
English Novel Summarization and Visualizations

App-solutely Mobile

Fatemeh Pesaran Zadeh, Sehee Kim

Contents

- Introduction
- Methodology
- Results
- Conclusion
- Project Management

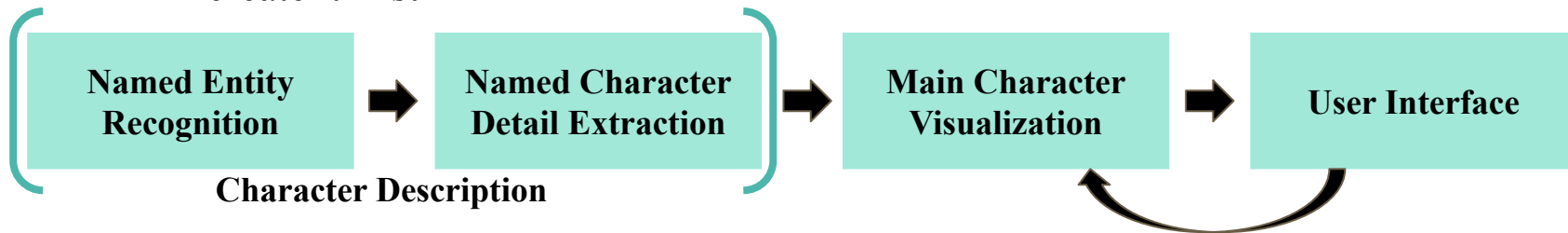
Introduction

Motivation

- More people are studying foreign languages
- Studying foreign languages with visual aids are effective
- We provide a way for visual aids to be more accessible
 - Cost-effective
 - Less time-consuming

Proposed Idea & Approach

- **Proposed Idea:** Provide illustrations for English novels
- **Up to midterm presentation**
 - Stories → Text Summarization → Text-to-Image Generation → Illustrations
- **Current approach**
 - Focused on visualizing the main character
 - Main character's appearance is a key factor when illustrating story
 - It should not change during in between illustrations → Find all information and create it first



Novelty

- Addition of the user interface
 - A console is provided so that the user has a chance to modify the caption after it is provided
 - Does not solely rely on the accuracy of the models
 - Ensures that the captions and the illustrations stay true to the original text

Methodology

Character Description - Named Entity Recognition

- Find main character within story → Extract description of main character
- **Named Entity Recognition**
 - Find main character
 - Fine-tuned Bert model for Named Entity Recognition

Character Description - Main Character Detail Extraction

- **Question Answering**

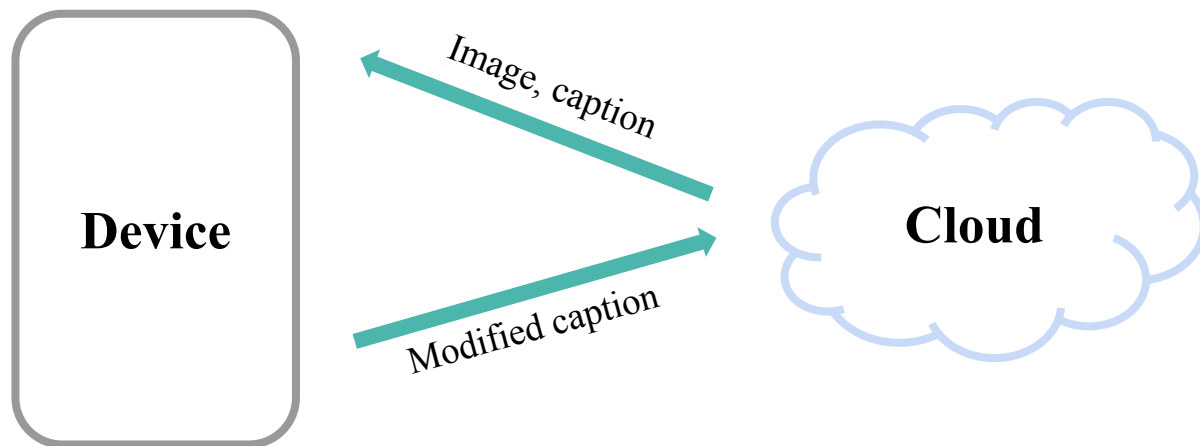
- Extract details about the appearance of main character
- Questions:
 - "Can you describe the main character's appearance?"
 - "Tell me about the main character's physical features."
 - "What are the distinguishing characteristics of the main character?"
- Prompts to guide the question-answering model → extract relevant information
- Gathered information → Concise sentence → Caption

