CNN            | 90% DDI prediction accuracy     | Limited features used      |
| Kourounis et al. (2023)    | Computer Image Analysis with Artificial Intelligence: A Practical Introduction to CNNs for Medical Professionals | CNNs                      | Helped medical image analysis   | Integration issues         |

|                                |  |                        |                          |                          |
|--------------------------------|--|------------------------|--------------------------|--------------------------|
| Al-Hussaeni et al. (2023)      | CNN-Based Pill Image Recognition for Retrieval Systems   | CNN+SVM, CNN+kNN       | 90.8% retrieval accuracy | Small dataset            |
| Zhang et al. (2022)            | CNN-DDI: A Learning-Based Method for Predicting Drug–Drug Interactions Using Convolution Neural Networks         | CNN-DDI                | Improved DDI prediction  | No unstructured data     |
| Ting et al. (2020)             | A Drug Identification Model Developed Using Deep Learning Technologies: Experience of a Medical Center in Taiwan | YOLO                   | 95.99% identification F1 | Limited hospital testing |
| Deshpande & Challagidad (2019) | Automatic Drug Identification Using Image Processing   | Image Processing       | Basic tablet info system | No AI predictions        |
| Ranjitha et al. (2018)         | Color and Shape Recognition of Pills Using Image Processing  | Color & Shape (OpenCV) | Low-cost recognition     | Not scalable             |

### 3. CNN Models Used in Drug Detection

CNNs are of many types, but the most commonly used types are the flat the deep, and wide convolutional neural networks. This has created a shift over time to the newer and better models in an attempt to enhance how machines will be able to recognize images of the drugs. Here's how they are grouped:

#### 3.1. Early CNN Models

Models that were first used in the air to pave the way for CNN usage in image recognition

- **LeNet**, A CNN that was created in the 1990s; hence, it is among the earliest CNNs that were devised. It

was employed for digit recognition that is similar to a ZIP code, and then it illustrated that CNNs are effective for image tasks.

- **AlexNet**: A breakthrough model in 2012. This revealed that DL was capable of achieving much better results compared to previous approaches in the process of image recognition.
- **VGG**: This model had 16 or 19 layers in total, and it inherited the ability to use small filters, 3x3 for superior image recognition. It was helpful in augmenting precision while demanding more computations than its predecessors.

Establishing these models created fundamental principles for imaging-based drug detection systems, especially for research and basic applications.

### 3.2. More Advanced CNN Models

- **ResNet (Residual Networks)**, Skip connections became a key feature of ResNet, creating the ability for models to achieve deeper learning without gradient vanishing issues. Through this mechanism, the model acquires the ability to detect detailed structures of medication along with their surface characteristics.
- **Inception**, the model's layer operates with various-sized filters simultaneously to discover features of different dimensions. Skip connections in this model serve as an essential tool for detecting small markings and different packaging styles present in drugs.
- **DenseNet**, Each layer of the structure links to every other layer, thus enabling easier reuse of features. The use of this method leads to precise forecasting while needing reduced parameters.

The advanced drug detection systems deploy high-performance models because of their superior capabilities.

### 3.3. Lightweight CNN Models

- **MobileNet**, designed for mobile devices and apps. It is small and fast, so it can run on smartphones. This makes it useful in mobile health apps or quick pharmacy tools.

Even with limited hardware, MobileNet can identify drugs with good accuracy, making it great for field use or remote areas.

### 3.4. Newer and High-Performance Models

- **EfficientNet**, the model supports trading off between size, speed, and accuracy performance. The system requires minimal resources to produce high-quality outcomes. This method demonstrates the capability for drug detection functionality in cloud computing platforms, together with mobile systems.
- **Vision Transformers (ViT)**: A variant model that operates without standard CNN layers utilizes attention to conduct whole-image exploration. The model delivers impressive performance and achieves sophisticated image tasks, which include pill difference detection.

These contemporary techniques function as leading-edge methods for studying as well as implementing drug detection through images.

## 4. Preparing the Image Data

The preparation of image data stands as a vital requirement during the training process of any CNN model. The data preparation phase enables the system to absorb correct features effectively. First, the images are cleaned. The process of cleaning involves fixing issues related to poor lighting and blurry sections, and unwanted background elements to enhance the clarity of significant features (such as pill shapes and markings). Image processing techniques follow these steps, after which the model proceeds:

- **Resizing**, The model requires uniform image dimensions of 224 pixels by 224 pixels through the resizing technique.
- **Cropping**, the essential image section that contains pills or tablets, remains while all other areas are

truncated.

