Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3 (2025)

An Efficient Counterfeit Medicine Classification Forecasting System: A Structure based Deep Learning Technique

Binitha S Thomson[1] , Dr. W. Rose Varuna[2]

[1]Ph.D. Research Scholar, Department of Information Technology, Bharathiar University, Coimbatore-46., bini2796tha@gmail.com

[2]Assistant Professor, Department of Information Technology, Bharathiar University, Coimbatore-46.

rosevaruna@buc.edu.in

Article History:

Received: 12-01-2025

Revised: 15-02-2025

Accepted: 01-03-2025

Abstract:

Identification of fake medicine images with chemical structure by using different compounds or molecular compositions that are different from the legitimate products. Thus, by using better computational approaches to look into structural features can help in the detection of counterfeit medicine. The proposed work in this paper enhances the ability to distinguish between the original and fake product that is crucial for product legitimacy identification and consumer health especially with industries dealing with pharmaceutical products and construction materials while public safety demands counterfeit medicine identification. The current techniques prove to be inefficient and the precise outcomes can be achieved by using complex calculations derived from the chemical structures. The aim of this study is to develop an efficient system based on Graph Neural Networks (GNN)'s to classify and predict counterfeit medicines for addressing the global counterfeit medicines issue. This paper proposes the classification system of counterfeit medicines based on the chemical structure and its forecast is assisted by the Deep Learning method known as GNN. The given methodology incorporates pre-processing steps which enhance structural characteristics of chemical compounds. The edge detection algorithms such as the Canny edge detector emphasize the prominent structural features. The morphological operation, dilation, and erosion are used for improvement of these features. The proposed chemical structure-based counterfeit medicine image detection using Canny Edge Detector with Graph Neural Networks (CED-GNN) is found to be better than the existing techniques with a maximum accuracy of 81.91%.

Keywords: Counterfeit Medicine, Graph Neural Network, Edge Detection, Chemical Structure, and Morphological Feature.

Introduction

Counterfeit medicine image detection has emerged as an important problem in many fields such as e-commerce, social networks, and digital forensic [1]. The availability of better image editing tools and Generative Adversarial Networks (GAN) has facilitated the creation of fake images, which in turn has caused misrepresentation, loss of money, and privacy violations. Thus, detection mechanisms must be reliable and capable of protecting against these threats. Tasks such as node classification, link prediction, and graph classification are very efficient when solved by Deep Learning (DL) models [2].

Counterfeit medicine are identified based on the chemical composition of the said drugs. Some of the methods that may be employed in this process include spectroscopy, chromatography and mass spectrometry [3]. Spectroscopy is a method of studying the effect of light on chemical substances and depends on the absorption spectra to infer the molecular structure of the compound. Chromatography enables one to distinguish between chemical compounds in a mixture by separating them. Mass spectrometry quantifies and qualitatively determines the mass to charge ratio of the ions.

These methods help in giving a true account of the chemical make-up of a medicine and can be used to compare with a standard sample of the same medicine [4]. Even the smallest of variations in the chemical make-up or molecular structure of the materials is a sign of fakes. This comparative analysis assists in weeding out fake drugs in the market hence protecting the consumers from the effects of counterfeit medicine. Advanced statistical tools are very useful in identifying counterfeit medicine in the market [5].

Counterfeit medicine-based image prediction using chemical structure is based on the chemical structures of images to predict counterfeit medicine images. This method engages the structure of the chemical compound in question and assesses the image attributes including atomic arrangement, spatial disposition, and bonding [6]. The proposed method of the deep learning can enhance the levels of accuracy and performance of the counterfeit medicine image prediction systems and can show high values for the use in the pharmaceutical and scientific research [7].

Chemical structures in the sense that the structures are unique and different assist in improving deep learning algorithms for fake drugs detection [8]. It can be seen that Deep Learning (DL) algorithms is more suitable for large and complex data such as spectroscopic scans. These algorithms are trained with large spectroscopic data of the genuine and counterfeit medicines and thus are able to learn different patterns or characteristics of chemical structures. Once trained, DL algorithms are capable of fast and accurate classification of new samples for the identification of defects of desired chemical composition [9].

In contrast to the previously described monitoring techniques, this one is quite helpful as deep learning models are able to process large amount of data and detect changes that can be unnoticed by operators or other types of algorithms [10]. Thus, the application of deep learning in large scale of counterfeit medicines identification is a very efficient solution to fight against fake and potentially unsafe drugs. Perhaps, it will be most appropriate to state that DL models are the most involved in the identification of counterfeit medicine. Such models' ability to process large datasets and make precise predictions ensures that even miniscule differences in the concentration of chemicals are identified. This capability is very vital especially in pharma industry where the chemical compounds that are consumed have to be pure for the benefit of the consumer. Therefore, with the help of advanced techniques like the use of DL algorithms, scholars are capable of developing effective systems that not only detect the fake drugs but also prevent their distribution [11].

