

# NEWSFLIX



# THE TEAM

**Abhinav Sobi**

**U2103008**

**Alan Joseph**

**U2103021**

**Basil Eldho Joseph**

**U2103057**

**Daniel Robin**

**U2103072**

**Guide**

**Dr. Mary Priya Sebastian**

THE  
**PRESENTATION**  
**FOCUS**

01	Problem definition	21	Assumptions
02	Purpose & need	22	Work responsibilities
03	Project objective	23	Requirements
04	Literature survey	25	Gantt chart
06	Proposed method	26	Budget
07	Architecture diagram	27	Risk & challenges
09	Sequence diagram	28	Outcomes
10	Modules	29	Conclusion
11	Each module in detail	30	References

# “ PROBLEM DEFINITION

To develop a system that can automatically transform news articles into short, engaging videos by incorporating visuals and captions.



# PURPOSE & NEED

Traditional newspaper formats are unfavorable for elderly and for those with visual impairments, limiting accessibility and reach of the news.

News providers can't satisfy the growing demand for digital news content, creating engaging & quick news reels is labor-intensive and time consuming.

# OBJECTIVE

Develop a fully automated system using NLP and AI:



- Transform scanned newspaper articles into engaging, summarized videos.
- From the article extract text, summarize & classify the news.
- Generate video with relevant footage, audio of news summary, and synchronized subtitles.
- Implement a multi-view layout for consuming multiple news videos.

# LITERATURE SURVEY

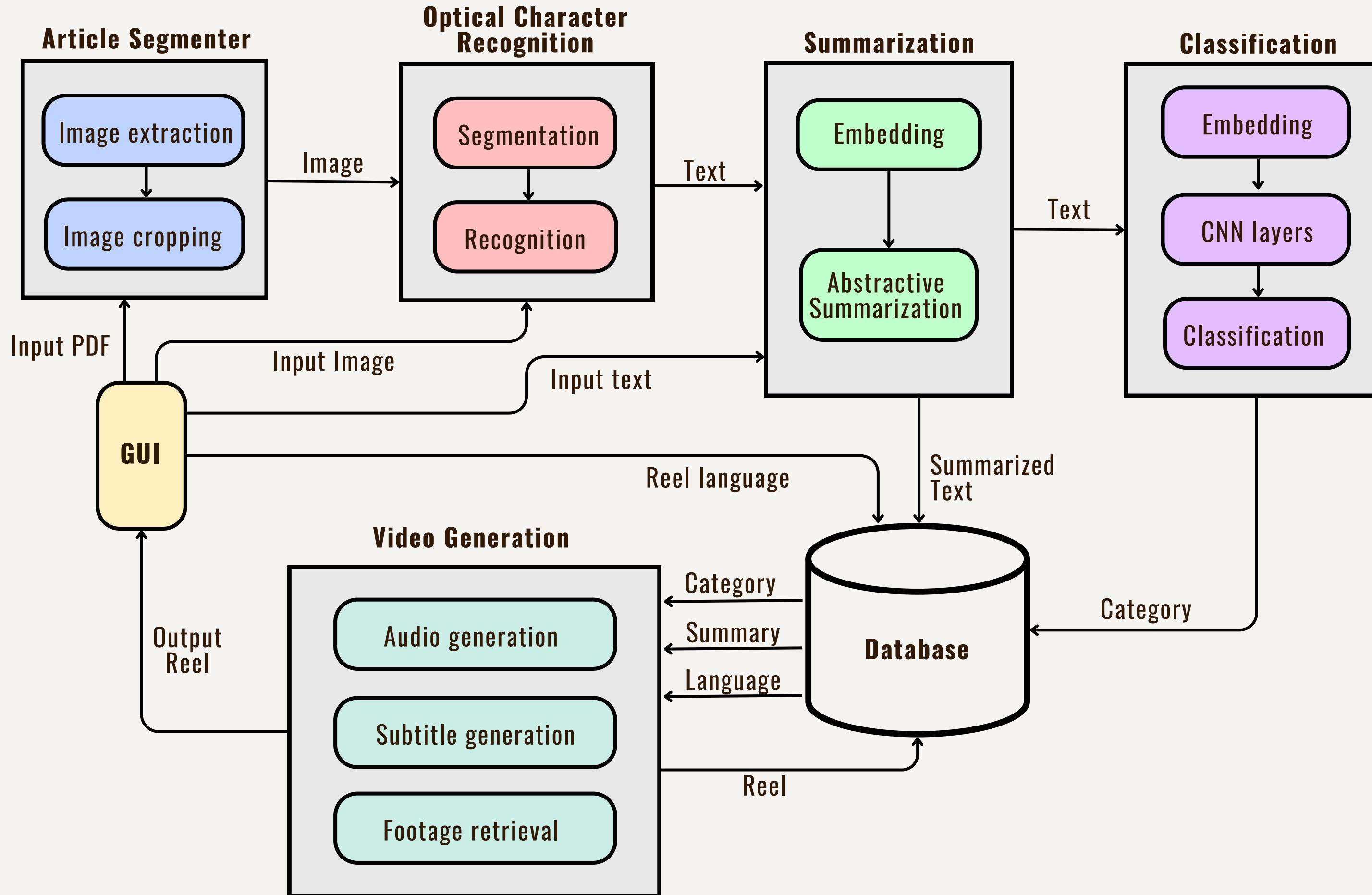
PAPER	ADVANTAGES	DISADVANTAGES
T. Ma et al.[1] T-BERTSum: Topic-Aware Text Summarization Based on BERT (2022)	<ul style="list-style-type: none"><li>• High Accuracy</li><li>• Pre-Trained</li><li>• Understands Context</li></ul>	<ul style="list-style-type: none"><li>• Complex architecture</li><li>• Sensitive to input quality</li></ul>
Y. Liu et al.[2] An Adaptive and Robust Edge Detection Method Based on Edge Proportion Statistic (2020)	<ul style="list-style-type: none"><li>• Clear single pixel edges</li><li>• Edges without breakages</li></ul>	<ul style="list-style-type: none"><li>• Time consuming</li><li>• Needs high computational power</li></ul>

PAPER	ADVANTAGES	DISADVANTAGES
Y. S. Chernyshova et al.[3] Two-Step CNN Framework for Text Line Recognition in Camera-Captured Images (2020)	<ul style="list-style-type: none"><li>• Multilingual segmentation</li><li>• Light-weight and fast</li></ul>	<ul style="list-style-type: none"><li>• Single-Lang recognition</li><li>• Work only for line texts</li></ul>
X. Chen et al.[4] A Long-Text Classification Method of Chinese News Based on BERT and CNN (2022)	<ul style="list-style-type: none"><li>• Classifies multilingual text</li><li>• Identify local features</li></ul>	<ul style="list-style-type: none"><li>• Complete input not considered</li><li>• Full potential of BERT not explored</li></ul>
Q. Chen et al.[5] Scripted Video Generation With a Bottom-Up Generative Adversarial Network (2020)	<ul style="list-style-type: none"><li>• Good performance across datasets</li></ul>	<ul style="list-style-type: none"><li>• Training instability</li><li>• Complex architecture</li></ul>