- **Flipping and Rotating**, Flipping the images horizontally or vertically, along with rotating them, produces additional versions of the same image. Data augmentation is a technique that helps models acquire better data and avoid overfitting.

To avoid such time-consuming tasks and enhance efficiency, pre-trained models are employed. These ones are cognizant of more generalized image features from tens of thousands of images (like ImageNet). In using them, we do not need to go through the process of building a model from scratch, but just adapt them to be used for the purpose of detecting drugs, which is much faster and more efficient.

## 5. Measuring Performance

After training a CNN model, there is a need to evaluate how effective it is in solving the problem. This step is beneficial to ensure that the model is spot on in terms of identifying the drugs from the images obtained. There are a number of ways of defining performance, which are as follows:

- **Accuracy**, It tells you how many total forecasts were right and also how many per cent of tips were correct to an accuracy of the nearest tenth of one percentage point. It is plain and easy for intervention though sometimes it might not be sufficient if the provided data set is unproportional. For instance, some classes of the drugs may contain many images while others have very few.
- **Precision**, it demonstrates the actual number of the drugs that, according to the model, are indicated for the particular disease and should be contained in the list of reference drugs. Less false alarms mean that a high degree of accuracy has been achieved.

- **Recall**, it should inform us how many of the actual drugs were detected by the model. High recall means the possibility of the model identifying the wrong drugs is low because it failed to identify a small set of correct ones.
- **F1-Score**, this is the aspect of accuracy and completeness of results. It is advantageous, especially when we want both precision and recall values to be high. The higher the F1-score, the better; it means that there is a generally good level of completed tasks.

Here, it is possible to get more than simply an answer, yes or no, we can know how and why the model is right, where a problem might lie, etc.

## 6. Tools and Datasets

To construct a drug detection system using image processing, it is imperative to have appropriate software that is effective in analyzing images, and sample images that are ideal for the detection of such substances.

- **Software Tools**, Among deep learning frameworks, TensorFlow and PyTorch hold the status of being the most commonly used libraries. These tools facilitate the development of CNN models through their capability to design them and perform training and testing procedures efficiently. The toolkit contains a set of ready-made features that make it easy to develop networks and manage images and pre-trained models.
- **Datasets**, A model training process needs a substantial amount of drug image samples. Most commonly used datasets contain photographs of pills alongside their packaging types, as well as capsules and tablets for training. The various

types of drug images are presented in the following **Figure 1**. Two common sources of public data include the Pill Image Recognition Challenge Dataset, in addition to hospital-generated information databases. A quality dataset should contain images that are well-defined, in addition to presenting drugs from various vantage points with a diverse range of substances.

- **AI Apps for Pill Recognition**, The mobile applications MedSnap, combined with Pill Identifier, together with ID My Pill, employ Artificial Intelligence to identify pills through picture recognition. The app technologies enable health practitioners, along with patients, to evaluate medications through photography and then verify treatment safety. Powerful tools, together with good data sets, combined with advanced apps, enable the creation of dependable drug detection systems with quick results and precise identification.



**Figure 1.** Different types of drug images

## 7. Challenges

Despite the possibilities provided by CNNs, several issues are still relevant to drug detection using images:

- **Drugs that Look Very Similar**, Due to many pills and tablets' inclination to small, round, and similar colored forms, many issues can occur during their packaging. Depending on the size and hue, such variations as, for instance, thin lines or shades can hardly be noticed by the model or even the human eye.
  - **Poor Image Quality**, Someone may take the picture with a low resolution that blurs images or illuminates the image too much, or takes the image at the wrong angle; this way, the model might not identify the drug correctly. Both the exposure and sharpness of the images are the most important factors that influence the results.
  - **Not Enough Image Data**, Sometimes, the required pictures of some rare drugs may not be available enough to feed into the model. This means that any small data that is taken from the model will result in it being less accurate, and thus its efficiency decreases as it makes more mistakes.
- Hard to Understand AI Decisions**, Usually, such networks are called black boxes, because it is very challenging to explain why the CNN came to make a specific forecast. This lack of explainability turns into a challenge in such time-sensitive sectors as healthcare, where the decisions made have to be trusted and ensured.

## 8. Future Possibilities

Consequently, the use of Artificial Intelligence in the detection of drugs has a

very bright future in the near future, with many improvements lined up as follows:

- **Explainable AI (XAI)**, There are current efforts to understand and enhance the explicability of AI decisions. This would be especially useful for doctors and pharmacists to know why the AI has identified a pill as a particular drug, which will increase its reliability.
- **Privacy-Preserving Training**, In the future, the training of these models can take place locally, hence no health details of the user will be passed to the cloud. This contributes to the privacy of patients' information, hence enhancing the security.
- **Real-Time Detection**, Newer and lighter AI models will enable one to identify drugs through the use of a smartphone or a small handheld device. This should be quite useful in rural settings, pharmacies, or during emergencies.

