

Project 1: Signature Forgery Detection Using Graph Neural Networks

Abstract

Due to the rapid development of artificial intelligence technology, the usage of smart tools to modify or even fake documents is becoming more and more frequent. This project is aimed to automatically detect signs of tampering and forgery signature in different kinds of documents, such as manipulated checks, insurance and educational paperwork. The project will utilize graph neural network technology to screen, analyze and identify possible counterfeit documents, to reduce the potential possibility of property loss or social damage.

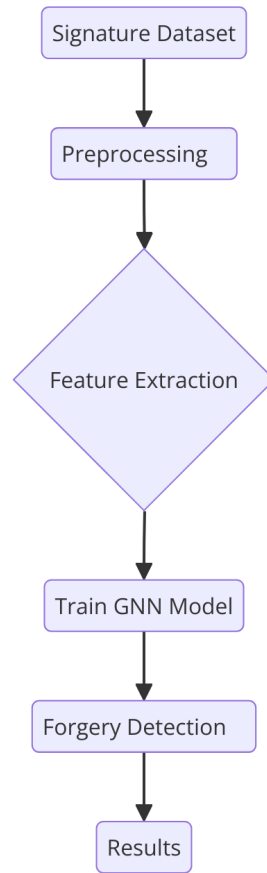


Fig.1. Flow diagram of research methodology

However, as far as graph neural networks are concerned, this technology is unable to realize all the requirements of the customers. Anti-counterfeiting and counterfeiting have been dynamic games since ancient times, and counterfeiters usually have the advantage of information asymmetry. Therefore, while using graph neural networks, it is also important to focus on game-theoretic strategies in the future. The LLM model can help to propose forward-looking defense strategies against counterfeiting as well as designing a more effective detection system.

Target User/Application

Financial institutions/Banks: Financial institutions like banks deal with a large number of financial documents, such as checks, loan applications, mortgage documents. Graph neural network detection can reduce financial fraud by automatically identifying defects or questionable areas and compare with existing databases.

Educational institutions and accreditation agencies: Universities like Boston University need to verify the authenticity of documents such as students' academic certificates and transcripts. Typically, such certificates have a security number or an official link. Single image neural network detection may not ensure 100% accuracy, but it can be effective for initial screening and can also increase automation efficiency.

Transnational corporations and international organizations: When dealing with international trade, cooperation agreements or outsourcing contracts, document forgery detection techniques can be used to verify the legitimacy of documents and prevent transnational fraud.

Insurance companies: Insurance companies need to handle many claims, which may involve forged physician signature in medical reports. The accuracy of claims review can lead to reduction of losses due to fraudulent behavior.

More Clients: Audit & Accounting Firms, Legal and judicial authorities etc.

ChatGPT Responses

Question: Can we use Graph Neural Network to detect signature forgery?

ChatGPT:

Yes, Graph Neural Networks (GNNs) can be used to detect signature forgery, although their application in this domain is still emerging. The unique ability of GNNs to process data structured as graphs can be advantageous for understanding the spatial relationships and structures present in handwriting and signatures. Here's how GNNs could be applied for signature forgery detection:

- **Signature Structure as Graph:** A signature can be viewed as a series of strokes or points that are connected in a specific sequence, forming a graph structure. GNNs are effective at learning patterns from data represented as graphs, such as nodes (representing points in the signature) and edges (representing strokes or connections between points).
- **Relational Data Representation:** GNNs can capture complex relationships between the spatial points of a signature, making them well-suited to detect subtle differences in stroke patterns, curvature, and spacing between points—features that are often indicative of forgery.

- **Invariance to Permutation:** Since GNNs are designed to be invariant to the ordering of graph nodes, they are robust to small variations in signature shape, which is critical for identifying both genuine signatures and forgeries.

Possible Approach for Using GNNs in Signature Forgery Detection

Step 1: Preprocessing (Feature Extraction)

- **Convert the signature image into a graph representation:** Use key points detection algorithms (e.g., SIFT, Harris corners) to extract important points from the signature image.
- **Create a graph where nodes represent these key points and edges represent the relationships between them based on the structure of the signature's strokes.**
- **The features of each node (e.g., location, curvature) can be included as node attributes.**

Step 2: Graph Construction

Once the signature is transformed into a graph:

- **Edges can represent the connectivity between sequential or neighboring strokes.** Edge weights could be based on spatial proximity, direction of the stroke, or curvature intensity.

