AI-Driven Story Teller

A major project report submitted in partial fulfillment of the requirement for the award of degree of

Bachelor of Technology

In

Computer Science & Engineering / Information Technology

Submitted by

Chirag Walia (201304)

Jagrit Kamra (201218)

Under the guidance & supervision of

Dr. Ekta Gandotra

Department of Computer Science & Engineering and Information Technology

Jaypee University of Information Technology, Waknaghat, Solan - 173234 (India)

TABLE OF CONTENTS

Cover Page

Certificate i

Declaration ii

Acknowledgement iii

List of Tables iv

List of Figures v

List of Abbreviations, Symbols or Nomenclature vi

Abstract vii

Chapter 1: Introduction 1-7

1.1 Introduction 1-3

1.2 Problem Statement 4

1.3 Objectives 5

1.4 Significance and Motivation of the Project Work 6

1.5 Organization of Project Report 7

Chapter 2: Literature Survey 8 - 20

2.1 Overview of Relevant Literature 8 - 20

2.2 Key Gaps in the Literature 20 - 21

Chapter 3: System Development 22 - 38

3.1 Requirements and Analysis 22 - 22

3.2 Project Design and Architecture 22 - 30

3.3 Data Preparation 31 - 31

3.4 Implementation 31 - 37

3.5 Key Challenges 37 - 39

Chapter 4: Testing 40 - 45

4.1 Testing Strategy 40 - 40

4.2 Test Cases and Outcomes 40 - 45

Chapter 5: Results and Evaluation 46 - 46

5.1 Results 46

Chapter 6: Conclusions and Future Scope 47 - 47

6.1 Conclusion 47

6.2 Future Scope 47

References 48 - 50

CERTIFICATE

This is to certify that the work which is being presented in the project report titled **"AI-Driven Story Teller"** in partial fulfillment of the requirements for the award of the degree of B.Tech in Computer Science And Engineering and submitted to the Department of **Computer Science And Engineering,** Jaypee University of Information Technology, Waknaghat is an authentic record of work carried out by Chirag walia(201304) and Jagrit Kamra(201218) during the period from July 2023 to December 2023 under the supervision of **Dr. Ekta Gandotra**, Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat.

Chirag Walia (201304) **Jagrit Kamra (201218)**

The above statement made is correct to the best of my knowledge.

Dr. Ekta Gandotra
Associate Professor
Computer Science & Engineering and Information Technology
Jaypee University of Information Technology, Waknaghat

Candidate's Declaration

I hereby declare that the work presented in this report entitled '**AI-Driven Story Teller**' in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology in Computer Science & Engineering / Information Technology** submitted in the Department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat is an authentic record of my own work carried out over a period from August 2023 to December 2023 under the supervision of **Dr. Ekta Gandotra** (Associate Professor, Department of Computer Science & Engineering and Information Technology).

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

(Student Signature with Date)

Student Name:

Roll No.:

(Student Signature with Date)

Student Name:

Roll No.:

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

(Supervisor Signature with Date)

Supervisor Name:

Designation:

Department:

Dated:

ACKNOWLEDGEMENT

We owe our profound gratitude and indebtedness to our project supervisor **Dr. Ekta Gandotra**, who took keen interest and guided us all along in our project work titled — **AI-Driven Story Teller**, till the completion of our project by providing all the necessary information for developing the project. The project development helped us in research, and we got to know a lot of new things in our domain. We are really thankful to her.

Group No. – 47

Name: Chirag Walia

Roll No.: 201304

Name: Jagrit Kamra

Roll No.: 201218

LIST OF TABLES

S. No	Title	Page No.
1	Different Literature survey with result and limitation	8-10
2	System Requirement	11
3	Model validate key points	31

LIST OF FIGURES

S. No.	Title	Page No.
1	Flowchart of model integration	2
2	Original Image	7
3	Gray Scale Conversion	7
4	Detected Text	7
5	Extracted Text	7
6	Model Architecture	12
7	COCO Dataset	15
8	Local Dataset	16
9	Flowchart of Login and Sign-up	17
10	Welcome Screen Design	18
11	UI Implementation (Image to story , combine Images , Output)	18-19
12	Set of Output Images	28-30
13	Graph of The Result Obtained	32-36

LIST OF ABBREVIATIONS

S. No	Title
1	GAN: Generative Adversarial Network
2	OCR: Optical Character Recognition
3	MERN: Mongo dB Express React Node
4	R-CNN: Region-Based Convolutional Neural Network
5	CNN: Convolution Neural Network
6	UI/UX: User Interface / User Experience
7	TTS: Text To Speech
8	API: Application Programming Interface
9	DIP: Digital Image Processing

ABSTRACT

In the realm of artificial intelligence, there is an equally important need for AI-Driven Story Teller. This innovative system harnesses the power of Generative Adversarial Network (GAN) to provide a narrative story to user input. The user interface prompts the user to provide a base input in the form of a set of images using which our story is generated. The design and implementation of such a platform is proposed in this study.

Need for this project varies from application to application, movie writers have trouble getting ideas or motivations for a small scene or script. This application can prove inspiring for them. It makes storytelling easier and more accessible. We use advanced technology and ideas from users to create engaging stories. To understand, and feel them. And a bonus implementation would be to train the model using generated stories based on feedback based output.

Our solution addresses various applications, including movie scriptwriting for filmmakers and writers in the entertainment industry, generating storylines for online game ideas, facilitating grammar learning for kids, creating digital marketing product ads, and generating personalized stories for tourism blogs. Leveraging OCR techniques and multiple image detection methods from various AI models, particularly YOLO8 and Open-AI, we plan to utilize open APIs to generate genre-based stories. Extensive testing, including unit testing, will be conducted on the models to ensure their effectiveness. The culmination of this effort will be a comprehensive full-stack application developed using the MERN stack (Node.js, React, Express, Mongo DB), hosted on AWS for easy accessibility by a diverse user base.

The performance of our model was evaluated through a survey involving faculty members from diverse Humanities and Social Sciences disciplines. Feedback was collected and scored based on predetermined criteria to assess the model's accuracy and effectiveness. Results indicate that 53% of respondents rated the model as "very good," while 32% rated it as "good," reflecting broad satisfaction among participants. These findings affirm the success of the project and its ability to meet the expectations of its intended audience.

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

We are providing a new and advanced technology with great storyline for a market that craves constantly new solutions. This MERN stack application (MongoDB, Express.js, React, and Node.js) is the best instance representing harmonization in back-end performance and front-end liveliness. In essence, using image to text transformation with GANs opens a whole new chapter in storytelling that goes way beyond tradition.

Technological Synthesis:

Therefore, we have chosen the MERN stack as the technical backbone for our web application where these elements combine together to yield smooth running system. The stores multimedia data in MongoDB, a NoQuery database, while providing efficient server-side development through Express.js. Therefore, react–a declarative java script library is very good at creating very responsive and rich UI's. Node.js coordinates these elements that forms the permanent foundation on which we built the new way of storytelling.

Revolutionizing Story Creation:

In a nutshell, our software revolutionizes the use of image-to-text conversion and GAN neural networks. These dynamic twins help extract the narrative potential from pictures and convert those pictures into textual data which forms the base of the storytelling process. This is then expanded with GAN integration that creates fresh images guided by an unfolding narrative as well as custom user feedback. The outcome is not merely a tale, but rather, an unfolding visual and narrative phenomenon which transcends traditional ways of thinking about stories.

Empowering User Engagement:

Its main function is to convert pictures into text data and use provided genres and lines for creation of unique stories. It involves the audience in activities that make them active agents and also helps mold the flavor in every story towards its direction. Furthermore, the sophisticated ability to create novel images stemming from the unfolding story makes for an exciting exploration of each tale.

Targeted Domains and User-Centric Features:

Our application is very flexible and it covers a range of fields addressing numerous end users. For the entertainment sector, it serves the ordinary people, movie lovers, and game players through development of narrative plots for films, T.V. series, gaming and narrative experience. A storytelling tool in children's education and entertainment, it is useful in creation of interactive and educative tales.

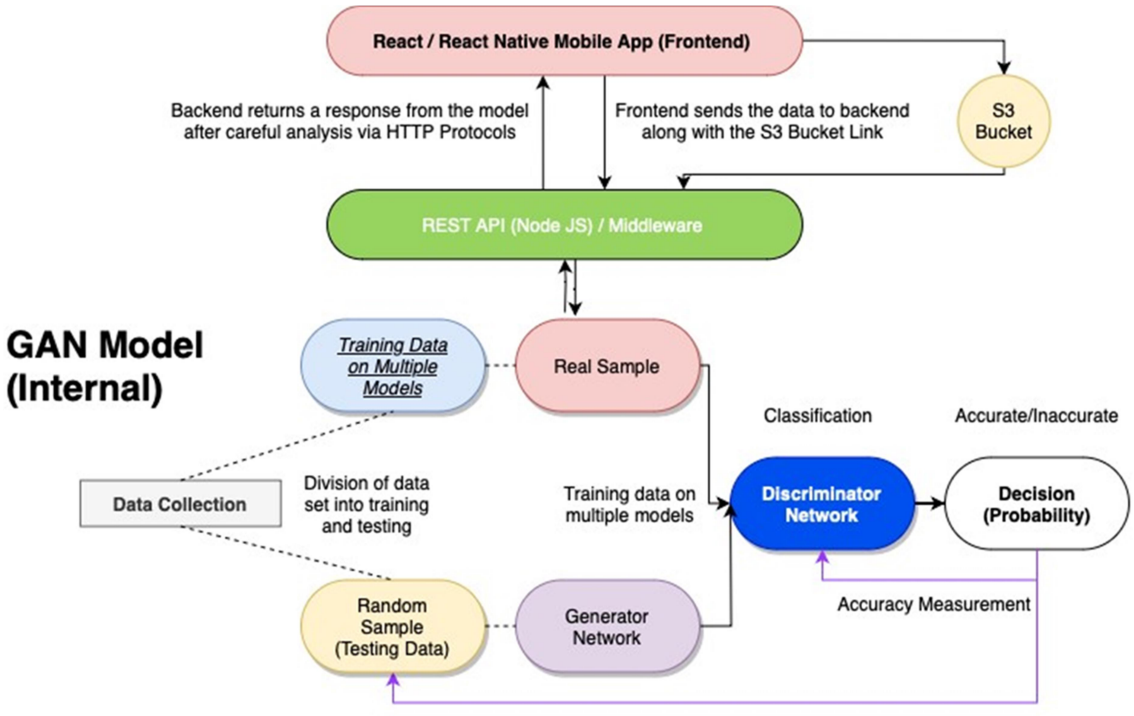


Figure 1: Flowchart of model integration

Figure 1 depicts the basic structure of our full stack application:

1. React is a frontend framework made in the JS Language which makes state management easily and it divides the code into different components.
2. We use S3 Buckets to host the images uploaded online, we plan to remove the image after use in future test cases to improve the security and costs of the same service.
3. We use Backend API's made in Node js using the express framework which controls middleware and services to communicate to the GAN Model..

GAN Model Internal Working (Overview)

A Generative Adversarial Network (GAN) for story generation works with two components:

Generator:

- Takes inputs like genre, keywords, and additional text.
- Generates synthetic stories based on these inputs.

Discriminator:

- Distinguishes between real and generated stories.
- Guides the generator by providing feedback on the realism of the generated stories.

Training Process:

- The generator aims to create stories that fool the discriminator.
- The discriminator is trained to correctly identify real and generated stories.

Output:

- Once trained, the generator produces realistic stories based on given inputs.

Fine-tuning is crucial, and the process repeats until the generator creates compelling, lifelike stories.

1.2 PROBLEM STATEMENT

The combination of picture analysis and narrative offers a new paradigm in the fast developing field of artificial intelligence, with broad connection in a variety of fields. Our AI- powered narrator, which creates stories from a collection of photos, is a flexible tool that can be used with a wide range of target users.