- **Combining Images with Other Information**, Someday, the systems may not only rely on the pictures of pills but also chemical structures, bar codes, or prescriptions. This would enhance the dependability of the process of identification of drugs by leadership.

## 9. Conclusion

This paper demonstrated that CNNs can be highly effective when used for drug detection from image processing. To be more precise, by processing the image data, with the help of numerous models, various drugs can be recognized in a short time. Pros involve solutions in the future, such as explainable AI and mobile real-time detection; however, challenges of similar appearance drugs, poor quality images, and lack of data still persist. In this way, it will be possible to make the identification of drugs safer, faster, and more effective in real-life healthcare, pharmacy, and emergency applications.

## References:

- Al-Hussaeni, K., Karamitsos, I., Adewumi, E., & Amawi, R. M. (2023). CNN-based pill image recognition for retrieval systems. *Applied Sciences*, *13*(5050). <https://doi.org/10.3390/app13085050>
- Asfand-e-yar, M., Hashir, Q., Shah, A. A., Malik, H. A. M., Alourani, A., & Khalil, W. (2024). Multimodal CNN-DDI: Using multimodal CNN for drug-to-drug interaction associated events. *Scientific Reports*, *14*(4076). <https://doi.org/10.1038/s41598-024-54409-x>
- Deshpande, L. R., & Challagidad, P. S. (2019). Automatic drug identification using image processing. *International Journal of Engineering Research & Technology (IJERT)*, *8*(15).
- Kourounis, G., Elmahmudi, A. A., Thomson, B., Hunter, J., Ugail, H., & Wilson, C. (2023). Computer image analysis with artificial intelligence: A practical introduction to convolutional neural networks for medical professionals. *Postgraduate Medical Journal*, *99*(1178), 1287–1294. <https://doi.org/10.1093/postmj/qgad095>
- Li, Y., Liu, B., Deng, J., Guo, Y., & Du, H. (2024). Image-based molecular representation learning for drug development: A survey. *Briefings in Bioinformatics*, *25*(4), bbae294. <https://doi.org/10.1093/bib/bbae294>

- Ranjitha, Kulkarni, S. S., Rashmi, K. H., & Athmashree, A. T. (2018). Color and shape recognition of pills using image processing. *International Journal of Engineering Research & Technology (IJERT)*, 7(8).
- Rehman, A. U., Li, M., Wu, B., Ali, Y., Rasheed, S., Shaheen, S., Liu, X., Luo, R., & Zhang, J. (2024). Role of artificial intelligence in revolutionizing drug discovery. *Fundamental Research*. <https://doi.org/10.1016/j.fmre.2024.04.021>
- Sakib, M. N., & Chowdhury, M. (2024). A machine learning-based approach for the detection of drugs in drug self. *Artificial Intelligence and Applications*. <https://doi.org/10.47852/bonviewAIA42022666>
- Ting, H.-W., Chung, S.-L., Chen, C.-F., Chiu, H.-Y., & Hsieh, Y.-W. (2020). A drug identification model developed using deep learning technologies: Experience of a medical center in Taiwan. *BMC Health Services Research*, 20(312). <https://doi.org/10.1186/s12913-020-05166-w>
- Zhang, C., Lu, Y., & Zang, T. (2022). CNN-DDI: A learning-based method for predicting drug–drug interactions using convolutional neural networks. *BMC Bioinformatics*, 23(Suppl 1), 88. <https://doi.org/10.1186/s12859-022-04612-2>

---

**Youa Raj Chettri**

MCA, School of IT,  
The ICFAI University, Sikkim,  
Gangtok, India,  
[yuvrazz2001@gmail.com](mailto:yuvrazz2001@gmail.com)  
ORCID 0009-0003-8540-3756

**Dheeraj Kumar Prasad**

MCA, School of IT,  
The ICFAI University, Sikkim,  
Gangtok, India,  
[dk384414@gmail.com](mailto:dk384414@gmail.com)  
ORCID 0009-0001-4296-4698

**Rahul Shah**

Assistant Professor,  
The ICFAI University, Sikkim,  
Gangtok, India,  
[rahul.shah@iusikkim.edu.in](mailto:rahul.shah@iusikkim.edu.in)  
ORCID 0000-0002-0605-5681

---