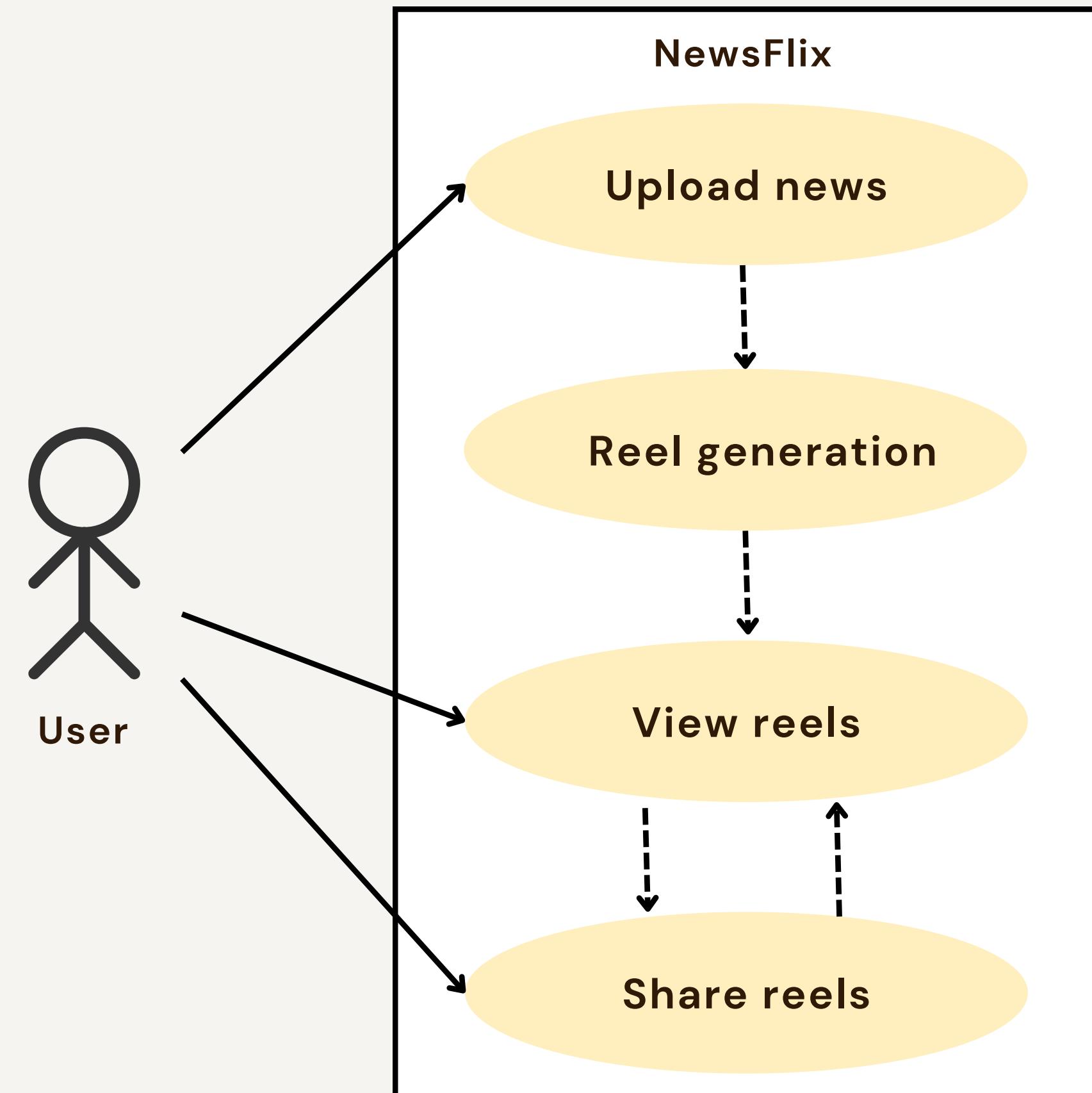
# PROPOSED METHOD

A SaaS that automates the creation of engaging short videos from newspaper articles, making news more accessible and appealing to the modern audience using NLP and AI technologies.