Main Character Visualization

- Used “prompt to prompt image editing with cross attention map” paper approach:
 - Enhanced Image Generation: The diffusion model enables precise editing of generated images for visually appealing results.
 - Flexibility and Adaptability: Easy modification of prompts allows fine-grained control over image generation.
 - Coherence and Consistency: Cross attention maps ensure visual continuity and coherence between prompts and images.
 - Seamless Integration: The approach seamlessly integrates textual prompts with image generation for a cohesive connection.
 - Creative Expression: The technique provides a unique and creative way to visually express the main character's attributes and story.

User Interface

- Information from question answering may not be totally accurate.
- Human involvement to complement information
- Receive caption and created image → Caption modification from user → Send modified caption back to cloud



Results

Visualization based on generated caption.



"Cinderella is Beautiful Kind-hearted, gentle, and hardworking"

User Interface

- Designed for use in mobile device
- **Original Plan:** Connect to cloud and move data back and forth between cloud and device
- **Current implementation:** on device only, video only design, not actual captions or image



Evaluation Results

- Since there is practically no way for quantitative evaluations on this topic, we had to use qualitative methods for evaluation.
- Extraction of physical description:
 - Adequate
 - Does not only extract the exact part
- Visualization:
 - Captures the main character
 - Low resolution images

Challenges

- **Time management**
 - We spent a lot of time on several cases of trial and error
 - Tried different methods such as various models and prompts
- **Changes in team members**
 - Decrease in team members during the progress of the project
 - Difficult to divide the workload

Conclusion

Conclusion

- **Goal:** Provide enhanced language learning experience
- **Tasks:** Named Entity Recognition, Text Summarization, and Image Generation
- **User Console:** Improve accuracy
- **Limitations:**
 - Time constraints
 - Generate only first caption, implement only the design of the console
 - Diffusion model
 - Not fine-tuned on images related to story illustrations
- **Future work**
 - Address limitations

Project Management

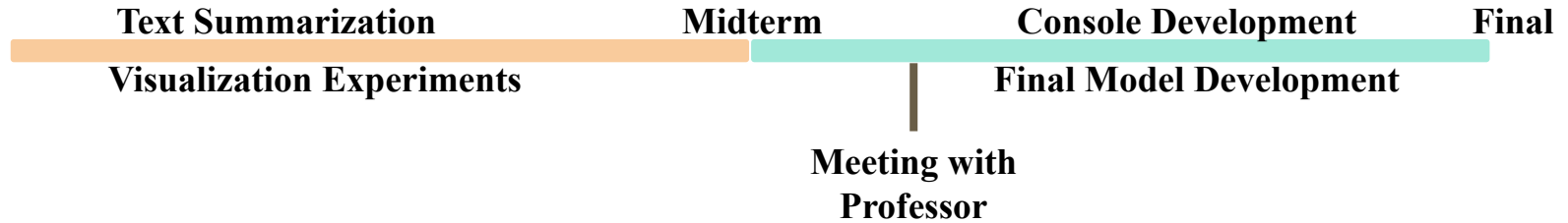
Project Management

- **Scope of the project**
 - **Before midterm:** Focus on building an acceptable text summarization model for story summarization in order to create captions
 - **Current:** Added user involvement to complement inaccuracies
 - Our main task is not text summarization
- **Roles and contributions:** Not strictly separated

	Up to Midterm	Up to Final
Fatemeh	Idea, experimenting on image generation using summaries	Final model development(name entity recognition,extracting features, image generating), Paper
Sehee	Idea, Text summarization experiment, Presentation	Mobile console development, Presentation