The counterfeit medicine image detection, especially in the pharmaceutical one, is greatly improve by the means of the deep learning techniques [12]. Analyzing the chemical characteristics of drugs with the help of spectroscopy, chromatography and mass spectrometry, as well as using such deep learning models, it is possible to detect most of the fake drugs and prevent their circulation. It promotes the

effectiveness of the counterfeit medicine detection systems making the products safer and healthier for the consumers. The incorporation of these sophisticated approaches into the counterfeit medicine identification systems is a plus in the fight against fake and potentially dangerous goods [13].

Another problem of the counterfeit medicine image detection is the dynamic development of the methods of image manipulation [14]. Since novel techniques of synthesizing realistic fake images are created, the detection algorithms should learn from these innovations. This requires the constant updating and training of Graph Neural Network (GNN) models on new and current data sets of different fields. To accelerate this process and improve the detection accuracy, deep learning, a technique where the pre-trained models are fine-tuned for specific tasks, can be used.

DL based on structural approach to counterfeit medicine image detection make use of the intrinsic structural properties of graphs for distinguishing between authentic images and their forgery counterparts. This technique concentrates on the features at node-levels as well as those of a graph which are usually tampered by counterfeiters during the creation of fake images. This paper aims at presenting the practical steps that are taken when implementing structure-based counterfeit medicine image detection systems. First, there is always a preprocessing stage that involves constructing a graph for the input images, normalization, and augmentation to enhance the model's resilience.

The processed graphs are then given to the GNN that produces a prediction of the probability of the image being fake. Thresholding and probabilistic inference can be applied to improve the results in the post-processing step. The incorporation of GNNs in structure-based image detection presents a solution, to combatting the increasing issue of image manipulation. By utilizing GNNs capabilities to detect abnormalities these systems can achieve high accuracy and adaptability. Research, in learning and the use of AI methods are improving counterfeit medicine image detection methods making them more effective and transparent.

The paper further followed as: the overview of chemical structure based counterfeit medicine prediction is detailed in Section 1, the related works with comprehensive analysis is depicted in Section 2, proposed methodology on Canny Edge Detector with Graph Neural Networks (CED-GNN) illustrated in Section 3, the outcome of Canny Edge Detector with Graph Neural Networks (CED-GNN) is compared with existing technique and illustrated in Section 4, and the research is concluded in Section 5.

Related Work

Some studies in this field have looked at methods of detecting counterfeit medicine such as near-infrared spectroscopy with Siamese networks, deep learning models for image-based drug identification, and chemometric methods with chromatographic fingerprints. Some issues arise in dealing with the new samples, the size of the model, and the generalization of the results with different types of drugs and diseases. The diversified studies of existing techniques are given in Table 1.

Table 1. Comprehensive Analysis of Related Works

Reference	Author Name and Year	Inference	Methods Used	Significance	Drawback	Research Gap

[15]	Gayialis et al. (2022)	Traditional traceability methods are ineffective against modern counterfeiting	Structured literature review, classification framework	Identified trends and best practices, highlighted blockchain and IoT as promising technologies	Traditional methods are easily falsified	Need for more advanced and secure traceability solutions
[16]	An-Bing et al. (2020)	Siamese-network with NIR spectroscopy effectively identifies counterfeit drugs	Near-infrared spectroscopy, Siamese-network modeling, 1D-CNN	High accuracy in identifying counterfeit drugs, efficient even with limited samples	Not effective for unknown samples beyond the scope	Improvement in handling unknown and unbalanced samples
[17]	Ting et al. (2020)	Deep learning-based model for drug identification is highly accurate	Deep learning (YOLO), image analysis	High F1 score, effective in preventing LASA errors	Requires high-quality images, longer training times for back-side models	Further improvement in model efficiency and real-time integration
[18]	Simonovsky & Komodakis (2018)	GraphVAE for small graph generation offers potential in molecule generation	Variational autoencoder, graph generation	Advances generative models for graphs, potential in drug discovery	Challenges in scalability and complexity of generated graphs	Enhancing scalability and application to larger datasets
[19]	Kakio et al. (2017)	Combined handheld Raman spectroscopy and X-ray CT effectively discriminate falsified medicines	Handheld Raman spectroscopy, X-ray CT	Nondestructive, accurate discrimination of falsified medicines	Limited to specific types of pharmaceuticals	Broader application across diverse pharmaceutical products
[20]	Deconinck et al. (2012)	Chemometrics and chromatographic fingerprints	Chromatographic fingerprints, chemometric techniques	High correct classification rates for genuine and	Limited to specific drug types, requires extensive	Expansion to other drug categories and

		are effective for counterfeit medicine detection		counterfeit samples	sample preparation	simplified sample preparation

The research gap focused in this article is in relation to several weaknesses observed in earlier studies. Unknown and unbalanced samples are also well handled by GNNs as pointed out by An-Bing et al. (2020) [16]. This structure-based approach holds across various drugs, minimizing the vast sample preparation mentioned by Deconinck et al. (2012) [20]. Moreover, GNNs solve the scalability and complexity problems in the generation of graphs, as mentioned by Simonovsky and Komodakis (2018) [18], as well as the versatility allows for the use of the model in different pharmaceuticals, which responds to the broader application needs discussed by Kakio et al. (2017) [19].