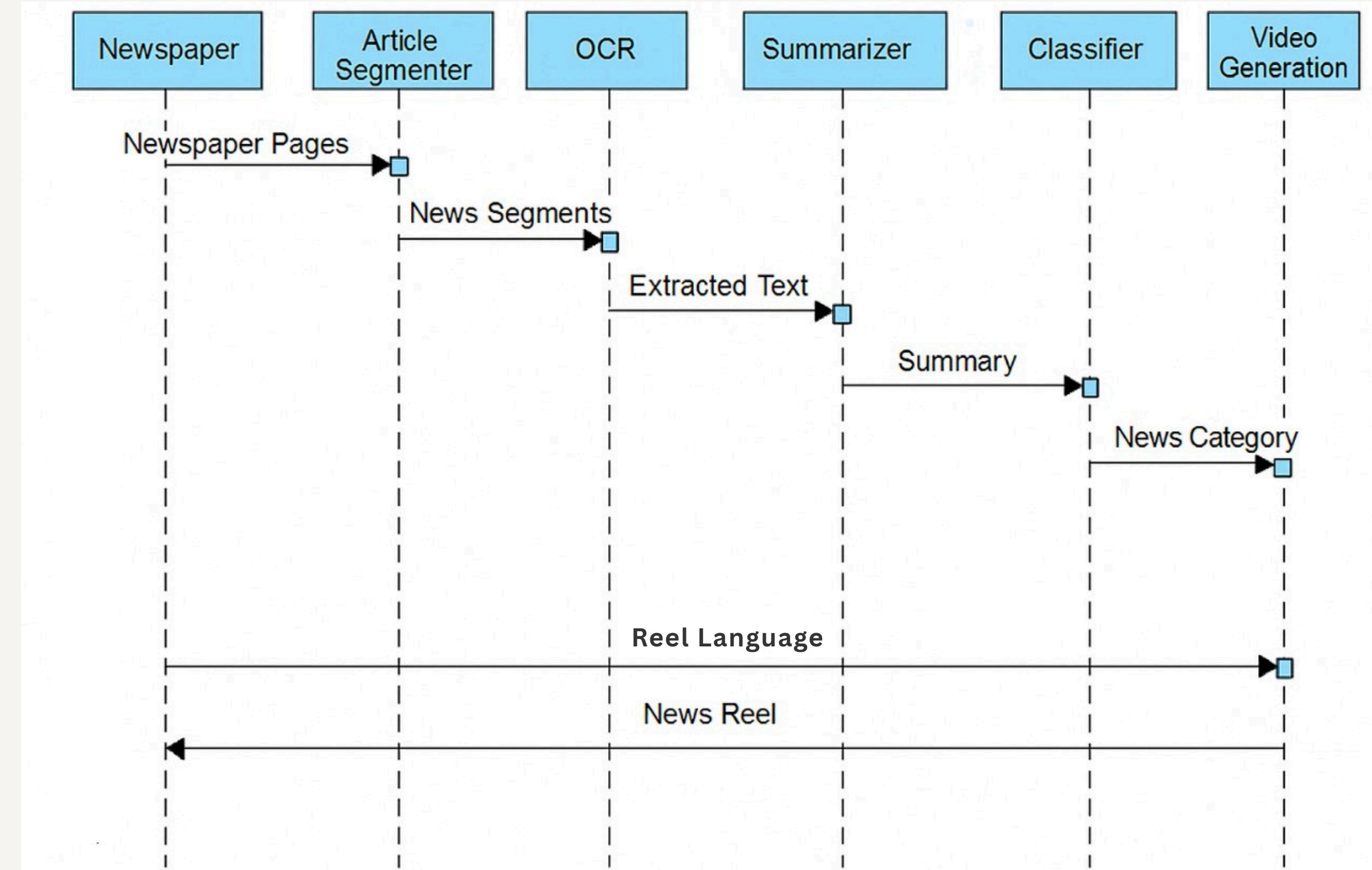
# ARCHITECTURE DIAGRAM



# USE CASE DIAGRAM



# SEQUENCE DIAGRAM



ARTICLE  
SEGMENTATION

*CLASSIFIER*

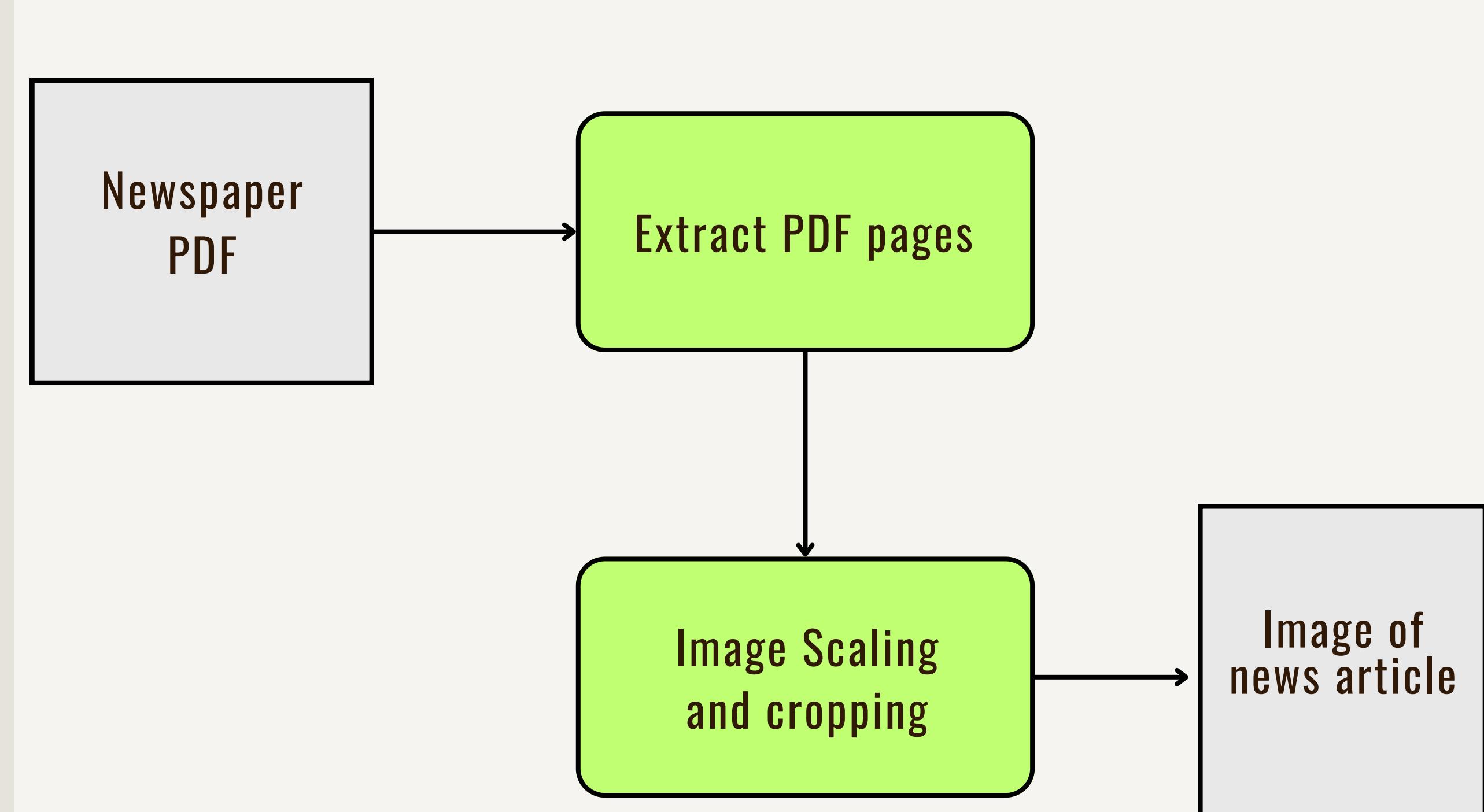
VIDEO GENERATOR

*OCR*

SUMMARIZER

THE  
MODULE WISE  
DIAGRAMS

# ARTICLE SEGMENTER



# ARTICLE SEGMENTER

Input: PDF file uploaded by the user  
Processing the PDF

- Extracts images from PDF pages using PDF.js and displays them in a carousel slider.