This design is based on Semantic Understanding as the gentle process that makes a computer know what's moving in an image. However, the system attempts to provide a strong base upon which fascinating stories may be created by taking into account more than just the superficial appearance. However, making the computer understand images is more difficult because the computer should be able not only to recognize but also study pictures deeper than simply learn them.

Another problematic area is the structure of the narrative for our AI- driven fibber. The objective of this system is not only to narrate tales, but also tales that correspond with the logic of the stoner. The narrator must be interested in what the computer story has to say about the set of sophisticated images used to ensure that the final product is satisfactory. It's a challenging thing to do in the creative-artificial world because it involves mixing up the psycho-delines and writing a fascinating story.

Contextual Conclusion pushes the design's complexity indeed further by exploring the idea of comprehending passions and feelings in visual content. The computer's job is to interpret the possible emotional countries of the people portrayed in the filmland and use this understanding to tell the story in a way that makes sense for the environment. This means that emotional intelligence must be incorporated into the liar process in order for the AI to give stories a mortal- suchlike quality that's sensitive to the craft of the scenes it's chronicling.

To put it simply, our AI- powered fibber breaks through traditional boundaries by using a complex commerce between semantic understanding, narrative structure, and contextual conclusion to not only produce stories but also tell them. This invention has implicit uses in a number of fields, including remedy and education as well as entertainment.

1.3 OBJECTIVES

The main objectives of are project are –

1. Narrative GAN: To implement Image-to-Text Conversion using the GAN model and generate a story based on it.

2. Deployment Algorithm[19]: To develop and deploy the Algorithm using RESTful apis on a react web app.

3. Image union Application: To implement union image generation from image conversion and user add-on text.

1.4 SIGNIFICANCE AND MOTIVATION OF PROJECT WORK

An AI driven storyteller which generates stories from photos has importance in various fields and provides customised tale telling experiences to different users. These possibilities are enormous ranging from entertainment, education to advertising and therapeutic stories.

In the Entertainment Industry, our system has the implicit to revise how stories are drafted for pictures, television shows, and videotape games. The engagement of the general public, liar suckers, moviegoers, and gamers is heightened as the AI weaves dynamic narratives grounded on visual content, creating interactive and immersive gests .

In Children's Education and Entertainment, the AI fibber becomes a precious tool for creating educational and interactive liar gests . Children, parents, and preceptors can profit from engaging narratives that grease literacy and allure youthful minds.

The projects goal is to bring a modern era of narrative story generation with a fusion of images and AI that creates stories that resonate across multiple domains . our AI-Driven story teller aims to change the boundaries of storytelling.

In conclusion ,the development of AI-Driven story teller is by the motivation of providing dynamic but also inclusive and personalised ways of generating stories to the user and treating it as the primary objective. Moreover the system tries to ensure a simple and interactive user interface.

1.5 ORGANISATION OF PROJECT REPORT

The report it organised as follows:

- Chapter-01 presents with the introduction of the study, along with the problem statement, objectives, significance, and organisation of the project.

- Chapter-02 outlines the existing related work in the field of image to text conversion, it further presents the outputs of their analysis which we compare and use in our project.

- Chapter-03 puts forward the system that we have developed conversion from image to text. This is the chapter that includes the requirements, analysis, the design of the project and its architecture and the implementation of the project.

- Chapter-04 gives the Testing strategy or the tools used in the project.

- Chapter-05 presents the results and evaluation of the project.

- Chapter-06 presents the conclusion of the study along with future scope.

CHAPTER 2: LITERATURE SURVEY

2.1 Overview of Relevant Literature:

Janvalkar et al. [1], introduce the concept of text-to-image conversion and highlights the challenges associated with extracting text from images using Optical Character Recognition (OCR) technology. The emphasis is on the limitations of OCR in considering objects, leading to the suggestion that comprehensive consideration of all elements in an image is necessary for accurate information extraction.

Figure 2 : Original Image[1]

Figure 3 : Gray Scale Conversion[1]

Figure 4 : Detected Text[1]

```
ans =

HANDICAPPED
PARKING
SPECIAL PLATE
REQUIRED
UNAUTHORIZED
VEHICLES
MAY BE TOWED
AT OWNERS
EXPENSE
```

Figure 5: Extracted Text[1]

Similarly **Gala et al. [2],** proposes a method for text reading and extraction from images, coupled with text-to-speech conversion. The process involves gray scale conversion,

MSER feature detection for text region identification, and the application of Optical Character Recognition (OCR) for digitalizing written words. The authors discuss the MSER approach's strength in detecting text regions based on stroke width variation, presenting it as a robust method for text extraction.

Similarly **Farhani et al. [3],** deals into the escalating usage of digital technology and the application of Tesseract for Optical Character Recognition (OCR). It explores various OCR methods, including connected components, sliding window, hybrid, edge-based, color-based, texture-based, and stroke-based approaches. Each method is discussed in terms of its characteristics and applications, providing a comprehensive overview of the diverse techniques in the field.

Similarly **Varma et al. [4],** an effective approach is suggested for text recognition and extraction from images and text to speech conversion. The incoming image is firstly enhanced by employing gray scale conversion. Afterwards, the text regions of the enhanced image are detected by employing the Maximally Stable External Regions (MSER) feature detector. The next step is to apply geometric filtering in combination with stroke width transform (SWT) to efficiently collect and filter text regions in an image. The non-text MSERs are removed by using geometric properties and stroke width transform.

Pawar et al. [5], focus of this article is on proposing a novel approach for word recognition through a digital image-based text recognition technique called LTCS. The LTCS process involves pre-processing images, image segmentation, feature extraction, classification, and post-image processing. The paper addresses issues related to text identification and features, presenting a classification device based on pixel data in digital images. However, the authors note that the efficiency of their approach may vary in terms of timing for digital image conversion.

Similarly **Patil et al. [6],** suggested an effective method for text detection and extraction from photographs, as well as text to audio conversion, is proposed. Gray scale conversion is used to improve the incoming image first. Following that, the improved image's text sections are discovered using the Maximally Stable External Regions (MSER) feature detector. To efficiently gather and filter text sections in a picture, use geometric filtering in

conjunction with the stroke width transform (SWT). Geometric attributes and the stroke width transform are used to remove non-text MSERs.

Tamilselvi et al[7], a Logical Text Classification Strategy (LTCS) is introduced to perform an effective text recognition process using digital images. The proposed LTCS process the input image based on certain characteristics such as: Image Pre-Processing, Segmenting the Image, Extracting the Features, Classification Principle and the Image Post-Processing.

Saravanam et al[8], Introduced GPT-3 model is an AI that takes a string of text as a prompt (input) and purposefully predicts which word or a string of text should or is most likely to come next. To achieve this, OpenAI had GPT-3 learn from billions of information in words across the internet, within news articles, forum posts, websites, etc.

Table 1: Literature surveys with results and limitations

S. No.	Paper Title [Cite]	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Results / Observations	Limitations
1.	Text Recognition from an Image [1]	International Journal of Engineering Research and Application (2014)	OCR(Optical Character Recognition)	Extract text from image	Does not take objects into consideration
2.	Picture Tales: An Approach for Story Generation Using a Series of Images [2]	Institute of Electrical and Electronics Engineers (2021)	RNN , CNN	Descriptions aided the model in identifying the theme accurately .	Sometimes generated story are not accurate, cross Verify from the Database.

3.	Image to Text Conversion: State of the Art and Extended Work [3]	Institute of Electrical and Electronics Engineers (2017)	PTT Conversion using preprocessing, edge detection, segmentation, feature extraction,.	Pre-processing improved quality of image which enabled clear extraction of information.	Complex to implement
4.	Text Extraction From Image and Text to Speech Conversion[4]	International Journal of Engineering Research & Technology (2021)	OCR, Javascript, PostgresSQL	Gains maximum productivity in, graphic integrity, multidimensional ity	OCR may struggle with non-standard or less commonly used language
5.	Image to Text Conversion Using Tesseract [5]	IJRESM,2019	OCR, TTS, Tesseract	Different methods of OCR like sliding window based method	OCR may not accurately recognize characters,.
6.	Image Text Extraction and Text-to-Speech Conversion[6]	The International Journal of Creative Research Thought(2019)	MSER , SWT, TTS	Segmenting the textual content by MSER.	Sensitivity to scale and complex parameter tuning required.
7.	A Novel Text Recognition Scheme using Classification Assisted Digital Image Processing Strategy[7]	Institute of Electrical and Electronics Engineers (2022)	Text recognition, Text Segmentation, Text Enhancement	Efficiently recognizes text in images and eliminates the noise level over the pixel points.	Text segmentation algorithms may sometimes create too many segments, breaking down words or characters excessively.

8.	GPT-3 Powered System for Content Generation[8]	Institute of Electrical and Electronics Engineers (2022)	OpenAI GPT-3	Produced non-repetitive stories, although only to an extent, served to be effective.	Language model has limited input and output sizes

2.2 Key Gaps in the Literature

The works done in the past in this field is highly dependent on the quality and quantity of the training data available which limits its overall performance. The data being centralized, if altered externally intentionally or unintentionally or the data being noisy or the quantity being limited, can negatively impact the accuracy and performance of the models. Some proposed methods (such as faster R-CNN[12]) could be computationally expensive and thus cannot be utilized in real-time applications where the hardware resources are limited. The proposed methods may not be suitable in real world scenarios where there is a constant variation in the climatic and environmental conditions. Lack of proper privacy mechanism can cause sharing of sensitive user information and can be challenging.

CHAPTER 3: SYSTEM DEVELOPMENT

3.1 REQUIREMENTS AND ANALYSIS

Software Resources	• Node.JS (JavaScript Runtime Environment) • NPM (Node Package Manager) • Chrome v116+

Hardware Resources	• A computer with
	  ○ i5 6th Gen+ CPU
	  ○ Integrated GPU
	  ○ 8 GB RAM
	  ○ Windows 10 or above

3.2 PROJECT DESIGN AND ARCHITECTURE

Project Design:

The proposed federated learning system will consist of four main components:

- Research and evaluation of use cases and edge points.
- Building and Testing of G.A.N model based on the dataset.
- Designing Restful API's to integrate with the model.
- Designing & Integration of the Frontend with the backend.

Project Architecture:

The primary aim of this university project is to develop a sophisticated application capable of seamlessly transforming uploaded images into captivating stories. This involves a multi-faceted process where the image is first uploaded to an S3 bucket, its link is sent to an AI model, and the resulting description and text are utilized for story generation based on selected genres and additional prompts.