Proposed Methodology- A Structure based Deep Learning Technique

This research framework employs GNNs to classify and estimate the fake drugs depending on the structural features of chemical compounds. The steps comprise data collection whereby a dataset of the chemical compounds, both real and counterfeit, is compiled. They are depicted in structures including molecular graphs for each compound of the compounds list. The proposed framework involves the use of the Canny edge detector to enhance the structural parts, morphological processes such as dilation, and erosion to enhance and to define these parts. The processed data is then converted into graph representations; nodes are atoms and edges are bonds, the graph is then passed through a Graph Neural Network (GNN) for classification. The overall research work is given in Figure 1 and the original Famotidine structure is diverse angle as well as three-dimensional view is given as input.

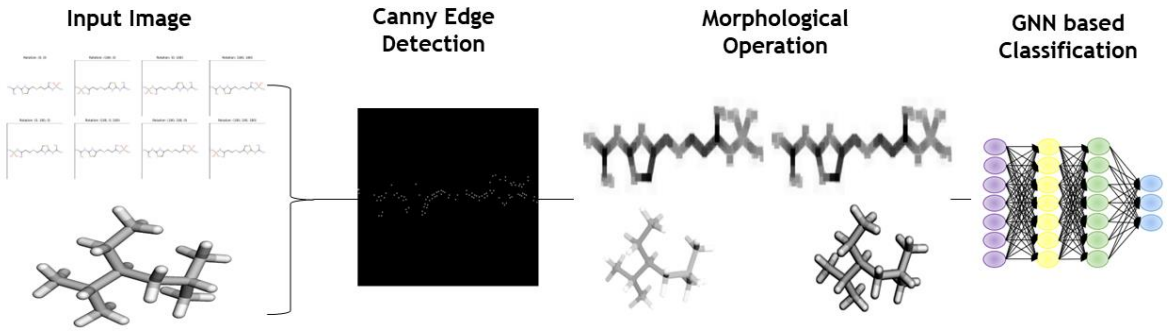

Figure 1. Methodology of CED-GNN

Edge Detection

Edge detection is performed by applying Gaussian smoothing to filter out noise, and then Sobel operators to determine strength and orientation of the edges. Non-maximum suppression enhances the edges by eliminating non-maximum pixels while hysteresis edge tracking utilizes high and low thresholds to confirm the edges' identification.

Noise Reduction

The Gaussian filter is applied to smooth the image and minimise the noise. The Gaussian filter is determined as:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \text{---------(1)}$$

where σ is the standard deviation of the Gaussian distribution, and x and y are the coordinates of the pixel in the image.

The smoothed image I_s is acquired by convolving the original image I with the Gaussian filter G:

$$I_s = I \times G \text{---------(2)}$$

Gradient Calculation

Compute the intensity gradient of the smoothed image to identify edges. The gradients in the x and y directions (G_x and G_y) are calculated using Sobel operators:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \times I_s \text{---------(3)}$$

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \times I_s \text{---------(4)}$$

The gradient magnitude G and direction θ are computed as:

$$G = \sqrt{G_x^2 + G_y^2} \text{---------(5)}$$

$$\theta = tan^{-1}\left(\frac{G_y}{G_x}\right) \text{---------(6)}$$

Non-Maximum Suppression

Suppress non-maximum edges to thin out the edges. For each pixel, compare the gradient magnitude to its neighbors in the direction of the gradient:

$$NMS(x,y) = \begin{cases} G(x,y) & if\ G(x,y) is\ a\ local\ maximum\ along\ \theta(x,y) \\ 0 & otherwise \end{cases} \text{---------(7)}$$

Edge Tracking by Hysteresis

Identify final edges using two thresholds, T_{low} and T_{high}:

$$E((x,y) = \begin{cases} 1 & if\ NMS(x,y) > T_{high} \\ 0 & if\ NMS(x,y) < T_{low} \\ connected\ to\ a\ strong\ edge & if\ T_{low} \le NMS(x,y) \le T_{high} \end{cases} \text{---------(8)}$$

This results in the final set of edges, where strong edges are identified directly, and weak edges are included if they are connected to strong edges.

Morphological Operation

Morphological operations, such as dilation and erosion, are fundamental techniques used in image processing to enhance and clarify structural details within images.

Dilation is a morphological operation that expands the boundaries of regions of foreground pixels (typically white) in an image. It achieves this by adding pixels to the boundaries of objects in the image. The dilation of an image I by a structuring element B is defined as:

$(I \oplus B)(x, y) = \bigcup_{(s,t)\in B} I(x - s, y - t)$---------(9)

where $\oplus$ denotes the dilation operator, (x,y) are pixel coordinates, and B is the structuring element, which is a small binary image.

Erosion is another morphological operation that shrinks the boundaries of regions of foreground pixels in an image. It removes pixels from the boundaries of objects in the image. The erosion of an image I by a structuring element B is defined as:

$(I \ominus B)(x, y) = \bigcup_{(s,t)\in B} I(x + s, y + t)$---------(10)

where $\ominus$ denotes the erosion operator. The image preparation process is given in Algorithm 1.