Step 3: Model Training

- **Use a GNN architecture (such as Graph Convolutional Networks or Graph Attention Networks) to learn patterns in genuine signatures.**
- **Train the GNN on a labeled dataset containing genuine signatures and known forgeries.**
- **The GNN can learn to differentiate the subtle differences in stroke patterns, pressure, and flow between authentic signatures and forgeries.**

Step 4: Forgery Detection

- **The trained GNN can classify whether a given signature is authentic or forged based on its learned knowledge of spatial and structural patterns.**

Challenges:

- **Data Representation:** Transforming signature data into graph form can be complex, and selecting the right graph structure (i.e., defining nodes and edges) is critical for GNN performance.
- **Limited Datasets:** GNNs require a well-curated dataset of both genuine and forged signatures. Acquiring a large and diverse dataset can be challenging, particularly for rare types of forgeries.
- **Computational Complexity:** GNNs, like many neural networks, can be computationally expensive to train, especially on large graph structures or datasets.

While GNNs can be powerful, other deep learning techniques, such as **Convolutional Neural Networks (CNNs)**, have been widely used for image-based tasks like signature verification and forgery detection. In some cases, hybrid models combining GNNs and CNNs could also be effective, where CNNs handle image feature extraction and GNNs model the relational structure between these features.

Conclusion

Graph Neural Networks can be a valuable tool for detecting signature forgery by modeling the spatial and structural relationships within a signature. Although the application of GNNs in this field is still evolving, their potential for capturing complex relationships in handwriting makes them a promising approach. However, success will depend on proper graph representation, dataset availability, and model training.

Analysis of LLM results

From my perspective, LLM like ChatGPT can provide initial guidance and advice on the structure and planning for the project. They can offer explanations for the amateur or those who have no knowledge of it at all. In this project, when I first thought of this research direction, I wasn't sure if signature forgery analysis was attainable with GNN. In this circumstance, LLM could help determine the feasibility of the project. Obviously, LLM can outline and explain some of the details of the section of the project.

But for deeper analysis, LLM cannot do the job of analyzing very specific and detailed points. For example, while organizing the graph relations, what is the specific relation between edges and points. LLM outputs can be highly context-dependent, because context may change over time, thus making the analysis methods obsolete and inaccurate. LLM has provided me with an outline, but I still hold doubts about how to do graphical relationship plotting and how to express them in mathematical relationships.

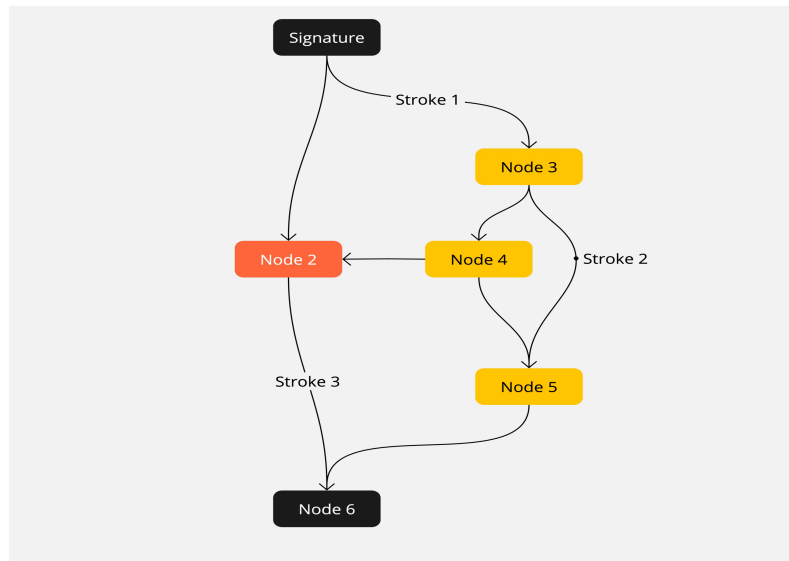


Fig.2. Assumed Signature Analysis with Nodes and Edges

Figure 2 shows the pattern of nodes that I hypothesized the GNN would need to learn. It is impossible to be specific in the results of LLM as to how this relationship would be proposed and created. Thereafter, the correctness and validity also require verification.

In summary, LLM is excellent at offering ideas and addressing processes. They can enlighten me a lot and search for technical specifications and tools that the project is depending on. However, accuracy and timeliness can often be a problem. We can't count on LLM for projects and research, they are merely tools and advisors to query and search when I have concerns.

Resources:

Kong, B., Ai, T., Zou, X., Yan, X., & Yang, M. (2024). A graph-based neural network approach to integrate multi-source data for urban building function classification. *Computers, Environment and Urban Systems*, 110. <https://doi.org/10.1016/j.compenvurbsys.2024.102094>

Phan, H. T., Nguyen, N. T., & Hwang, D. (2023). Fake news detection: A survey of graph neural network methods. In *Applied Soft Computing* (Vol. 139). Elsevier Ltd. <https://doi.org/10.1016/j.asoc.2023.110235>

Lin, X., Wang, S., Deng, J., Fu, Y., Bai, X., Chen, X., Qu, X., & Tang, W. (2023). Image manipulation detection by multiple tampering traces and edge artifact enhancement. *Pattern Recognition*, 133. <https://doi.org/10.1016/j.patcog.2022.109026>