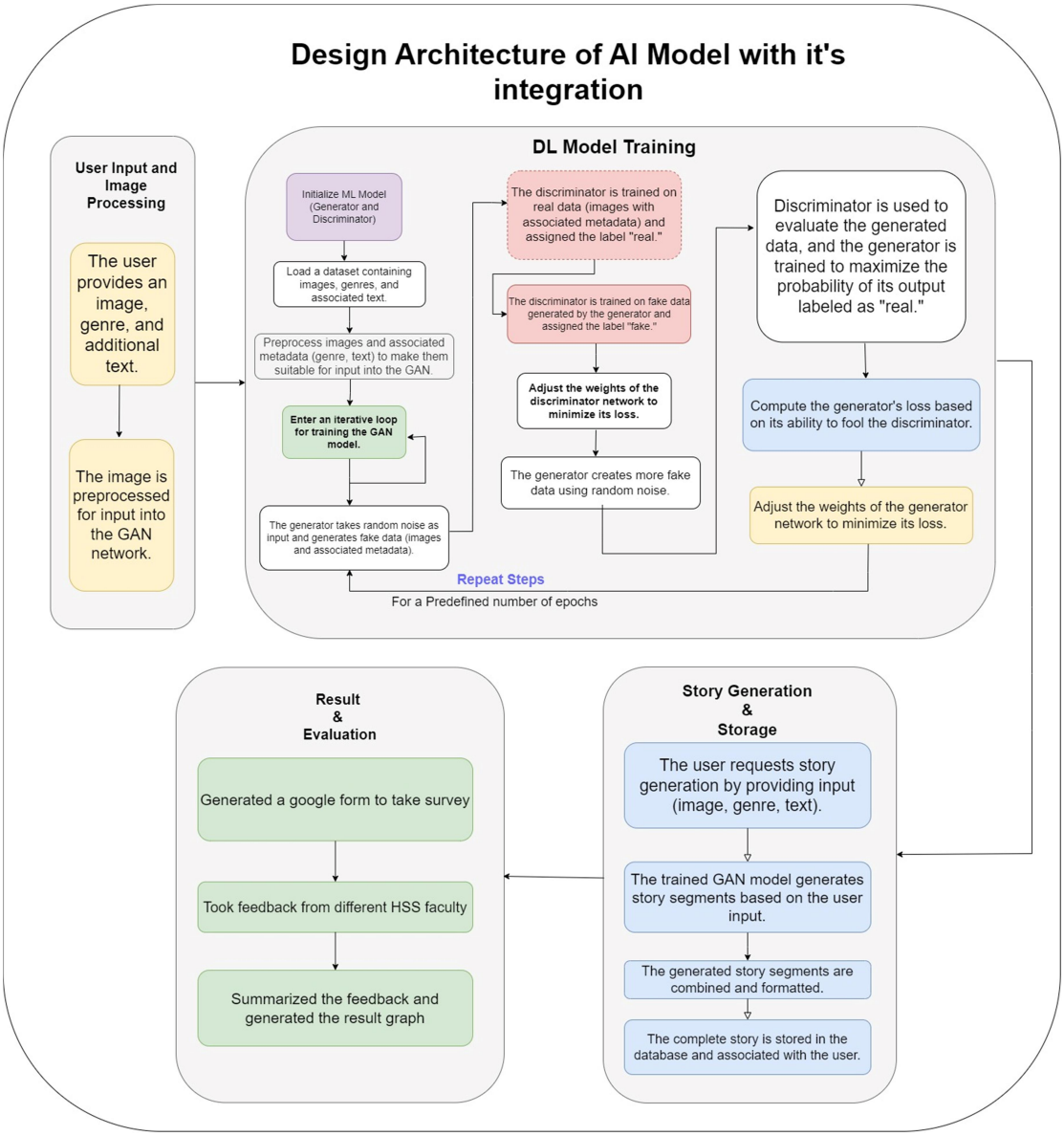


Fig 6: Model architecture

The flowchart is divided into 3 major phases:

1. User Input and Image Processing

2. DL Model Training

3. Story Generation with data storage

Phase 1: User Input and Image Processing

First a user uploads the image and then chooses one genre while offering additional input of extra relevant information or text. Therefore, an essential phase that involves the preprocessing of this user input prior to its incorporation into the GAN network becomes crucial. It is expected that all changes like scaling and normalization of all images along with others modifications would go through it as they have to become consistent and congruous with neural networks in this case.

Phase 2: GAN Network Training

- **Initialization and Data Loading:** Therefore, before starting the training of the GAN, both the generator and the discriminator network need to be initialized. Then the text along with genres, and images are loaded into a dataset for more pre-processing. This loaded data is the basis of the learning process underlying the GAN.
- **Preprocessing of the Data and Parameter Configuration:** Preprocessing ensures that the loaded photographs and metadata conform to the GAN for that reason. The initialization of important training parameters like learning rate, batch size, and epochs helps in guiding the overall process flow of training.
- **GAN Training Loop:** A GAN then enters a continuous loop where it produces synthetic data with the help of the generator, trains the discriminator on both genuine and false data, adjusts the weights of their parameters to minimize losses. This loop enhances the GAN's ability to generate realistic and situational appropriate data.
- **Discriminator Training:** The discriminator is educated through used and genuine information in order for it to distinguish between real and manufactured materials. These losses are used to adjust the discriminator weights and thus increase the accuracy of the system.
- **Generator Training:** The generator is simultaneously trained to maximize the likelihood that the discriminator will classify its output as "real". The discriminator and generator engage in combative play that enhances the generator's ability to generate genuine material.

- **Continuous Improvement:** Steps 3 to 5 are repetitively executed for a predefined number of epochs, allowing the GAN to continually improve its performance. The generator learns to generate diverse and contextually relevant story segments.
- **Saving Trained Model:** Upon completion of the training process, the GAN model, now equipped with learned patterns, is saved for utilization in the subsequent story generation phase. The trained model encapsulates the knowledge acquired during training, enhancing its ability to generate coherent stories based on user input

Phase 3: Story Generation and Storage

After GAN training, the user can request story generation by providing input. The trained GAN model then performs inference, generating story segments based on the user's input. These generated story segments are integrated, formatted, and subsequently stored in a database, associated with the respective user. This ensures that the generated stories can be retrieved and presented to the user upon request.

Now we expand more on the data preparation part.

3.3 DATA PREPARATION

We have used a combination of 2 datasets for the implementation purpose of the project i.e., MSCOCO and COCO128 and a local dataset.

- **MSCOCO:** Microsoft COCO dataset is a pillar of computer vision world because it contains thousands of photos with detailed descriptions related to different visual problems solving skills. It is no exaggeration that this dataset, which is jointly put together by Microsoft researchers and some academics, has turned out to be one of significant importance in imaging caption, object detection, and even **image segmentation[**18].

  Indeed there are enormous and varied pictures embedded in MSCOCO that represent actual life scenes and items. The dataset is made up of over 328 thousand images and it stands out as one which concentrates on capturing common situations. Images accompanied by several human produced captions intertwine to create diverse multi-lingual labels matching the visual data. MSCOCO proves how strong is the cooperation between the people working for computer vision. They managed

to create a huge database of images and give the world a standard which many researchers have used to improve their work about image understanding. The rich annotations, various images, setting standards and measuring competition, MSCOCO serves as a base point toward pushing further computer vision's boundaries and innovating with regard to the comprehension of visual grounds.

- **COCO 128:** The COCO 128 dataset constitutes an important reference for the development and assessment of object detection algorithms concerning images. It has 128 high-resolution images that have been manually extracted form COCO dataset with various types of scenes and common objects for autonomous driving vehicles. Each of these images is very carefully tagged with tight boundary boxes and labels corresponding to every observed object in the scene, serving for reference in training and evaluation of object detectors. Compared to its parent dataset which is of a much larger size, the COCO 128 dataset is of a reduced size and thus convenient for initial training validation before further commitment through extensive training using the larger COCO dataset.

  The reason why COCO 128 is suitable for Ai driven story teller lies in its wide variety of reality-based situations. These pictures capture different things like pedestrian, vehicle, traffic signposts and white lines that formulate a route map which is necessary for the information extraction .Additionally, the smaller dataset is easier to train and validate on, allowing for quick experimentation and growth.

Figure 7: COCO dataset [14]

- Along with this we used a local dataset where images are collected from different sources with proper annotation and segregated into different folders i.e. train, test and validation. Each folder is further divided into two sub folders, images and labels. The labelling of the images is done using roboflow annotate.

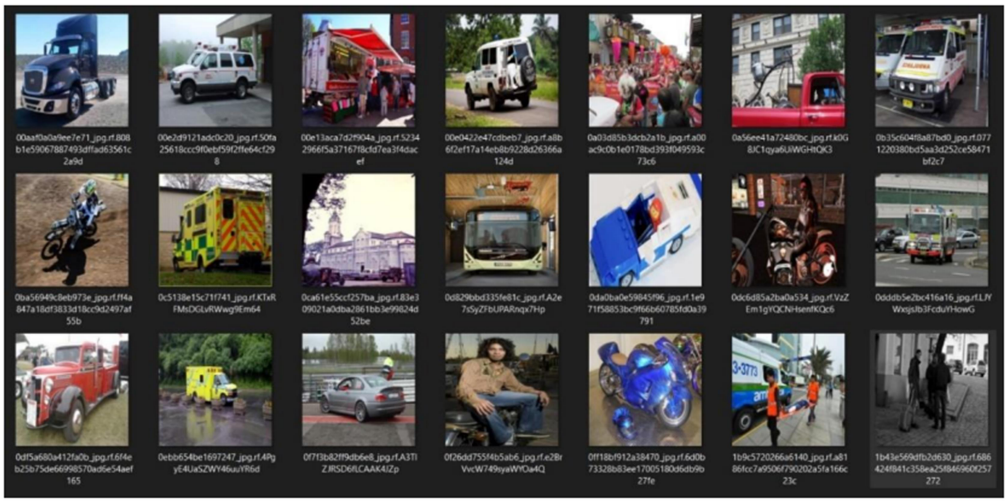


Figure 8: Local dataset

3.4 IMPLEMENTATION

The developmental part of the project is divided into two parts:

- Frontend
- Backend

Frontend

Web application that encompasses user authentication and predicts machine learning and offers useful information which is presentable to users in friendly way easily and clearly. When a user opens an app, it offers him a well elaborated guide from registration through to login.

1. User Authentication[16]

- **Welcome Screen:** On its opening page an online service provides you with a friendly atmosphere and gives a simple opportunity either to log in or create your own account instantly.

We worked on the following processes:

- The login and signup process
- Management of password during login and signup

Encryption: Image uploads to S3 are secured using HTTPS, ensuring the confidentiality and integrity of data during transmission.

Login: During signup process, we take the email from the user and check it against our database. If the email is not present then we continue with the process, otherwise we throw error.

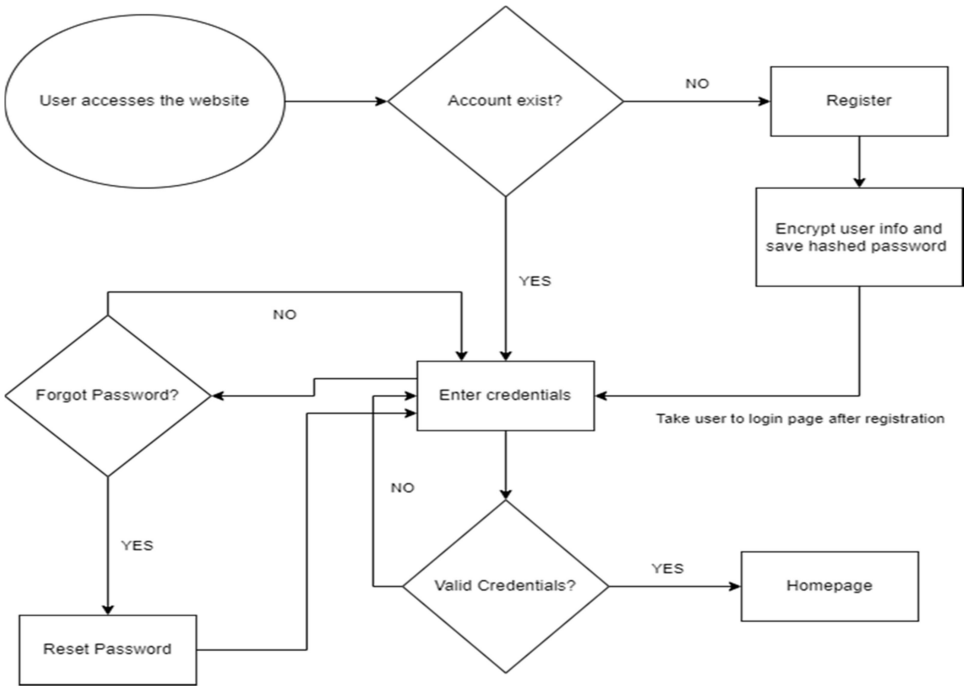


Figure 9: Flowchart of Login and Sign-up[17]

The user accesses the website, after this it is checked whether the account already exists on the system or not.