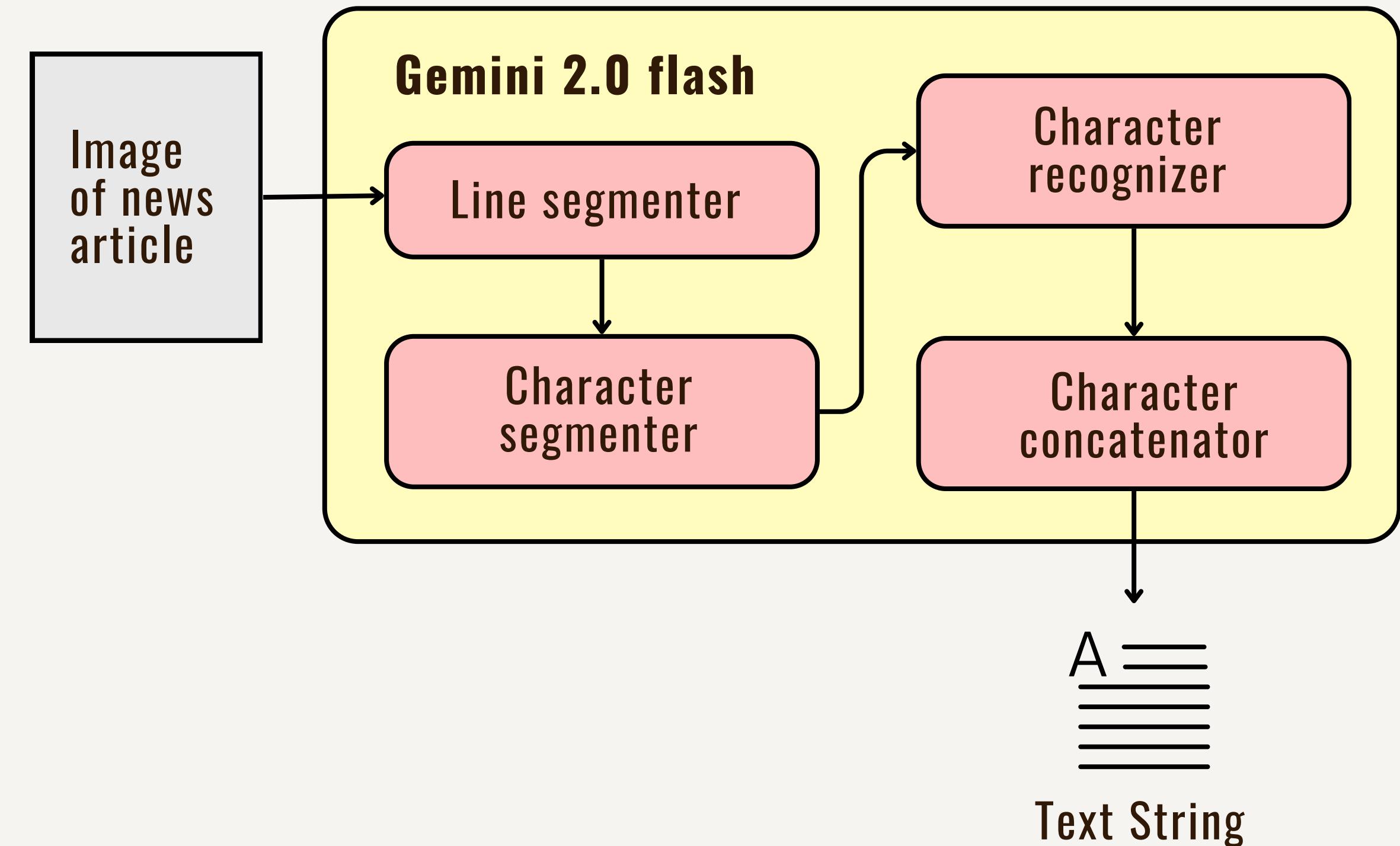
Cropping an Image

- User selects an image, opens a modal, and draws a crop box.
- Cropped image is extracted using a canvas and displayed.

Sending to OCR Module

- Cropped image is converted to a Blob, then a File object which is then sent to the OCR module

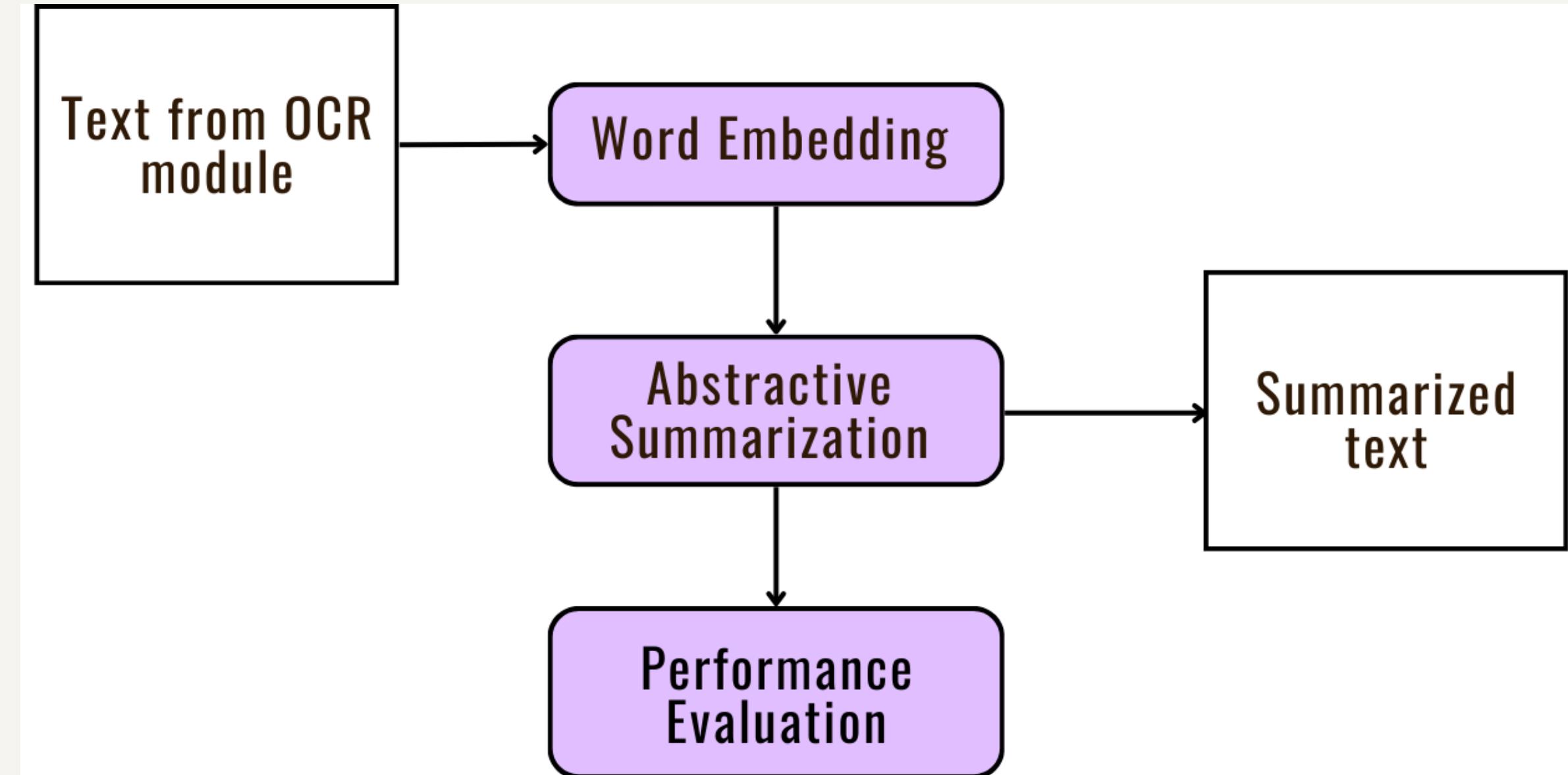
# OCR MODULE



# OCR MODULE

- Input: Output sub-image from the segmenter module.
- Gemini 2.0 flash is used for text extraction.
- Line segmentation using baseline and capline.
- Character segmentation using a language independent CNN.
- Character recognition using a separate CNN.
- All characters are concatenated to get the output text.