Algorithm 1: Image Preparation Process

1. Data Gathering

   Input: Dataset of chemical compounds (both authentic and fake)

   Output: Molecular graphs for each compound

   1.1 Obtain dataset of chemical compounds

   1.2 Represent each compound as a molecular graph (nodes: atoms, edges: bonds)

2. Preprocessing

   Input: Molecular graphs

   Output: Preprocessed molecular graphs

   2.1 Apply Gaussian filter for noise reduction

      Function GaussianFilter(I, σ):

         I_s = Convolve(I, G)

         return I_s

   2.2 Compute intensity gradients using Sobel operators

      Function ComputeGradients(I_s):

         G_x = SobelX(I_s)

         G_y = SobelY(I_s)

         return G, θ

```
2.3 Perform non-maximum suppression
    Function NonMaximumSuppression(G, θ):
        NMS = ZeroMatrix(size(G))
        for each pixel (x, y) in G:
            if G(x, y) is a local maximum along θ(x, y):
                NMS(x, y) = G(x, y)
        return NMS
2.4 Apply edge tracking by hysteresis
    Function HysteresisEdgeTracking(NMS, T_low, T_high):
        E = ZeroMatrix(size(NMS))
        for each pixel (x, y) in NMS:
            if NMS(x, y) > T_high:
                E(x, y) = 1
            elif T_low <= NMS(x, y) <= T_high and connected to strong edge:
                E(x, y) = 1
        return E
2.5 Apply morphological operations
    Function MorphologicalOperations(E):
        E_dilated = Dilation(E, structuring_element)
        E_eroded = Erosion(E_dilated, structuring_element)
        return E_eroded
```

GNN based Classification

To classify original and counterfeit medicine images based on chemical structure graphs using a Graph Neural Network (GNN), particularly focusing on Graph Convolutional Networks (GCNs), here's a structured approach:

Graph Convolution Layer: Aggregate information from a node's neighbors using a weighted sum. The process is given in Equation 11.

$$h_v^{(k+1)} = \sigma\left(\sum_{u \epsilon N(v)} \frac{1}{|N(v)|} \cdot h_u^{(v)} \cdot W^{(k)} + W_0^{(k)} \cdot h_v^{(k)}\right) \text{---------(11)}$$

where $h_v^{(k+1)}$ is the feature vector of node v at layer v, $N(v)$ denotes the neighbors of node v, σ represents an activation function like ReLU, $W^{(k)}$ and $W_0^{(k)}$ are learnable weight matrices.

Pooling Layer: Reduce the graph size while preserving key features. The process is given in Equation 12.

$$h_G = POOL\left(\{h_v^{(K)} | v \in G\}\right)\text{---------(12)}$$

where h_G represents the pooled graph features, and $h_v^{(K)}$ are the features of nodes after K layers.

Fully Connected Layer: Perform final classification using softmax activation. The process is given in Equation 13.

$$\hat{y} = softmax(W_{fc} \cdot h_G + b_{fc})\text{---------(13)}$$

where $\hat{y}$ is the predicted probability vector, W_{fc}, and b_{fc} are the weights and biases of the fully connected layer. The tuning parameter for GNN is given in Algorithm 1.

Loss Function: Use cross-entropy loss for classification tasks is given in Equation 14.

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(\hat{y_i}) + (1 - y_i)\log(1 - \hat{y_i}))\text{---------(14)}$$

where y_i is the true label, and $\hat{y_i}$ is the predicted probability for the label i.

Optimization: Update model parameters using backpropagation and stochastic gradient descent (SGD) is given in Equation 15.

$$\theta \leftarrow \theta - \eta \cdot \nabla_\theta \mathcal{L}\text{---------(15)}$$

This structured approach using GCNs allows effective classification of images based on their chemical structure graphs, distinguishing between original and counterfeit medicines. Adjustments in network depth (number of layers), activation functions, and regularization techniques can further optimize performance based on specific dataset characteristics and computational constraints. The entire procedure counterfeit medicine classification is given in Algorithm 2.

Algorithm 2: Counterfeit Medicine Classification using GNN

1. Graph Construction
   Input: Pre-processed molecular graphs
   Output: Graph representations for GNN
   1.1 Convert pre-processed data into graph representations (nodes: atoms, edges: bonds)
2. GNN-based Classification
   Input: Graph representations
   Output: Classification results (authentic or fake)
   2.1 Define Graph Convolution Layer
      Function GraphConvolutionLayer():
         $h_v^{(k+1)} = \sigma\left(\sum_{u \epsilon N(v)} \frac{1}{|N(v)|} \cdot h_u^{(v)} \cdot W^{(k)} + W_0^{(k)} \cdot h_v^{(k)}\right)$
         return H_next
   2.2 Define Pooling Layer
      Function PoolingLayer(H):
         H_G = POOL(H)