- If **YES,** then the user is asked to provide credentials and it is cross checked from the database, if the credentials are found valid the user will be redirected to the homepage.
- If **NO,** then the user is asked to register and this information is encrypted and save hashed passwords. Then the user is again redirected to registration page.

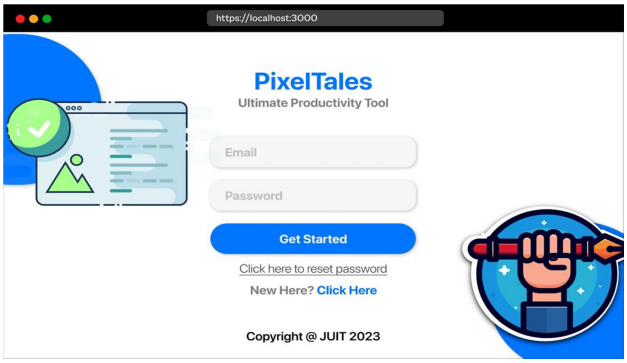


Figure 10 : Welcome screen design

- **Login Process:** The site provides a login form that allows previously signed up users to key in their email address and passwords. As they tried to fulfil their desire for The web app carries out a robust authentication of the credentials against the back end REST API database.
- **Signup Process:** Creating an account only requires a user's email address, preferred username, and a strong password. The web application on the other hand stores the users information in an encrypted way, safely into the databases.

2. Home Dashboard

Home dashboard is central access point for current sensed sensor information and machine forecasts. It consists of two distinct pages:

• **Image to Story:** The user provides a set of images as input and an extra prompt is available to add specific information that would change the story accordingly. The user can also provide an genre of the story as an input.

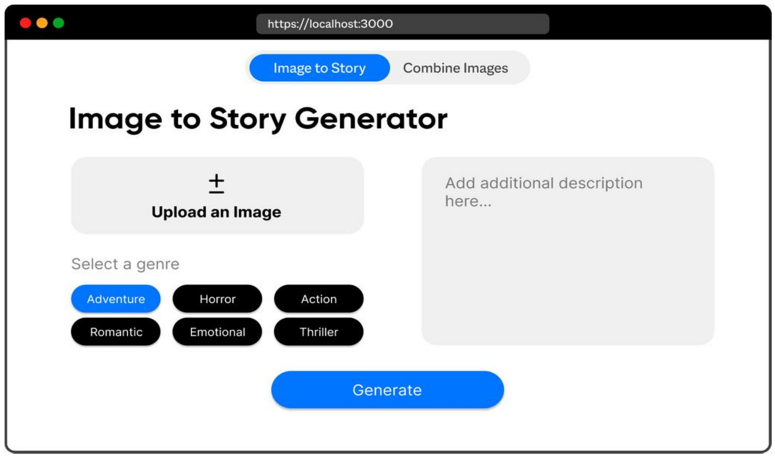


Figure 11: UI implementation
(Image-to-story)

• **Combine Image:** The user gives a set of images and all the images are combined to a single image merging all the characteristics of all the images.

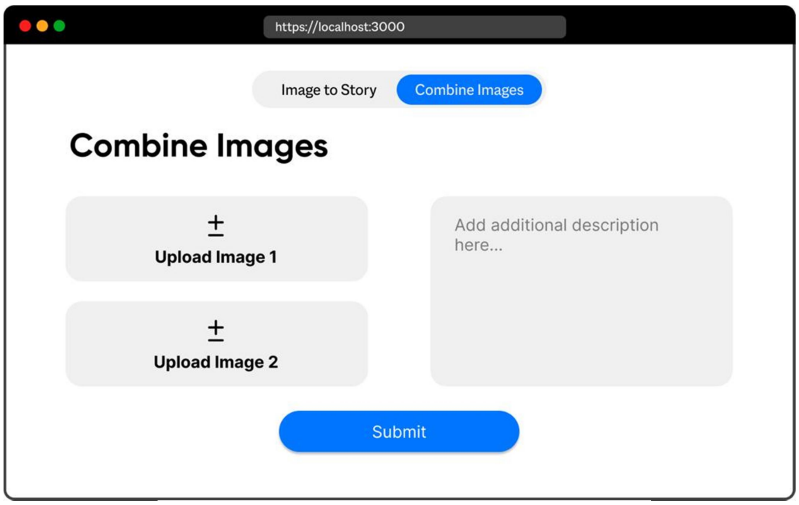


Figure 12: UI implementation (combine image)

• **Output Window:** The generated story is displayed in the output window.

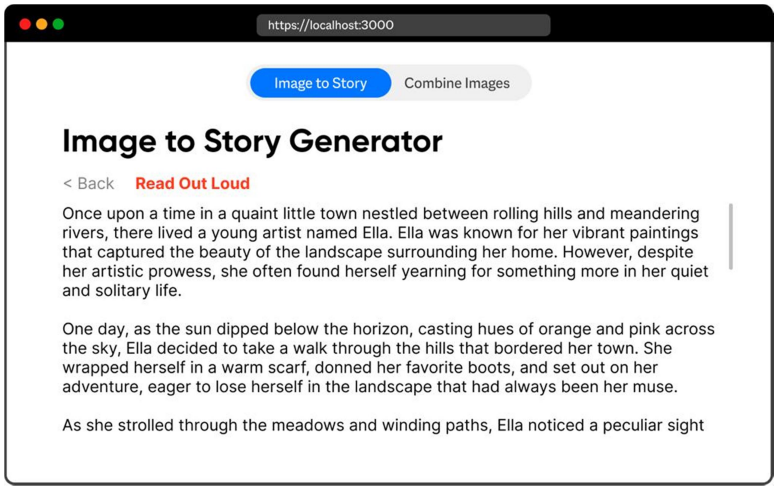


Figure 13: UI implementation (Output)

BACKEND:

- Backend is responsible for handling image processing, communicating with the AI model, and orchestrating the story generation process.

```
pip install ultralytics
```

- **Executing the model (YOLO) [15]**

```
from ultralytics import YOLO

model = YOLO("yolov8n.pt")

model.train(data="coco128.yaml", epochs=5)
metrics = model.val()
```

```
Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
  1/5         0G      1.166      1.461      1.243        294        640: 100%|██████████| 8/8 [02:01<00:00, 15.19s/it]
            Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 4/4 [00:38<00:00,  9.54s/it]
              all        128        929      0.629      0.549      0.596      0.443

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
  2/5         0G      1.219      1.463      1.271        251        640: 100%|██████████| 8/8 [01:52<00:00, 14.08s/it]
            Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 4/4 [00:39<00:00,  9.76s/it]
              all        128        929      0.623      0.577      0.616      0.462

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
  3/5         0G      1.123      1.374      1.215        158        640: 100%|██████████| 8/8 [01:53<00:00, 14.22s/it]
            Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 4/4 [00:37<00:00,  9.36s/it]
              all        128        929      0.651      0.575       0.64      0.478

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
  4/5         0G      1.131      1.292      1.218        256        640: 100%|██████████| 8/8 [01:54<00:00, 14.36s/it]
            Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 4/4 [00:38<00:00,  9.53s/it]
              all        128        929      0.679      0.604      0.654      0.488

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
  5/5         0G      1.181      1.369      1.239        248        640: 100%|██████████| 8/8 [01:57<00:00, 14.72s/it]
            Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 4/4 [00:38<00:00,  9.59s/it]
              all        128        929      0.691      0.584      0.656      0.488
```

```
results = model("https://static1.squarespace.com/static/5de4842Trade+100_0452.jpg")
```

```
Downloading https://buffer.com/cdn-cgi/image/w=1000,fit=contain,q=90,f=auto/library/content/images/si
100%|██████████| 105k/105k [00:00<00:00, 46.9MB/s]
image 1/1 /content/free-images.jpg: 448x640 1 person, 1 cell phone, 336.2ms
Speed: 4.3ms preprocess, 336.2ms inference, 4.5ms postprocess per image at shape (1, 3, 448, 640)
```

Step 1: Starting the server:

Implemented a configuration of Express server for handling and processing requested data by specifying various routes and middlewares for effective handling. • Middleware: Used of Express middleware to parse JSON and carry out the Cross-Origin Resource Sharing

(CORS) function, thereby establishing the backend security and interoperability. • Routes: Well-defined routines for imaging upload, AI model communication, and story creation and to have a neat back-end arrangement. • Environment Variables: Secured storing of sensitive info with dotenv library boosting the app's stance on security. • Error Handling: Include extensive error handling procedures to guarantee strong resilience to failures and graceful degradation when failures occur.

```
jagritkamra@Jagrits-MacBook-Air backend % npm run dev

> backend@1.0.0 dev
> nodemon index.js

[nodemon] 3.0.1
[nodemon] to restart at any time, enter `rs`
[nodemon] watching path(s): *.*
[nodemon] watching extensions: js,mjs,cjs,json
[nodemon] starting `node index.js`
Listening on port:4000
Write a Action story on Two Astronaut, Mars, walking around,spacesuits with addons such asAdd something such as Alien Attack
```

Step 2: Imports secrets and sends data to our model

```
import dotenv from "dotenv";

dotenv.config();

const apiUrl = process.env.API_URL;
const apiKey = process.env.API_KEY;

const ModelCaller = async (ModelPrompt) => {
  console.log(ModelPrompt);
  const requestData = {
    model: "gpt-3.5-turbo",
    messages: [{ role: "user", content: ModelPrompt }],
    temperature: 0.7,
  };
```

```
  try {
    const response = await fetch(apiUrl, {
      method: "POST",
      headers: {
        "Content-Type": "application/json",
        Authorization: `Bearer ${apiKey}`,
      },
      body: JSON.stringify(requestData),
    });

    const data = await response.json();
    const assistantMessage = data.choices[0].message.content;

    return assistantMessage;
  } catch (error) {
    console.error(error);
    return "Error 400";
  }
};

export default ModelCaller;
```

Step 3: Generates the output from our model of specific genre defined in “Genre” variable

```
import express from "express";
import dotenv from "dotenv";
import ModelCaller from "./temp.js";

dotenv.config();

const PORT = process.env.PORT || 6000;

const app = express();

app.use(express.json());

app.get("/generate", async (req, res) => {
  const Genre = "Romantic";
  const ModelDescription =
    "Honeybee, Cute Girl, Walking Dead,lipstick, lamborghini";
  const ModelAddons = "Brutality, Blood";

  //Formatted String
  const ModelPrompt =
    "Write a " +
    Genre +
    " story on " +
    ModelDescription +
    "with addons such as" +
    ModelAddons;
  const result = await ModelCaller(ModelPrompt);
  res.json(result);
});

app.listen(PORT, () => {
  console.log("Listening on port:" + PORT);
});
```

Step 4: Doing a GET request to our API endpoint using Postman.