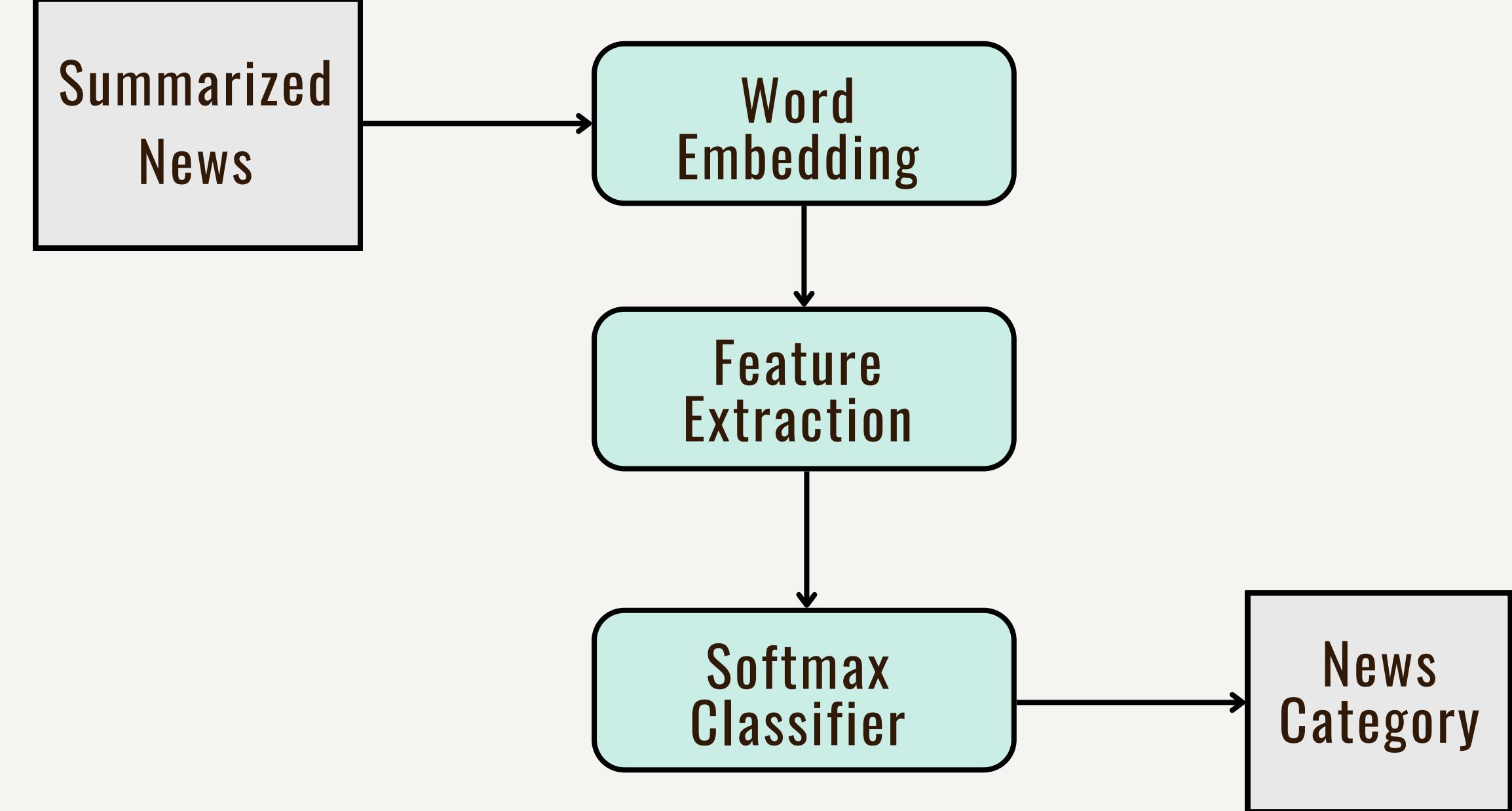
# SUMMARIZER



# SUMMARIZER

- Input: Text recognized by OCR module
- BART tokenizes input text and converts the tokens into word embeddings.
- Uses a bi-directional encoder (like BART) and a left-to-right decoder (like GPT).
- Implements Abstractive Summarization. Instead of copying sentences, BART rephrases and compresses the text into a coherent summary.
- Fine-tuned on SAMSum dataset.
- Output: Summarized text of each article