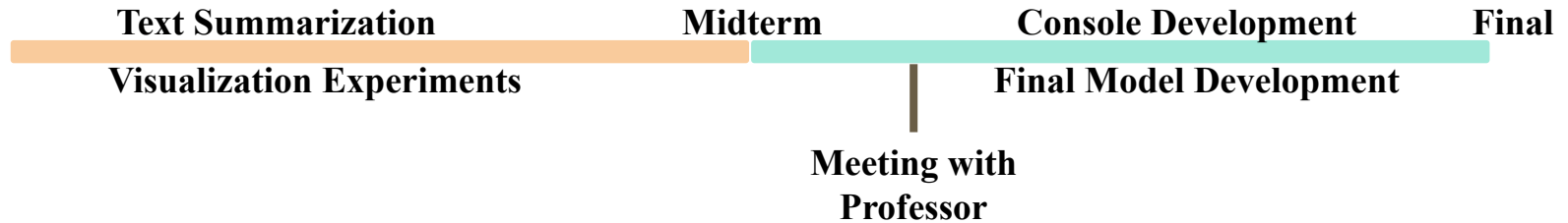
Project Management

- **Timeline**



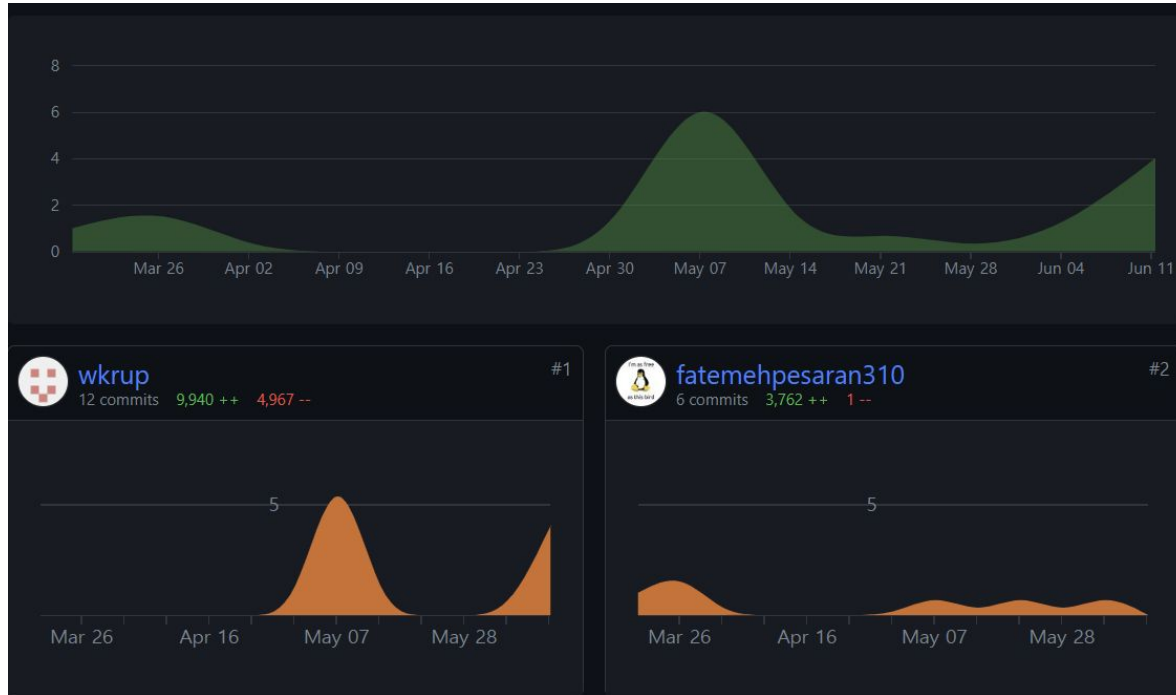
Project Management

- **Timeline**



Project Management

- **Git usage stats:** Contributions to main



Lessons and Reflections

- **Time Management**

- We focused too much on experimentation and spent too much time on this part
- Needed firmer restrictions on designated schedules

- **Preparation for the unexpected**

- Many unexpected events occurred that threw us off track

Thank You for Listening