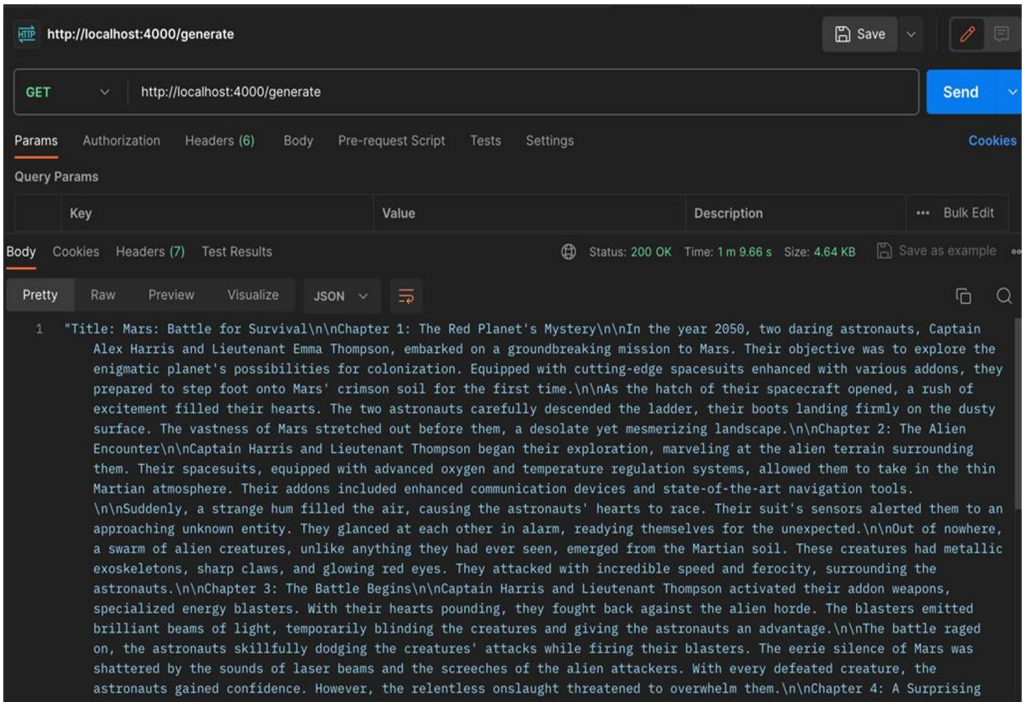


Image Upload (AWS S3): S3 Bucket Setup: Established an S3 bucket to serve as a secure repository for storing and retrieving uploaded images. • AWS SDK Usage: Leveraged the AWS SDK to interact with S3, managing access credentials securely and facilitating seamless integration. • Pre-signed URLs: Implemented the generation of pre-signed URLs, ensuring secure and temporary access for image uploads.

3.5 KEY CHALLENGES

- **Semantic Understanding:** Computers can do so many things, however, it doesn't know what occurred in the image to tell a story.
- **Narrative Structure:** It's still quite difficult to show these pictures in such a way they would attract attention, but tell a coherent story.
- **Contextual Inference:** In addition, the computer should have a tale about whether those persons in the picture are happy or not.
- **Diverse Image Types:** However, a computer cannot tell whether these stories are true or original.
- **Language Generation:** Reading should not become boring when they realise that a machine wrote the story. The story should also be exciting enough as one reads through it.

CHAPTER 4: TESTING

4.1 TESTING STRATEGY

The primary objective of the testing strategy are to:

- Ensure individual components work independently and produce expected results.
- Validate the overall user experience and functionality under realistic conditions.
- Streamline repetitive and critical testing processes for efficiency and consistency.

Testing Approach:

The testing procedures will comprise both a simulation as well as a conventional testing technique.

1. SIMULATION TECHNIQUE

- **Edge Case Testing:**
  - Check whether the app is able to manage rare or strange cases.
  - Try using weird image formats, bizarre user inputs, and irregular use scenarios.
- **Unit Testing:**
  - Test individual components and functions in isolation to ensure they produce the expected output.
  - Focus on testing image-to-text conversion, GAN integration, and other core functionalities separately
- **Usability Testing:**
  - Assess the overall usability of your Application.
  - Ensure you focus on intuitive navigation, clear commands, as well as user friendly interfaces.

- **Regression Testing:**
    - It is also advisable to run tests to confirm that the functionalities that were added or had upgrades remain intact after such changes are implemented.
    - Avoid introducing unintended side effects.

2. TESTING PLAN

- **End-to-End Testing:**
    - Simulate a range of user scenarios over the entire life cycle of your applications.
    - This includes picture submission, user-input addition, and story verification.
- **Performance Testing:**
    - Responsiveness and scalability of the system in different workloads.
    - Analyse the images' processing time, story creation process, and so on that are necessary for other key tasks.
- **User Acceptance Testing (UAT):**
    - The last thing is to involve genuine users so that they can test the application and offer some feedback.
    - They assess user satisfaction, easiness of usage, and entire UX.
- **Security Testing:**
    - Make sure to point out and fix possible gaps in the security, particularly if you process user information.
    - Ensure image uplaods and users interactions are secure.
- **Integration Testing:**
    - Assess how various modules work together in your system.
    - Make sure that elements like image to text conversion and GAN integration are smoothly compatible with each other.

CHAPTER 5: RESULT AND EVALUATION

5.1 RESULT

The project aims to address the mentioned issues and objectives by developing an AI-driven storyteller that turns photos into engrossing narratives. The ultimate objective is to further image-to-text generation research while offering a wide audience an engaging and participatory storytelling experience

We aim to analyze the overall performance of the model. The survey will involve collecting feedback from faculty members across various Humanities and Social Sciences (HSS) disciplines.

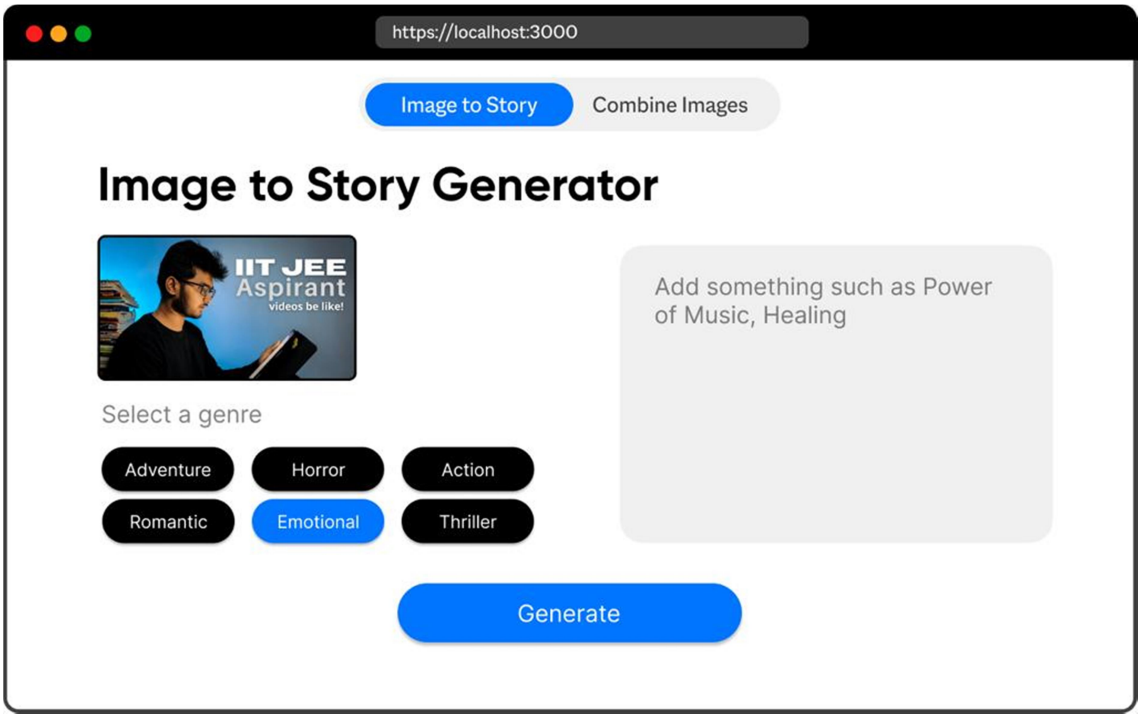
Participants will be asked to evaluate the generated responses in terms of accuracy of:

- Context of the story with respect to the image.
- Literature quality of the story.
- Is the image relevant to the story generated.
- Is the title relevant to the story generated.
- Is the genre provided by user reflected in the story generated.
- How is the quality of image detection.
- Is the "Pacing" and "Coherence" of the story appropriate.
- "Originality" and "Plot complexity" of the story.
- Is the summary related to the story .
- Overall satisfaction of the story generated.

Each response will be assigned a score based on these criteria, and the aggregate scores will be used to calculate the model's performance and accuracy.

Some of the generated responses are:

INPUT 1: Given an image of IIT-JEE aspirant, a story is generated for respective of his journey.



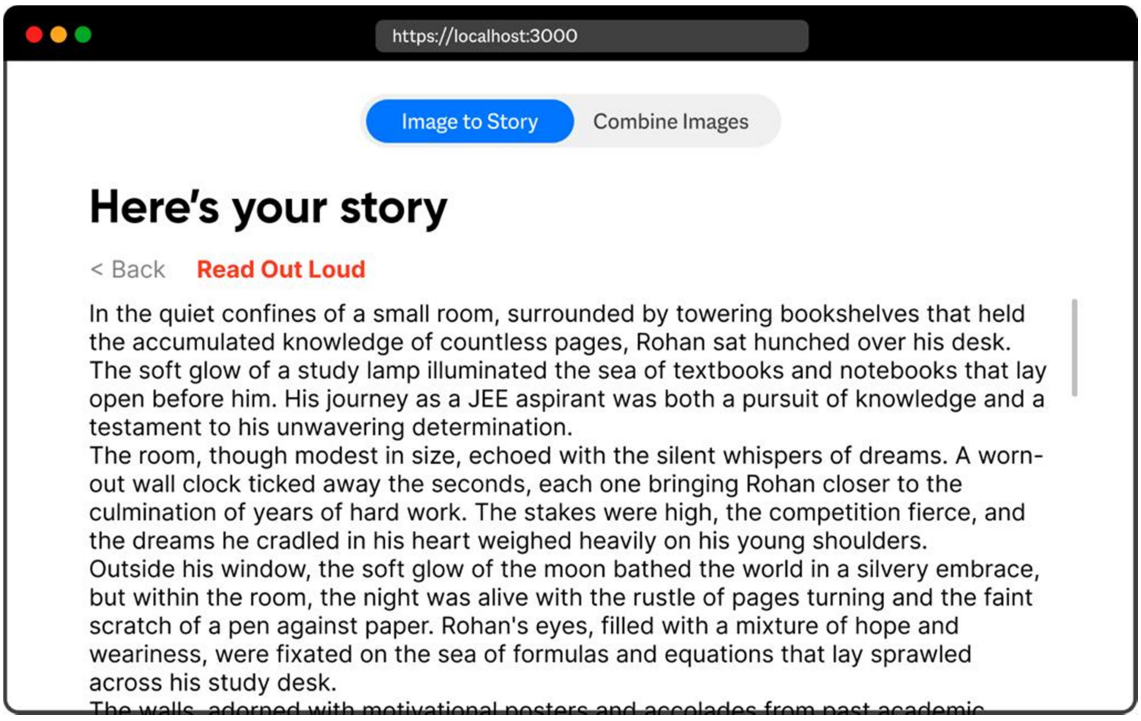
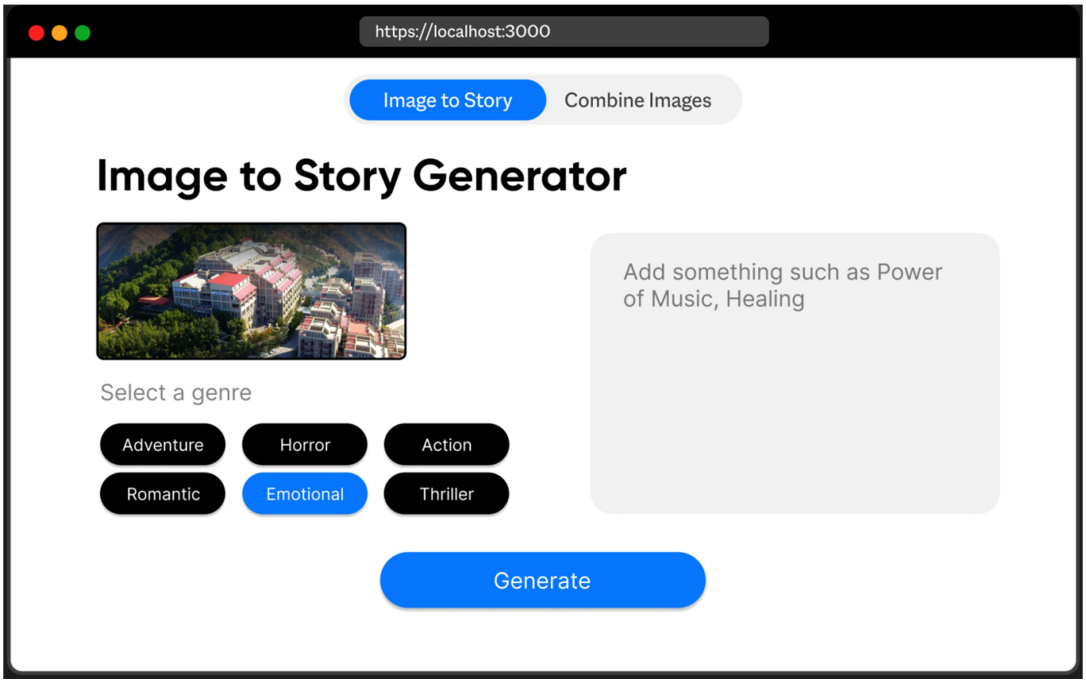


Figure 15 : Story generated about an IIT-JEE aspirant

INPUT 2: Given an image of JUIT, how can a campus turn into story.