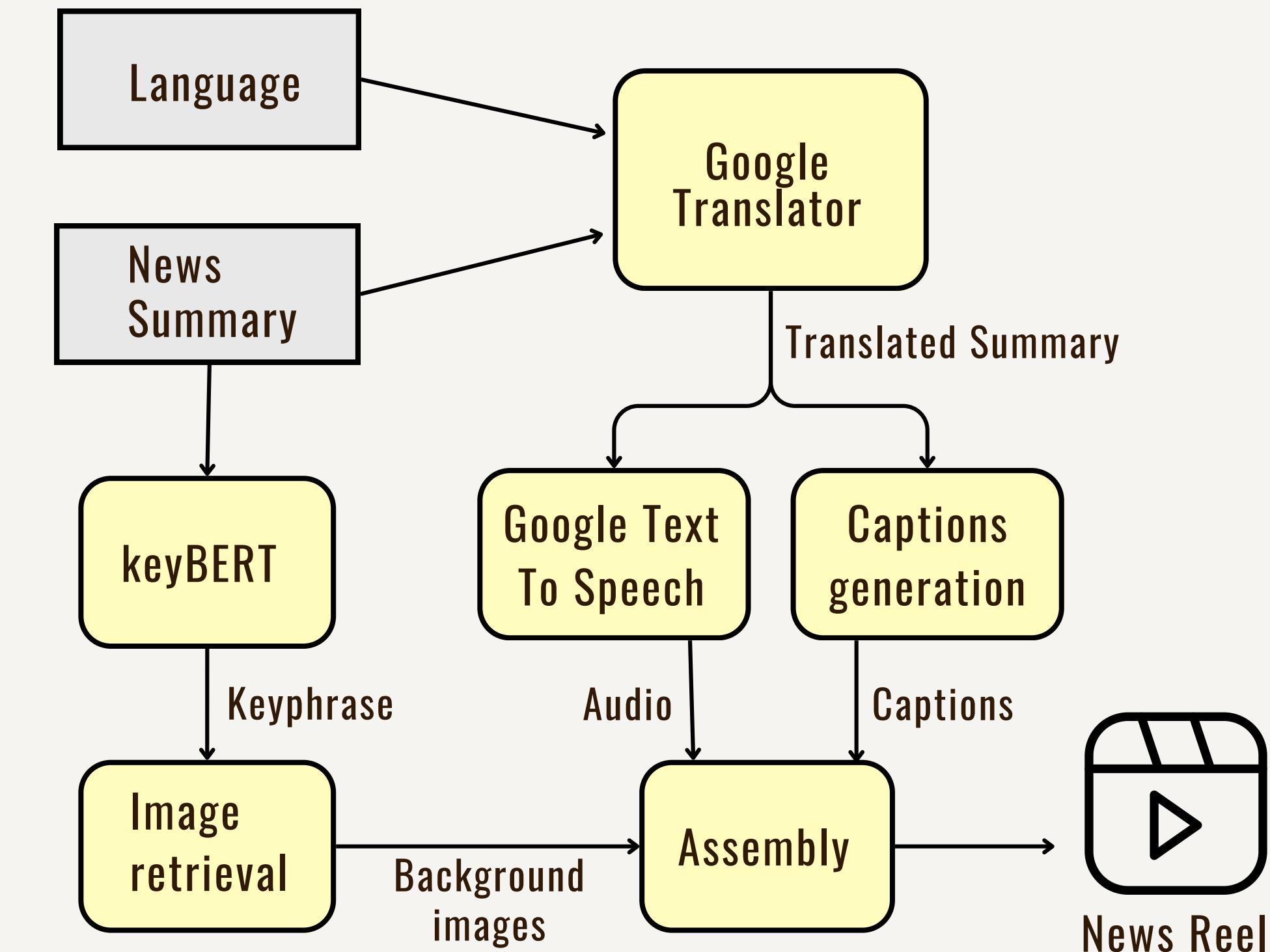
# CLASSIFIER



# CLASSIFIER

- Input: Summarized text of each article
- Using RoBERTa model, word embedding is done. Feature extraction is also done and its information is combined with the word embedding.
- The feature vector obtained is now classified using a classifier( Softmax or Sigmoid ).
- Output: Articles are classified into their respective classes

# VIDEO GENERATOR



# VIDEO GENERATION

---

**Algorithm 1** News Reel Generation

---

**Input:** *Summary* and *Category* of news article, Required output *language*

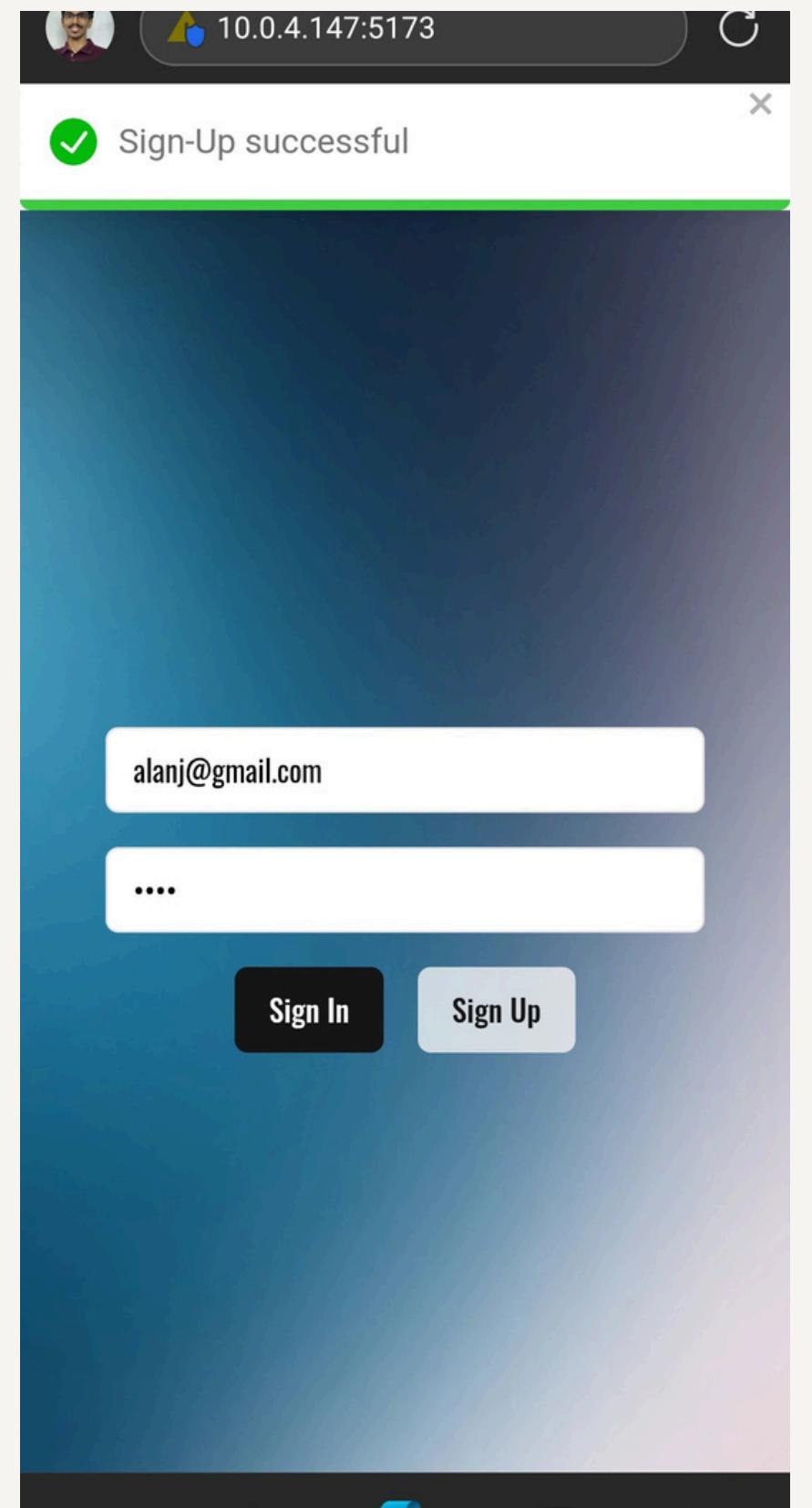
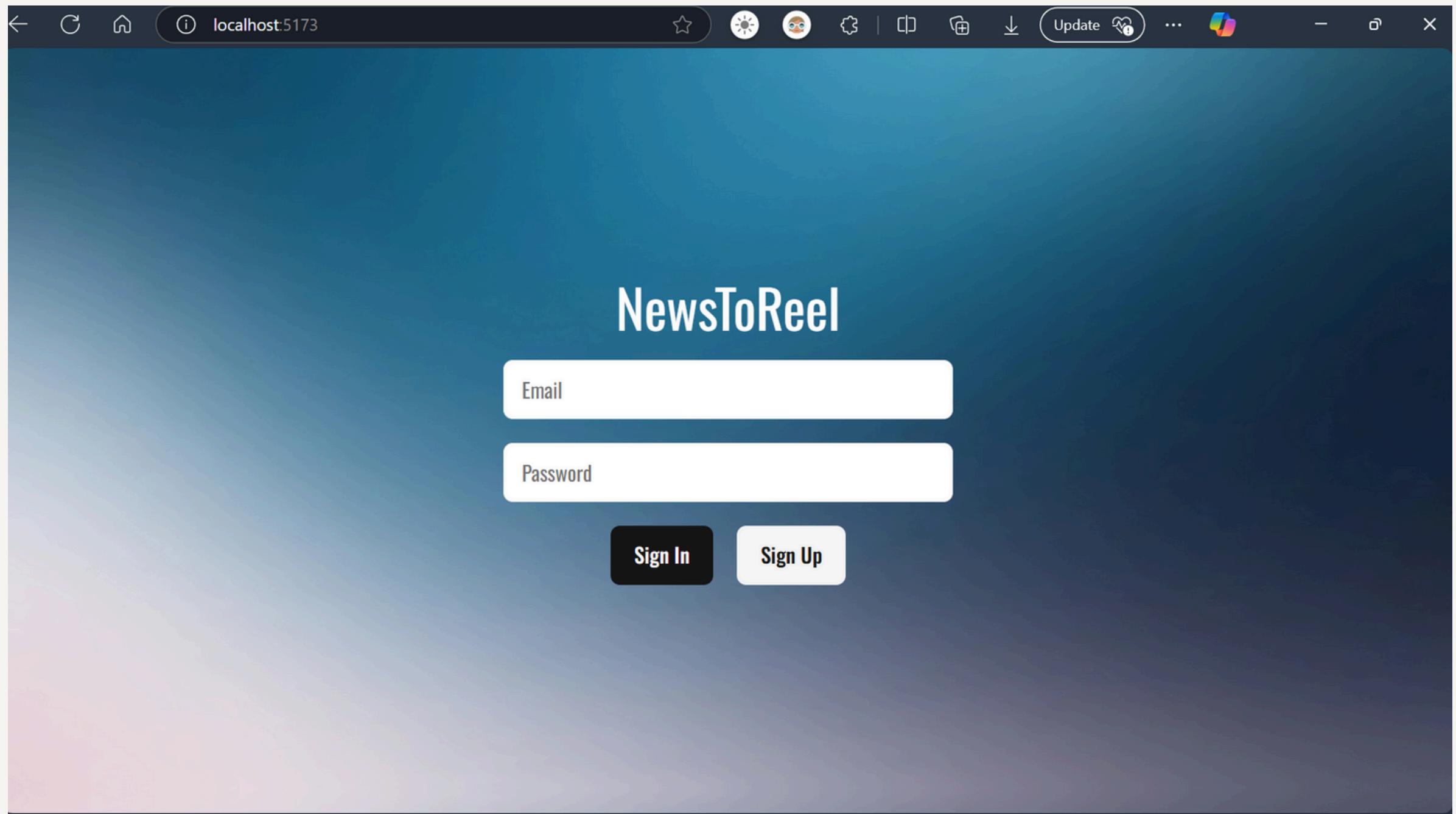
- 1: Extract a *keyphrase* from the *Summary*.
- 2: **if** *language* is not english **then**
- 3:     Translate the *Summary* to *language*.
- 4: **end if**
- 5: Convert *Summary* text to audio using gTTS, save it as reel.mp3.
- 6: Speed up reel.mp3 by 10% to reduce duration.
- 7: Declare *duration* as duration of reel.mp3.
- 8: Initialize *video* with 1080x1920p resolution.
- 9: Calculate no.of images,  $n \leftarrow \lceil \text{duration}/7 \rceil$
- 10: **for**  $i = 1$  to  $n$  **do**
- 11:     Fetch  $i$ th image under *keyphrase* using Pexels API.
- 12:     Download the image as  $i.jpg$ .
- 13:     Append  $i.jpg$  as background of *video* for 7 seconds.
- 14: **end for**
- 15: Transcribe reel.mp3 using Whisper to get synchronized captions.
- 16: Add captions to *video* using the font suitable for *language*.
- 17: Set reel.mp3 as the audio for *video*.
- 18: Write *video* at 1 fps and save it as reel.mp4.