$h_G = POOL\left(\{h_v^{(K)} | v \in G\}\right)$
return H_G
2.3 Define Fully Connected Layer
Function FullyConnectedLayer():
$\hat{y} = softmax(W_{fc} \cdot h_G + b_{fc})$
return $\hat{y}$
2.4 Define Loss Function (Cross-Entropy)
Function CrossEntropyLoss(y, y_hat):
L = -1/N * sum(y * log(y_hat) + (1 - y) * log(1 - y_hat))
$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i))$
return L
2.5 Define Optimization (SGD)
Function Optimization(L, θ, η):
θ = θ - η * ∇_θ L
$\theta \leftarrow \theta - \eta \cdot \nabla_\theta \mathcal{L}$
return θ
2.6 Training Loop
Initialize weights: W, W_0, W_{fc}, b_{fc}
for each epoch in num_epochs:
for each batch in dataset:
H = InitialNodeFeatures(batch)
A = AdjacencyMatrix(batch)
H = GraphConvolutionLayer(H, A, W, W_0)
H_G = PoolingLayer(H)
y_hat = FullyConnectedLayer(H_G, W_{fc}, b_{fc})
L = CrossEntropyLoss(y, y_{hat})
H_G = PoolingLayer(H)
y_{hat} = FullyConnectedLayer(H_G, W_{fc}, b_{fc})
L = CrossEntropyLoss(y, y_{hat})
θ = Optimization(L, θ, η)
return TrainedModel(θ)

Result and Discussion

When it comes to using the Python-based analysis for the dataset with 2,431,025 chemical compounds (drug details), the first step is data gathering from the sources such as ChEMBL and PubChem, and the second step is the thorough data pre-processing to clean and format the dataset. Exploratory data analysis (EDA) is then performed to identify the distributions and patterns of data and to guide the process of extracting molecular descriptors with the help of RDKit or Open Babel. The models using TensorFlow or PyTorch are trained and tested using the conventional 60:20:20 division for training, testing, and validation datasets [21]. The model parameters are as follows: 4 classes, batch size of 128,

hidden node count of 128, sequence length of 1 (assuming single node features per sequence), 100 epochs, optimizer as RELU Graph Neural Network (GNN), cross-entropy loss function, 3 layers, learning rate of 0.0011, and input node counts of 18, 15, and 16. The experiments were conducted using a CPU device.

The performance of the model is evaluated using various evaluation metrics on the testing and validation data and hyperparameter tuning is deployed to increase the accuracy and reliability of the model. Finally, using the trained models helps in tasks like predicting properties of compounds or in drug discovery, which shows the use of Python to deal and analyze big chemical data for pharma related research. The proposed CED-GNN is compared with existing techniques namely 1D-CNN [16] and YOLO [17].

The performance is investigated with the assistance of performance metrics namely accuracy, precision, and F1-score. Accuracy measures the proportion of correct predictions among all predictions. Precision is the ratio of true positive predictions to the total predicted positives, indicating how many selected items are relevant. F1-score is the harmonic mean of precision and recall, balancing both metrics for overall performance assessment. The input of the original Famotidine image is given in Figure 2, CED is given in Figure 3, and the morphological operation is given in Figure 4.

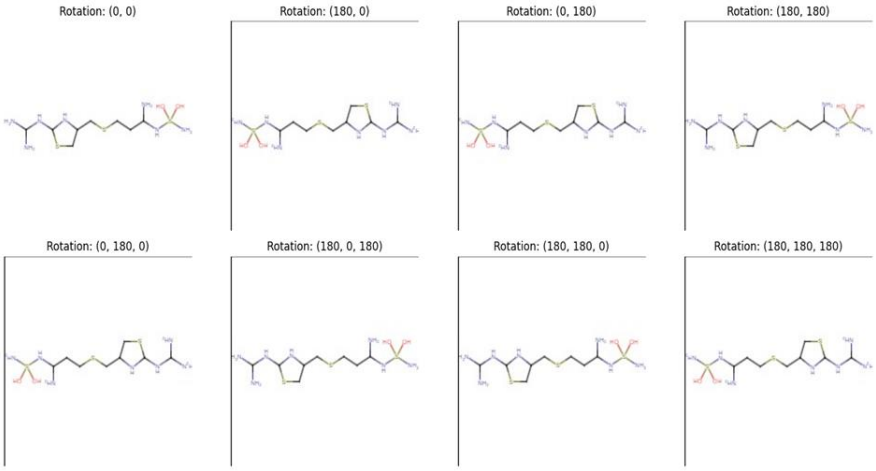

Figure 2. Input Image in Diverse Angles

The above figure shows the input image from different angle to give an idea about the versatility of the input image.

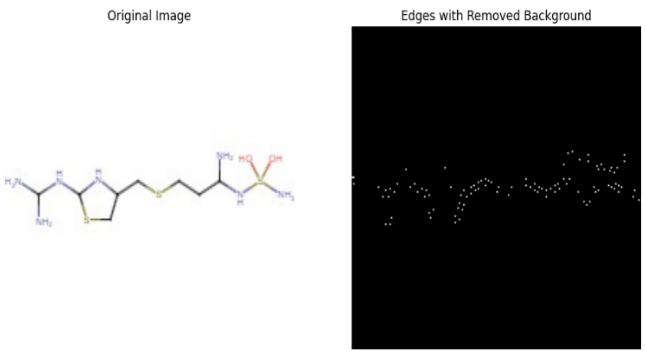

Figure 3. Canny Edge Detector of Famotidine

In figure 3, Canny edge detection technique is illustrated, which shows that Canny detects edges sharply and accurately in the image.