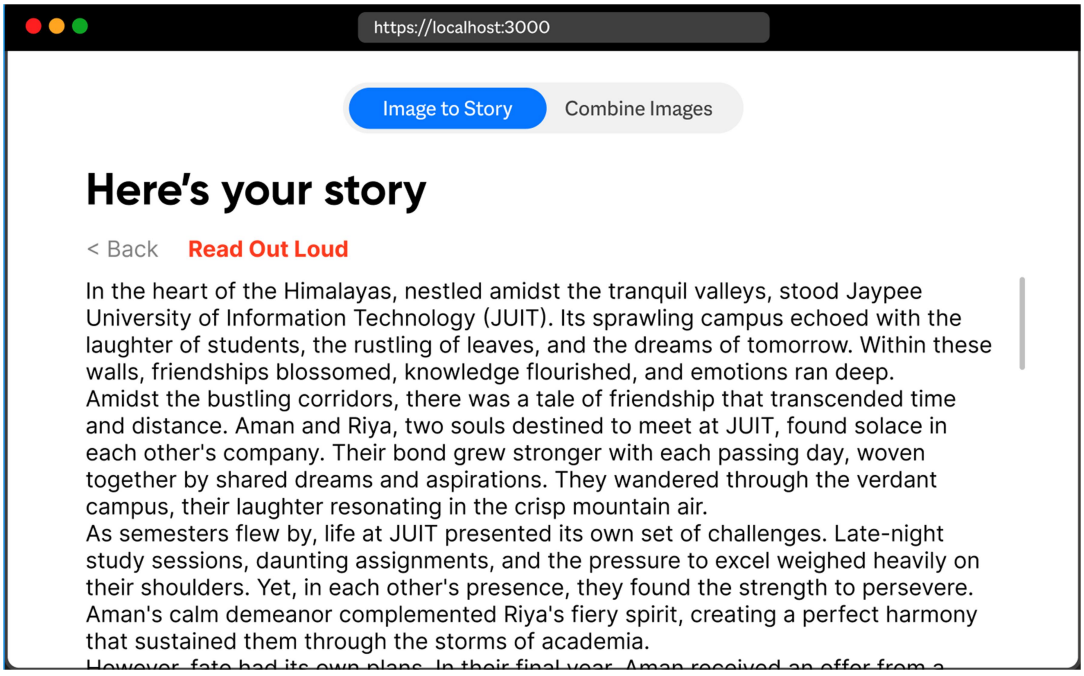
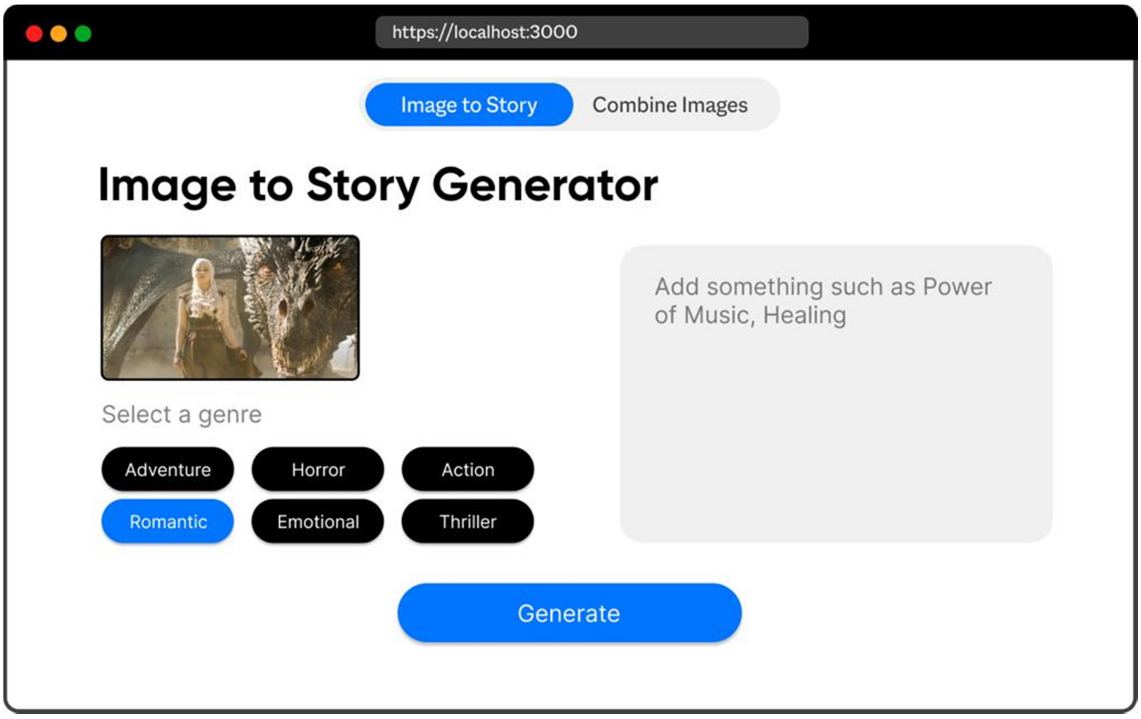


Figure 16 : Story of JUIT

INPUT 3: What if the web-series GOT is written by an AI.



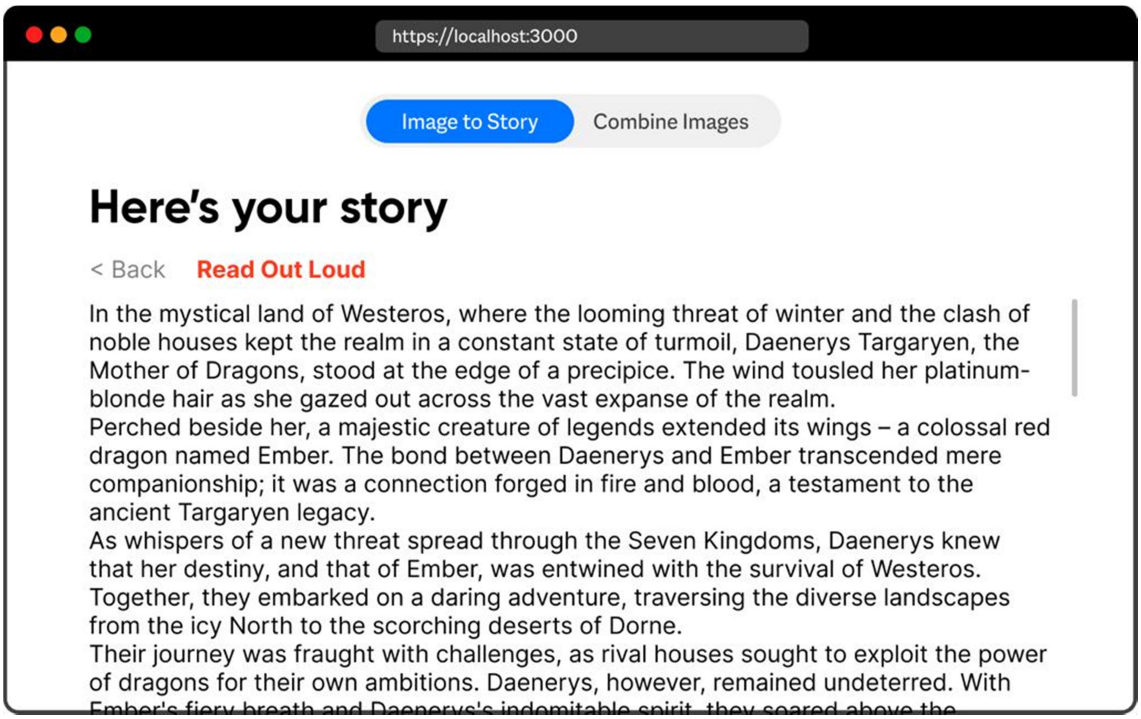


Figure 17: Story of G.O.T

Model Validate:

To generate the results, we administered a feedback form to gather insights from 50 participants, encompassing both teachers from various faculty departments and students. The form incorporated a rating scale ranging from 1 to 5, with 5 indicating the highest rating and 1 denoting the lowest. Using these feedback result we have create graphs that demonstrates our work.

We have shared the form to 50 people out of which 37 people responded us with the feedback.

	Evaluating points	% ratio				
		1	2	3	4	5
1	Context of the story	4	6	10	35	46
2	Literature quality	2	4	13	25	57
3	Image relevancy	2	8	11	31	49
4	Title relevancy	1	10	7	31	57
5	Genre relevancy	4	3	13	29	51
6	Quality of image detection	1	7	9	29	55
7	Pacing & Coherence	1	6	11	30	52
8	Originality & Plot Complexity	1	11	9	26	54
9	Summary	5	7	7	36	45
10	Overall Satisfaction	1	8	10	27	55

The following are the result generated.