---

**Output:** News video file reel.mp4

# OUTPUTS

## Sign Up & Log In page



# OUTPUTS

Home  
page



# OUTPUTS

# Pages for user selection

# OUTPUTS

## PDF processing page

**pune**  
Pune, Friday, March 28, 2025 [www.freepressjournal.in](http://www.freepressjournal.in)

**UNDOCUMENTED | Lack of 'smashan dakhla' creates ordeal**

## Katkari community faces hassles during cremation

**Indu Bhagat**  
PUNE

Datta Baban Jadhav (65) belonging to the Katkari community, died on Mar 25 around 6pm due to an alcohol overdose. However, his body was only cremated the following day at 2pm. For over 18 hours, his body lay on the streets of Khadakwalsla in Pune. The relatives were unable to perform the last rites due to a lack of necessary documents. They claimed that when they approached officials for a 'smashan dakhla' (cremation certificate), they were asked for basic documents like an Aadhar card, which they could not provide as the community does not possess them.

Speaking to The Free Press Journal (FPJ), Gangu Datta Jadav, wife of the deceased, said, "My husband was an alcoholic. On Monday evening, after dinner, he struggled to perform the last rites as no one was ready to give us the materials required for cremation. The officials asked us for documents, but we are poor and don't know

"The authorities are performing their duties—cremation cannot be performed without a 'smashan dakhla', so we can't really blame anyone. These people do not have

**Ankit Shukla**  
PUNE

The Pune Mahanagar Parivahan Mahamandal Limited (PMPML) witnessed a total of 53,243 bus breakdowns over the last three years. According to data obtained by The Free Press Journal (FPJ), 14,015 PMPML buses broke down in 2022, including 4,067 self-owned and 9,948 leased ones. The number increased in 2023, with 17,784 buses breaking down—7,134 PMPML-owned and 10,649 leased. In 2024, the situation worsened, with a total of 21,444 breakdowns, including 6,349 self-owned and 15,085 leased buses.