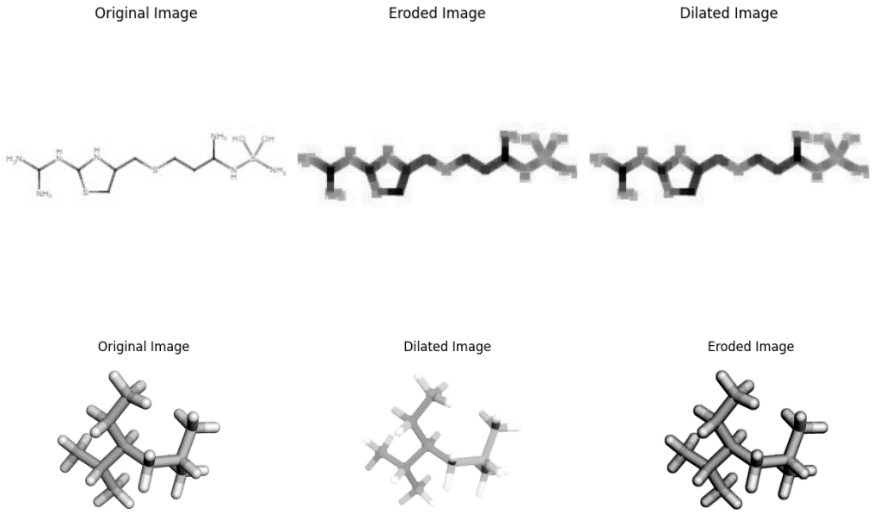

Figure 4. Morphological Operation - Dilation and Erosion of Famotidine

Figure 4 shows the morphological operations of dilation and erosion where dilation increases the size of the object boundary and erosion decreases it, which gives a better understanding of spatial transformations for feature enhancement and extraction in the field of computer vision. Each figure is used to demonstrate various processes of image manipulation starting from the capture of the image to the enhancement of edges and structural changes.

Table 2. Comparison of Accuracy

Epoch	1D-CNN	YOLO	CED-GNN
100	70.93	73.43	77.73
200	72.06	74.89	79.14
300	73.31	75.14	79.53
400	73.77	75.87	80.36
500	75.88	76.81	81.91

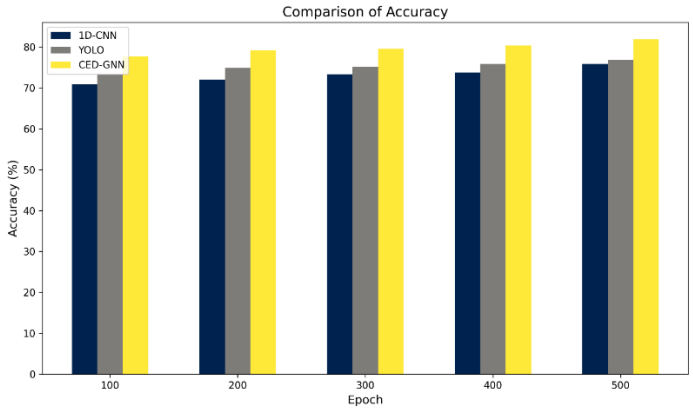

Figure 5. Comparison of Accuracy

From the Table 2 and Figure 5, it was observed that CED-GNN achieved higher accuracy than 1D-CNN and YOLO in all the epochs. Starting from 77.73% at 100 epochs, and which demonstrates a steady increase at each epoch, and ended at 81.91% at 500 epochs. The accuracy of YOLO grows progressively from 73.43% to 76.81% over 500 epochs. On the other hand, 1D-CNN has the lowest accuracy gain that begins at 70.93% and achieving 75.88% at 500 epochs. From this, it can be concluded that CED-GNN has a better learning capacity and generalization performance than the other models.

Table 3. Comparison of Precision

Epoch	1D-CNN	YOLO	CED-GNN
100	73.33	75.76	78.73
200	73.19	76.67	79.83
300	73.33	77.88	80.67
400	75.56	79.01	80.91
500	75.71	79.12	81.46

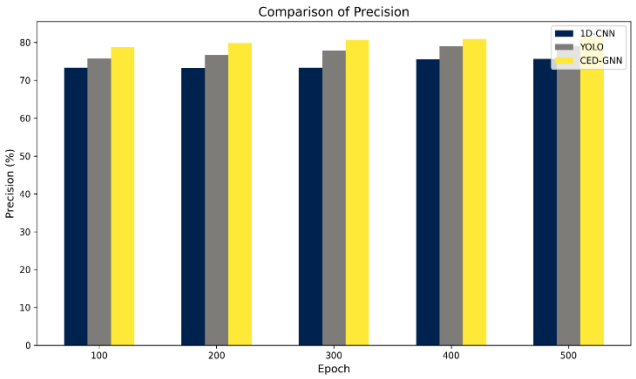

Figure 6. Comparison of Precision

Table 3 shows the precision metrics and Figure 6 represents the same. CED-GNN is always higher than the other models, beginning at 78.73% precision after 100 epochs and increasing up to 81.46% at 500 epochs. In the case of YOLO, the values also rise progressively from 75.76% to 79.12%. While 1D-CNN starts at a similar precision of 73.33%, and gets better relatively slower, attaining 75.71% by 500 epochs. It is believed that since CED-GNN has a higher precision, it can reduce false positive rates to the barest minimum.
Table 4 and figure 7 shows the F1-scores that is the harmonic mean of precision and recall.