Output 1:

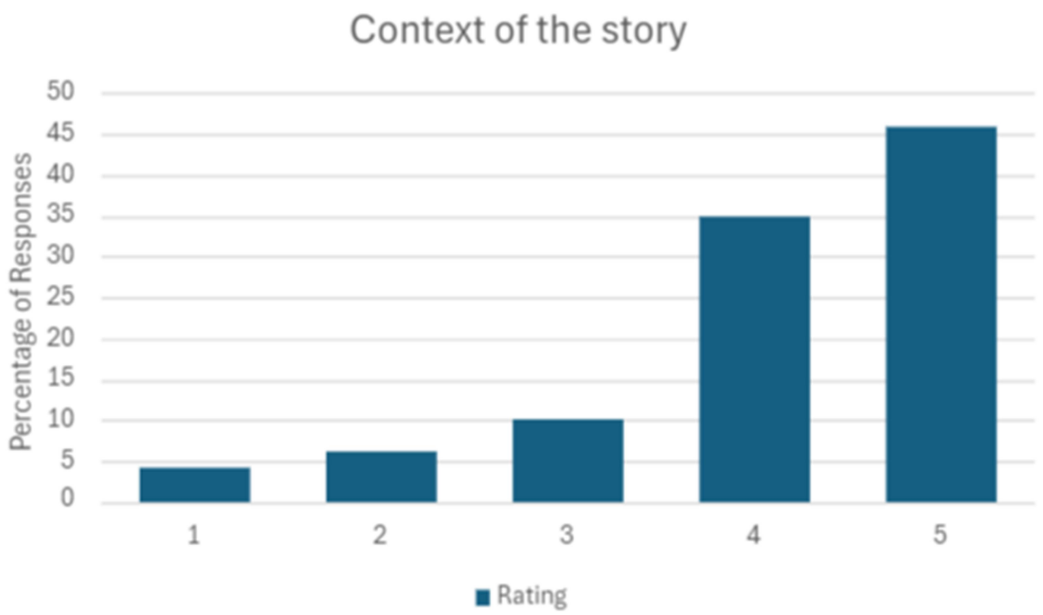


Figure 18: Result of Context of the story

Output 2:

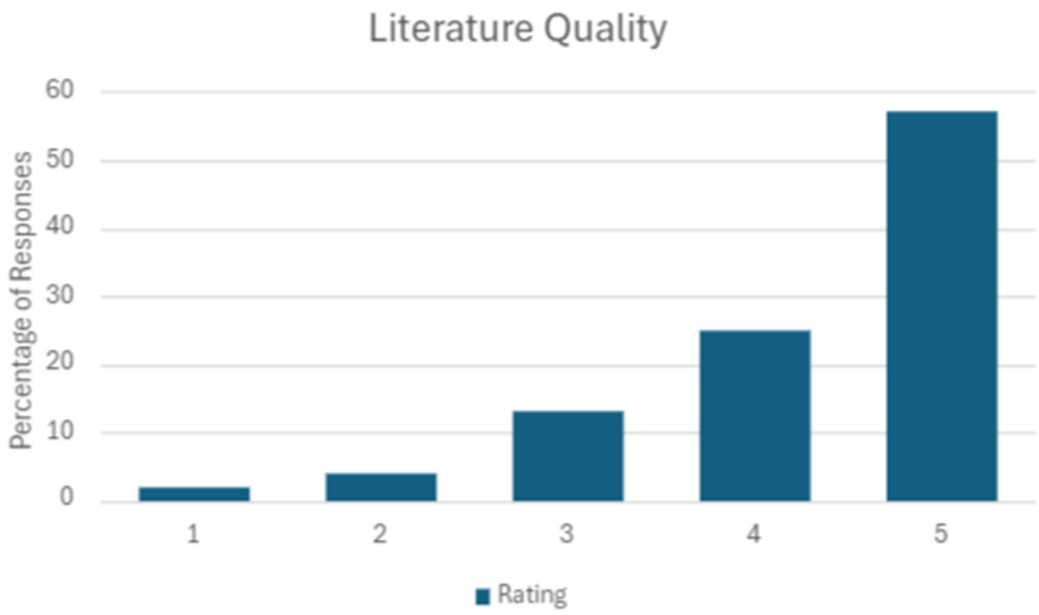


Figure 19: Result of Literature quality of the story

Output 3:

Figure 20: How much is TITLE relevant to the story

Output 4:

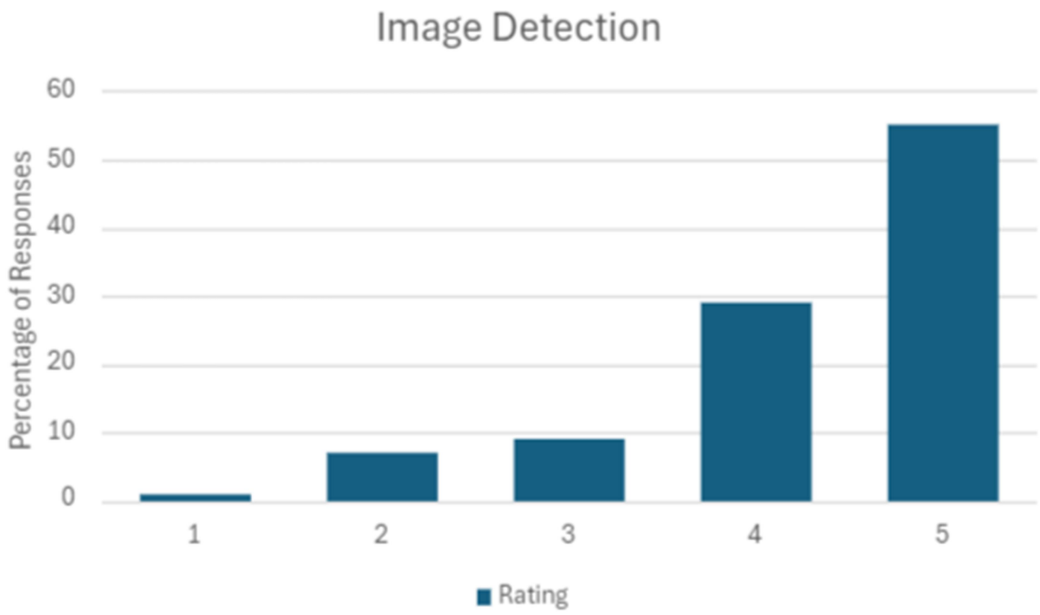


Figure 21: How good was the image detection.

Output 5:

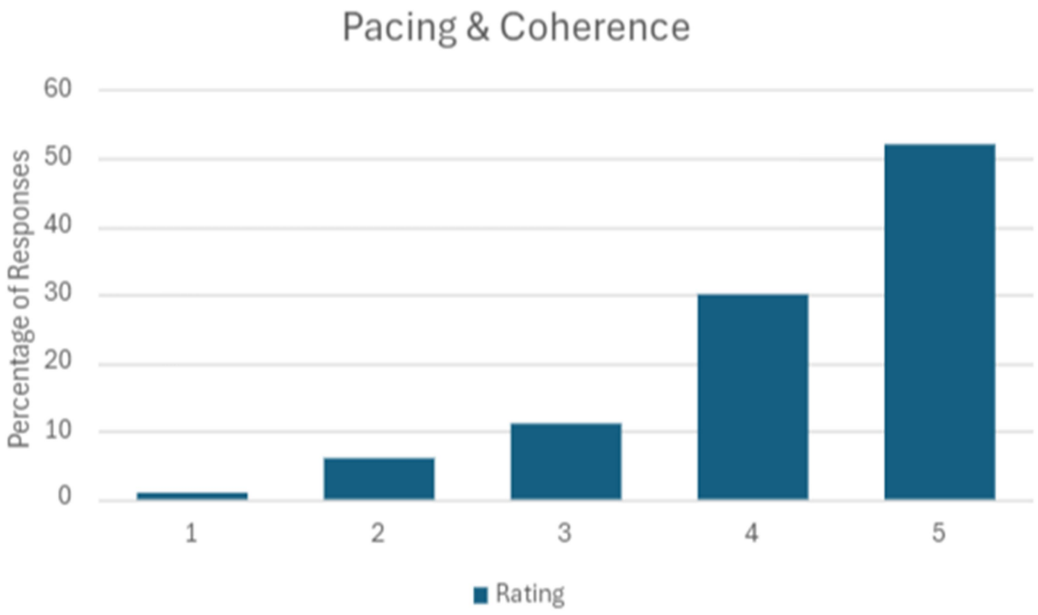


Figure 22: How good is the pacing & coherence of the story

Output 6:

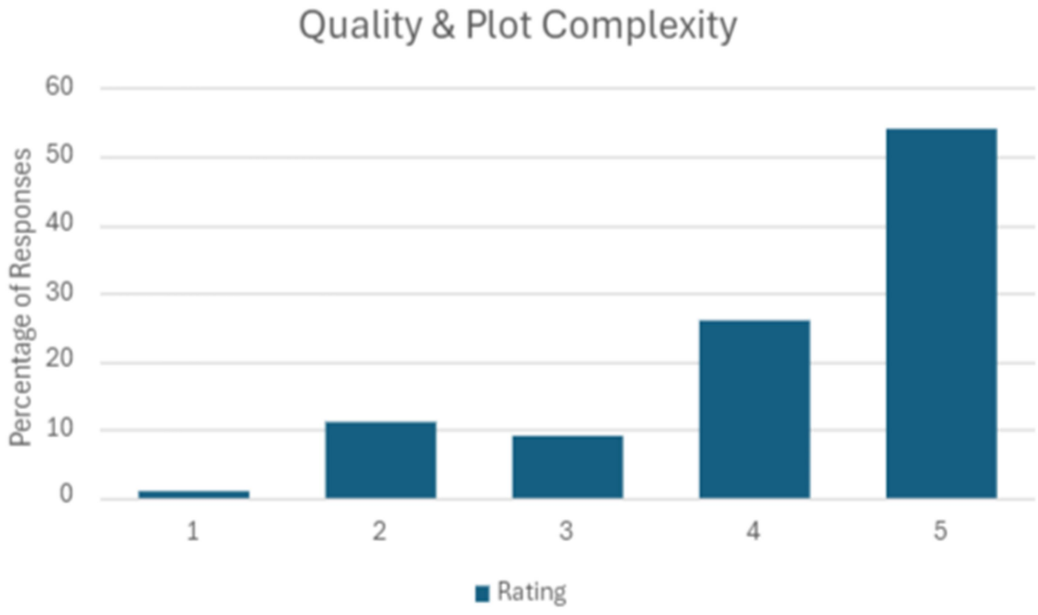


Figure 23: How good is the quality & plot complexity to the story

Output 7:

Figure 24: How much is GENRE relevant to the story

Output 8:

Figure 25: How much is IMAGE relevant to the story

Output 9:

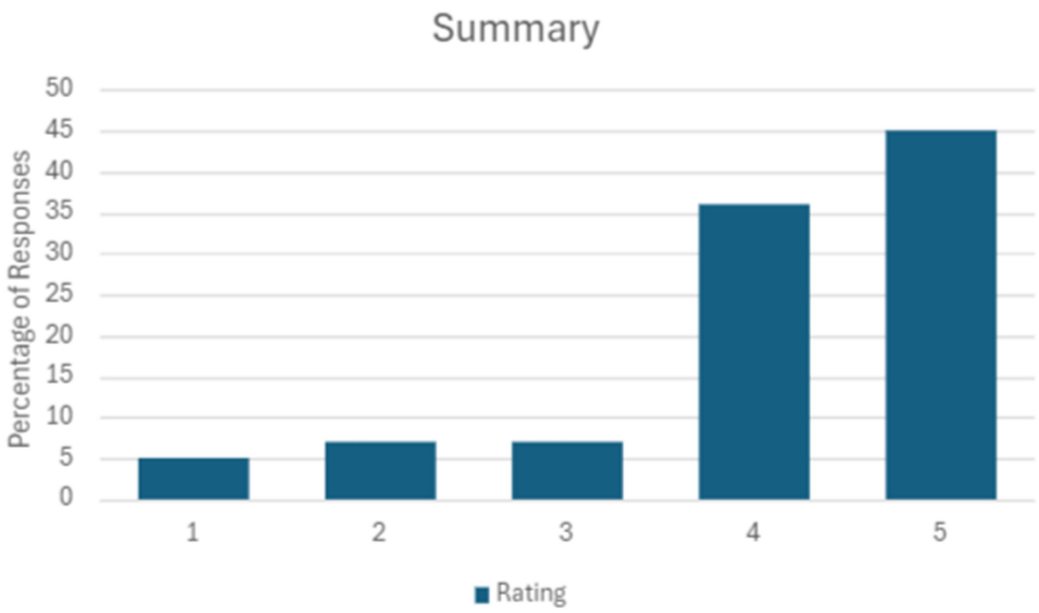


Figure 26: The summary of the story

Output 10:

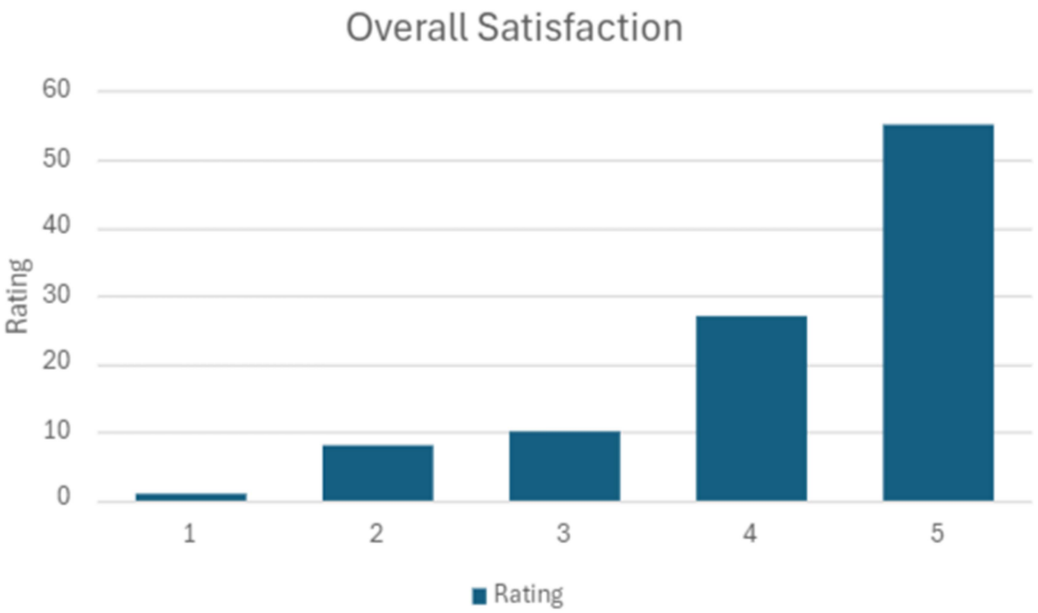


Figure 27: The overall evaluation of the story

The survey results show that a majority of respondents, approximately 53%, rated it as "very good," while 32% rated it as "good." This indicates widespread satisfaction among survey participants. Overall, the survey underscores the project's success and validates its ability to meet the expectations of its audience.

CHAPTER 6: CONCLUSION AND FUTURE SCOPE

5.1 CONCLUSION

The development of an artificial intelligence storyteller from pictures is a challenging task due to the need for skills in image understanding, putting together consistent stories, emotion reading, and adjustment to different visual cases. Multi-faceted approach adopted by the project aims to enrich the user experience and involvement with entertaining stories. Furthermore, the aim entails fun as well as learning and makes the process inclusive, allowing anyone to participate in storytelling. Also, the initiative seeks to enhance imaginative artistry, leading to uniquely personal narrative performances customised for specific likes.

5.2 FUTURE SCOPE

This project has significant future scope, and will lay the groundwork for more innovations and influence. The future for image fusion is to extend beyond merging images but rather enhancing the images using more advanced algorithms with greater scope for creativity. The Text-to-Speech Integration could provide for more natural and expressive voice synthesis ensuring greater accessibility and enriching user experience. Multi-Image to Text Conversion might even explore in depth, extraction of detailed written messages from several images. They are likely to integrate with various educational systems as they continue moving forward.

- **Image Fusion:** Combine one/more photos to result in a new mixed picture.

- **Text-to-Speech Integration[20]:** Develop an application that transforms text to speech thereby making accessibility as well as usability more meaningful.

- **Multi-Image to Text Conversion:** The functionality of the system should also be developed such that multiple images can be converted into textual information, among other applications.-

- **Integration with Educational Systems:** Work together with education systems toward the integration of this application on a user-friendly basis, in order to improve learning.

- **RealTime Collaboration:** The database stores the user data, which later can be used to change the prompt depending on requirements.

REFERENCES

[1] S. Janvalkar, P. Manjrekar, S. Pawar, L. Naik, “Text Recognition from an Image”, Int. Journal of Engineering Research and Application, vol. 5 (pp. 149-151), IJERA (2014) https://www.ijera.com/papers/Vol4_issue4/Version%205/Z04405149151.pdf

[2] K. Gala, M. Somaiya , M. Gopani , “Picture Tales: An Approach for Story Generation Using a Series of Images,“In IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), (pp. 1-5.) IEEE, 2021. https://ieeexplore.ieee.org/document/9573954

[3] N. Farhani, N. TERBEH, M. Zrigui "Image to Text Conversion: State of the Art and Extended Work,“In IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA) (pp. 937-943) IEEE.(2017) . https://www.researchgate.net/publication/323713259_Image_to_Text_Conversion_State_of_the_Art_and_Extended_Work

[4] T. Varma, S. Madari, L. Montheiro, R. Pooojary, “Text Extraction From Image and Text to Speech Conversion”,“International Research Journal of Engineering and Technology ” vol. 9 (pp. 5-9)IJERT (2021) https://www.ijert.org/text-extraction-from-image-and-text-to-speech-conversion

[5] N. Pawar, Z. Shaikh, P. Shinde, "Image to Text Conversion Using Tesseract“,” International Research Journal of Engineering and Technology ” vol. 6 (pp. 516-519) IRJET (2019)https://www.irjet.net/archives/V6/i2/IRJET-V6I299.pdf

[6] K. Sujatha, S. P. Patil, "Image Text Extraction and Text-to-Speech Conversion”,” In international journal of creative research thoughts (IJCRT)” vol. 9 pp. 1-8, 2019 IJCRT. https://ijcrt.org/papers/IJCRT2108410.pdf

[7] M. Tamilselvi , G. Ramkumar, “A Novel Text Recognition Scheme using Classification Assisted Digital Image Processing Strategy”,” In International Conference on Advances in Computing, Communication and

Applied Informatics pp. 1-6, IEEE Xplore 2022 https://ieeexplore.ieee.org/document/9752542

[8] S. Saravanan , K. sudha, "GPT-3 Powered System for Content Generation and Transformation" ,"In 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT) pp. 514-519 IEEE Xplore,2022 https://ieeexplore.ieee.org/document/9913526

[9] A. Kumar, "An efficient text extraction algorithm in complex images" ,"In sixth international conference on contemporary computing (IC3) (pp. 6-12). IEEE (2013).https://ieeexplore.ieee.org/document/6612171

[10] J. Wang, K. Markert, and M. Everingham, "Learning Models for Object Recognition from Natural Language Descriptions," Proc. British Machine Vision Conf., 2009.

[11] K. Bhunesh, M. Abinesh, S. Narmatha, and J. Kannan, "Unveiling the Power of Image to Text Conversion: Advancements, Challenges, and Applications", IJRESM, vol. 6, no. 9, pp. 29–37, Sep. 2023 https://journal.ijresm.com/index.php/ijresm/article/view/2812

[12] R. Chauhan, K. K. Ghanshala and R. C. Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 278-282,https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8703316&isnumber=8703199

[13] A. Mohan, C. Papageorgiou and T. Poggio, "Example-based object detection in images by components," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 4, pp. 349-361, April 2001, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=917571&isnumber=19822

[14] Lin, Maire, M. Belongie, S. Hays, P. Ramanan, D. Dollar, and Konklecker ,"Microsoft COCO : Common Object in Context" ,2014 In Proceeding of the European Conference on computer vision,vol.21, pp. 740-755, Springer 2014.

[15] C. Liu, Y. Tao, J. Liang, K. Li and Y. Chen, "Object Detection Based on YOLO Network," 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 2018, pp. 799-803, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8740604&isnumber=8740347

[16] S. W. Shah and S. S. Kanhere, "Recent Trends in User Authentication – A Survey," in IEEE Access, vol.7, pp. 112505-112519 , 2019, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8784263&isnumber=8600701

[17] M. L. Das, A. Saxena and V. P. Gulati, "A dynamic ID-based remote user authentication scheme," in IEEE Transactions on Consumer Electronics, vol. 50, no.2, pp. 629-631, May2004, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1309441&isnumber=29059

[18] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 7, pp. 3523-3542, 1 July 2022 https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9356353&isnumber=9788494

[19] A. Neumann, N. Laranjeiro and J. Bernardino, "An Analysis of Public REST Web Service APIs," in IEEE Transactions on Services Computing, vol. 14, no. 4, pp. 957-970,1July-Aug.2021, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8385157&isnumber=9507584

[20] C. Thu, T. Zin “Implementation of text to speech conversion ”, in International Journal of Engineering Research & Technology (IJERT), vol. 3 (pp. 2278-2304) 2014, https://www.academia.edu/download/64741818/implementation_of_text_to_speech_coversion_IJERTV3IS030548.pdf

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

PLAGIARISM VERIFICATION REPORT

Date: …….

Type of Document (Tick): PhD Thesis | M.Tech Dissertation/ Report | B.Tech Project Report | Paper

Name: _____ __**Department:** _____ **Enrolment No** _____

Contact No. _____**E-mail.** _____

Name of the Supervisor: _____

Title of the Thesis/Dissertation/Project Report/Paper (In Capital letters): _____

_____

_____

UNDERTAKING

I undertake that I am aware of the plagiarism related norms/ regulations, if I found guilty of any plagiarism and copyright violations in the above thesis/report even after award of degree, the University reserves the rights to withdraw/revoke my degree/report. Kindly allow me to avail Plagiarism verification report for the document mentioned above.

Complete Thesis/Report Pages Detail:

- Total No. of Pages =
- Total No. of Preliminary pages =
- Total No. of pages accommodate bibliography/references =

(Signature of Student)

FOR DEPARTMENT USE

We have checked the thesis/report as per norms and found **Similarity Index** at ……..(%). Therefore, we are forwarding the complete thesis/report for final plagiarism check. The plagiarism verification report may be handed over to the candidate.

(Signature of Guide/Supervisor) **Signature of HOD**

FOR LRC USE

The above document was scanned for plagiarism check. The outcome of the same is reported below:

<table>
<tr><th>Copy Received on</th><th>Excluded</th><th>Similarity Index (%)</th><th colspan="2">Generated Plagiarism Report Details (Title, Abstract & Chapters)</th></tr>
<tr><td rowspan="2"></td><td rowspan="4">• All Preliminary Pages
• Bibliography/Images/Quotes
• 14 Words String</td><td rowspan="2"></td><td>Word Counts</td><td></td></tr>
<tr><td rowspan="2">Report Generated on</td><td>Character Counts</td><td></td></tr>
<tr><td rowspan="2"></td><td>Submission ID</td><td>Total Pages Scanned</td><td></td></tr>
<tr><td></td><td>File Size</td><td></td></tr>
</table>

Checked by
Name & Signature **Librarian**

……

Please send your complete thesis/report in (PDF) with Title Page, Abstract and Chapters in (Word File) through the supervisor at plagcheck.juit@gmail.com

k

ORIGINALITY REPORT

7%	4%	4%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ijert.org Internet Source	2%
2	Shruti Saravanan, K. Sudha. "GPT-3 Powered System for Content Generation and Transformation", 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2022 Publication	1%
3	www.semanticscholar.org Internet Source	1%
4	M. Tamilselvi, G. Ramkumar, G Anitha, P Nirmala, S. Ramesh. "A Novel Text Recognition Scheme using Classification Assisted Digital Image Processing Strategy", 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), 2022 Publication	1%
5	Khushi Gala, Mansi Somaiya, Manvi Gopani, Abhijit Joshi. "Picture Tales: An Approach for	<1%

	Story Generation Using a Series of Images", 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), 2021 Publication	
6	Submitted to Leeds Trinity and All Saints Student Paper	<1 %
7	Submitted to Jaypee University of Information Technology Student Paper	<1 %
8	www.iieta.org Internet Source	<1 %
9	codeculturepro.medium.com Internet Source	<1 %
10	dspace.dtu.ac.in:8080 Internet Source	<1 %
11	pergamos.lib.uoa.gr Internet Source	<1 %
12	www.scitepress.org Internet Source	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		