Stegano - LLM
Ai Powered Steganography

Name : Divit Patidar
Net-ID : Dp78

Motivation

In an era where data privacy and secure communication are paramount, Stegano-LLM leverages the latest advancements in artificial intelligence and language models to revolutionize the field of steganography.

By seamlessly integrating large language models like Dolly-v-3B, this project introduces a novel approach to concealing sensitive information within natural language prompts generated by these AI systems.

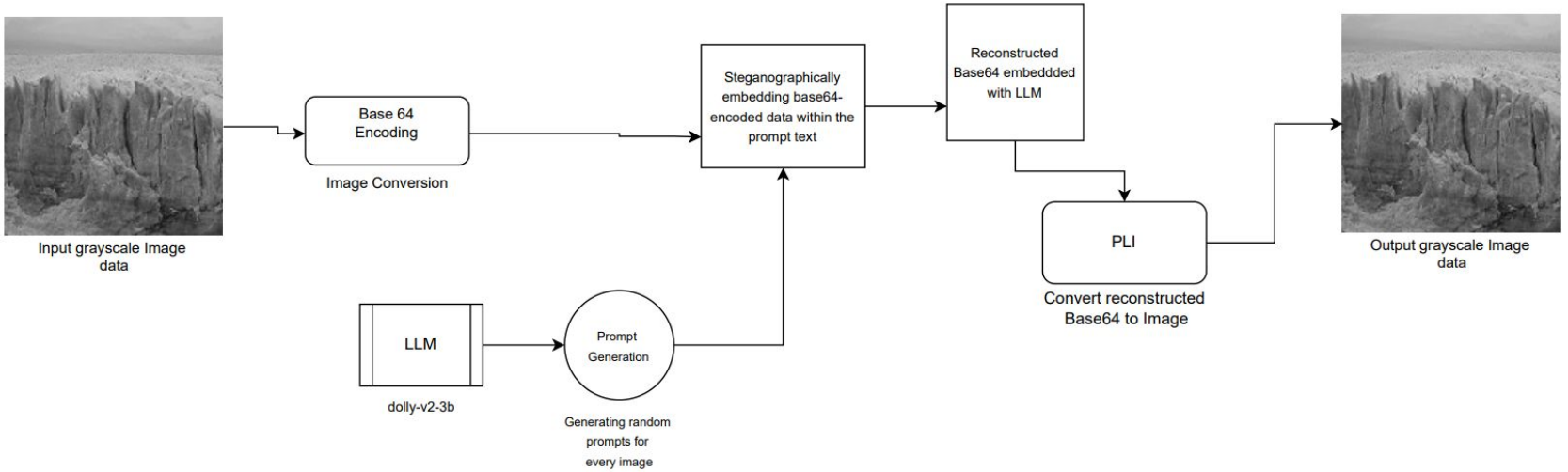
Through this innovative technique, data can be securely embedded and transmitted, virtually indistinguishable from regular text, offering a robust and adaptable solution for applications such as digital watermarking, covert communication, and tamper detection.

Stegano-LLM represents a significant step forward in the pursuit of information security in the digital age.

Problem Setup

The Stegano-LLM project tackles the limitations of traditional image steganography by developing adaptive techniques that leverage language models to robustly embed and securely conceal encoded image data within AI-generated textual prompts, enabling high-capacity, evolving, and indistinguishable data concealment.

Model



Here we use dolly-v2-3b LLM model from databricks this will generate unrelated descriptive prompts for images stored in the directory, the images that were converted into base64 & the prompts generated by the model are further embedded together the combination of data embedding within a text produces steganography technique.

Process

- Image Encoding: Converts images into base64 strings, turning binary data into text.
- Prompt Generation: Uses the Dolly-v-3B model to create artistic descriptions for each image, designed to divert from the actual content (enhancing steganography).
- Data Embedding: Embeds the base64 image data within these prompts at specific intervals or marked by delimiters, effectively hiding the image data within text.
- Data Reconstruction: Implements methods to extract and decode the embedded base64 data from the prompts, allowing for image retrieval.

Results

original and embedded prompts, with capabilities to extract embedded image data, demonstrating the use of advanced AI with steganography for secure data handling.

Input Image

Running through the process & reconstructing same image

Output Image

The above image shows after all the steps we get a reconstructed image which was same as input