Table 4. Comparison of F1-Score

Epoch	1D-CNN	YOLO	CED-GNN
100	77.13	75.9	78.67
200	77.36	76.45	79.09
300	78.9	76.76	79.88
400	79.06	77.1	80.98
500	79.6	78.65	82.99

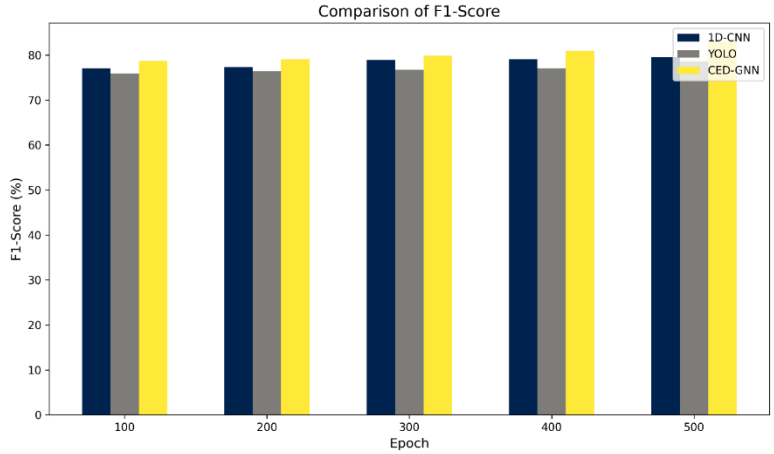

Figure 7. Comparison of F1-Score

CED-GNN performs better having a starting score of 78.67% at 100 epochs, it rises to 82.99% at 500 epochs. YOLO has a very similar trend and gradually increases from 75.9% to 78.65%. 1D-CNN, although increasing from 77.13% to 79.6% is still lower than the other models. This shows that CED-GNN has a better F1-score meaning it has a better balance in false positives and false negatives as compared to the other models.

CED-GNN proves to be more accurate, precise, and possess a higher F1-score than 1D-CNN and YOLO in all epochs. From the results presented above, it is possible to observe the training process improvements, which prove better learning and generalization of the superior model. YOLO is slightly better than 1D-CNN with moderate improvement, which means that it is reliable but less efficient than CED-GNN. 1D-CNN, as much as it has been improving gradually, is the worst-performing model of the three. As for the results, CED-GNN is shown to be the most suitable for tasks that demand high accuracy, precision, and equally important, the balance between them.

Conclusion

In this research article, the authors present the possibility of using Graph Neural Networks (GNNs) to address the problem of counterfeit medicine classification. Thus, based on the structural characteristics of chemical compounds, the proposed system should provide higher accuracy and system reliability compared to conventional approaches. At the center of this strategy is graph representation which uses nodes to depict atoms and edges to depict bonds of the chemical compounds. The GNNs which have the capacity to process graph-based data are used to capture the complex structure and features of such graphs. Such a representation makes it possible to distinguish the molecular features that define genuine medicines from counterfeit ones. The experimental results show that the proposed GNN-based system is better than traditional machine learning algorithms in accuracy, stability and scalability. The model's capacity to generalize across different types of medicines also strengthens its practical applicability. The application of GNNs in counterfeit medicine classification is a promising improvement. In addition to enhancing the classification performance, the structure-based approach also provides a more scalable and modifiable approach to tackle the problem of counterfeit medicines across the world, hence providing safer health care solutions. The future work might be directed

towards the fine-tuning of the GNN structure and towards the investigation of the possibility to combine it with other deep learning approaches for maximum accuracy of 81.91%

Reference

1. Staszak, M., Staszak, K., Wieszczycka, K., Bajek, A., Roszkowski, K., & Tylkowski, B. (2022). Machine learning in drug design: Use of artificial intelligence to explore the chemical structure–biological activity relationship. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *12*(2), e1568.
2. Chen, W., Liu, X., Zhang, S., & Chen, S. (2023). Artificial intelligence for drug discovery: Resources, methods, and applications. *Molecular Therapy-Nucleic Acids*, *31*, 691-702.
3. Huang, K., Xiao, C., Hoang, T., Glass, L., & Sun, J. (2020, April). Caster: Predicting drug interactions with chemical substructure representation. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 34, No. 01, pp. 702-709).
4. Yang, H., Lou, C., Li, W., Liu, G., & Tang, Y. (2020). Computational approaches to identify structural alerts and their applications in environmental toxicology and drug discovery. *Chemical Research in Toxicology*, *33*(6), 1312-1322.
5. Cheruto, M. (2021). *Analysis of counterfeit-drug-prevention strategies on financial performance of Nairobi's drug retailing pharmacies* (Doctoral dissertation, Strathmore University).
6. Saldívar-González, F. I., Aldas-Bulos, V. D., Medina-Franco, J. L., & Plisson, F. (2022). Natural product drug discovery in the artificial intelligence era. *Chemical Science*, *13*(6), 1526-1546.
7. Chen, Y., & Kirchmair, J. (2020). Cheminformatics in natural product-based drug discovery. *Molecular Informatics*, *39*(12), 2000171.
8. Nayarisseri, A., Khandelwal, R., Tanwar, P., Madhavi, M., Sharma, D., Thakur, G., ... & Singh, S. K. (2021). Artificial intelligence, big data and machine learning approaches in precision medicine & drug discovery. *Current drug targets*, *22*(6), 631-655.
9. Pérez Santín, E., Rodríguez Solana, R., González García, M., García Suárez, M. D. M., Blanco Díaz, G. D., Cima Cabal, M. D., ... & López Sánchez, J. I. (2021). Toxicity prediction based on artificial intelligence: A multidisciplinary overview. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *11*(5), e1516.
10. Elbadawi, M., McCoubrey, L. E., Gavins, F. K., Ong, J. J., Goyanes, A., Gaisford, S., & Basit, A. W. (2021). Harnessing artificial intelligence for the next generation of 3D printed medicines. *Advanced Drug Delivery Reviews*, *175*, 113805.
11. Wang, D., Liu, W., Shen, Z., Jiang, L., Wang, J., Li, S., & Li, H. (2020). Deep learning based drug metabolites prediction. *Frontiers in Pharmacology*, *10*, 1586.
12. Staszak, M., Staszak, K., Wieszczycka, K., Bajek, A., Roszkowski, K., & Tylkowski, B. (2022). Machine learning in drug design: Use of artificial intelligence to explore the chemical structure–biological activity relationship. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *12*(2), e1568.
13. Askr, H., Elgeldawi, E., Aboul Ella, H., Elshaier, Y. A., Gomaa, M. M., & Hassanien, A. E. (2023). Deep learning in drug discovery: an integrative review and future challenges. *Artificial Intelligence Review*, *56*(7), 5975-6037.

14. Hu, S., Chen, P., Gu, P., & Wang, B. (2020). A deep learning-based chemical system for QSAR prediction. *IEEE journal of biomedical and health informatics*, *24*(10), 3020-3028.
15. Gayialis, S. P., Kechagias, E. P., Papadopoulos, G. A., & Masouras, D. (2022). A review and classification framework of traceability approaches for identifying product supply chain counterfeiting. *Sustainability*, *14*(11), 6666.
16. An-Bing, Z., Hui-Hua, Y., Xi-Peng, P., Li-Hui, Y., & Yan-Chun, F. (2020). On-site identification of counterfeit drugs based on near-infrared spectroscopy Siamese-network modeling. *IEEE Access*, *9*, 3195-3206.
17. 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*, 1-9.
18. Simonovsky, M., & Komodakis, N. (2018). Graphvae: Towards generation of small graphs using variational autoencoders. In *Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part I 27* (pp. 412-422). Springer International Publishing.
19. Kakio, T., Yoshida, N., Macha, S., Moriguchi, K., Hiroshima, T., Ikeda, Y., ... & Kimura, K. (2017). Classification and visualization of physical and chemical properties of falsified medicines with handheld Raman spectroscopy and X-ray computed tomography. *The American journal of tropical medicine and hygiene*, *97*(3), 684.
20. Deconinck, E., Sacre, P. Y., Courselle, P., & De Beer, J. O. (2012). Chemometrics and chromatographic fingerprints to discriminate and classify counterfeit medicines containing PDE-5 inhibitors. *Talanta*, *100*, 123-133.
21. https://www.ebi.ac.uk/chembl/web_components/explore/compounds/
22. R. Renukadevi et al. “An Improved Collaborative User Product Recommendation System Using Computational Intelligence with Association Rules”, Communications on Applied Nonlinear Analysis**,** Volume 31, Issue 6s (2024), pp. 554 - 564**.** https://doi.org/10.52783/cana.v31.1243
23. K T, S., P, S. S., Kumar E, B., L R, S., J, V., & R., N. (2024). Experimental Assessment between Dissimilar Techniques and Methodologies to Sports Knee Injury using Magnetic Resonance Imaging. *South Eastern European Journal of Public Health*, 1635–1644. https://doi.org/10.70135/seejph.vi.2167
24. R. Renukadevi et al. “Optimized Computational Intelligence with Association Rules for a Collaborative Consumer Product Recommendation System”, Advances in Nonlinear Variational Inequalities, Volume 31, Issue 3s (2025), pp. 375 – 385. https://doi.org/10.52783/anvi.v28.3065
25. Prakash, G., P. Logapriya, and A. Sowmiya. "Smart Parking System Using Arduino and Sensors." *NATURALISTA CAMPANO* 28 (2024): 2903-2911.