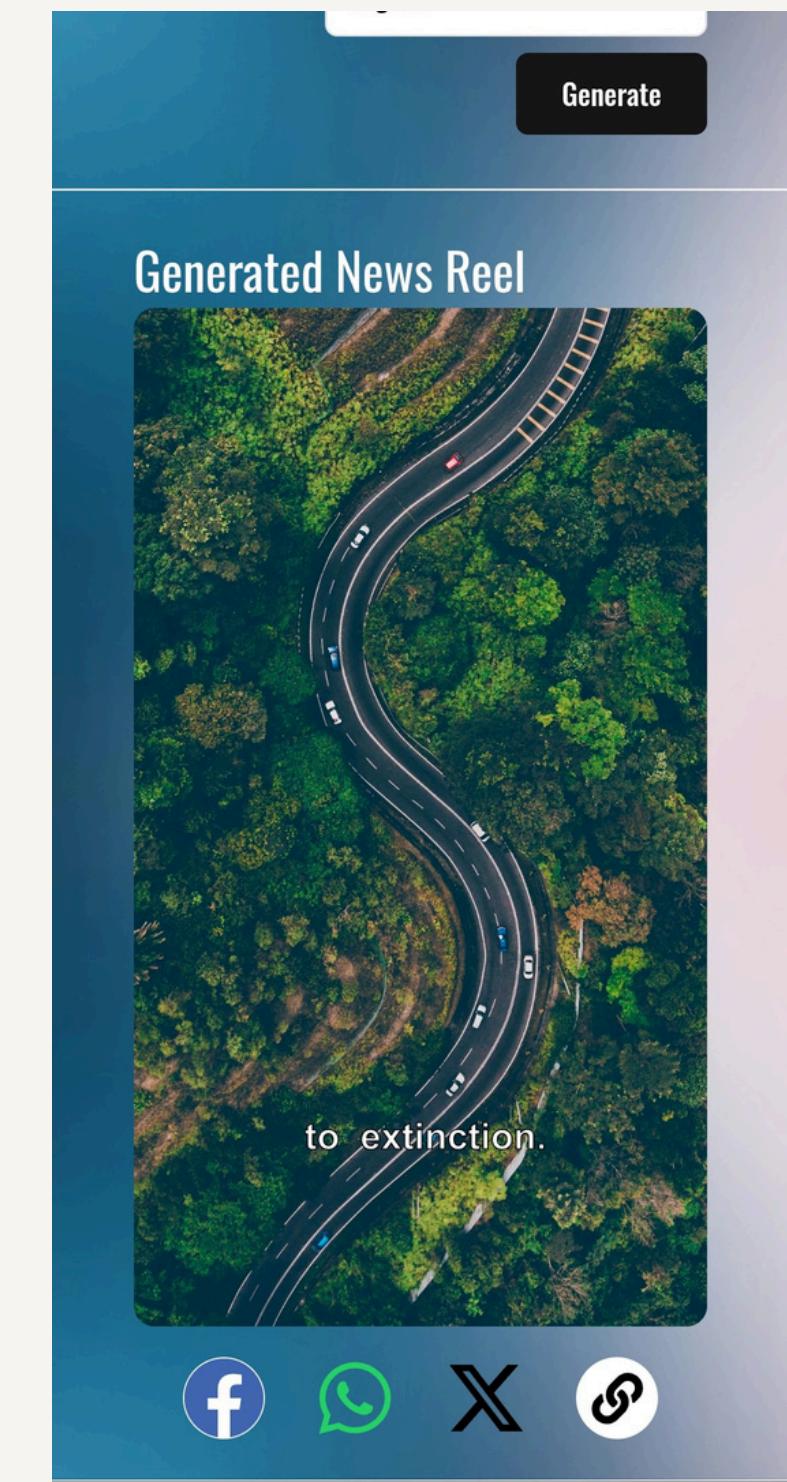
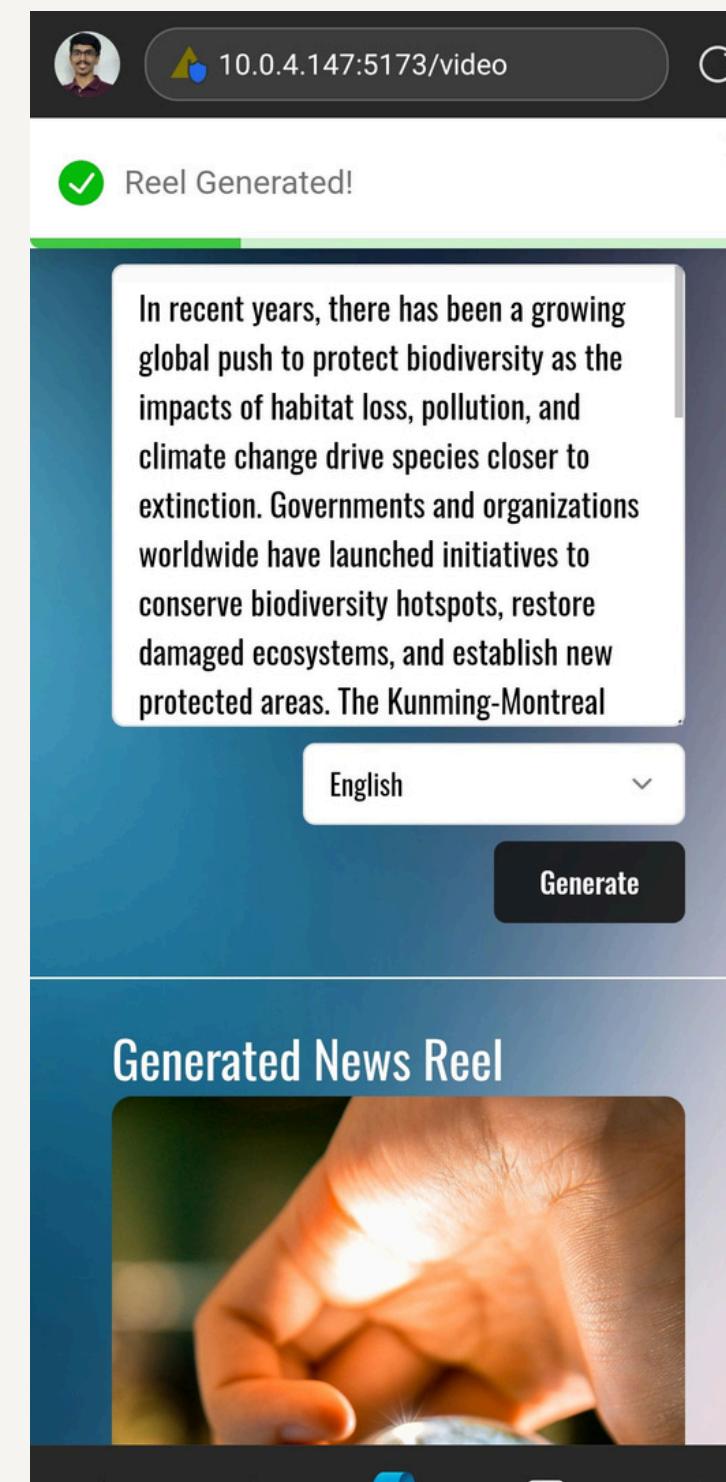
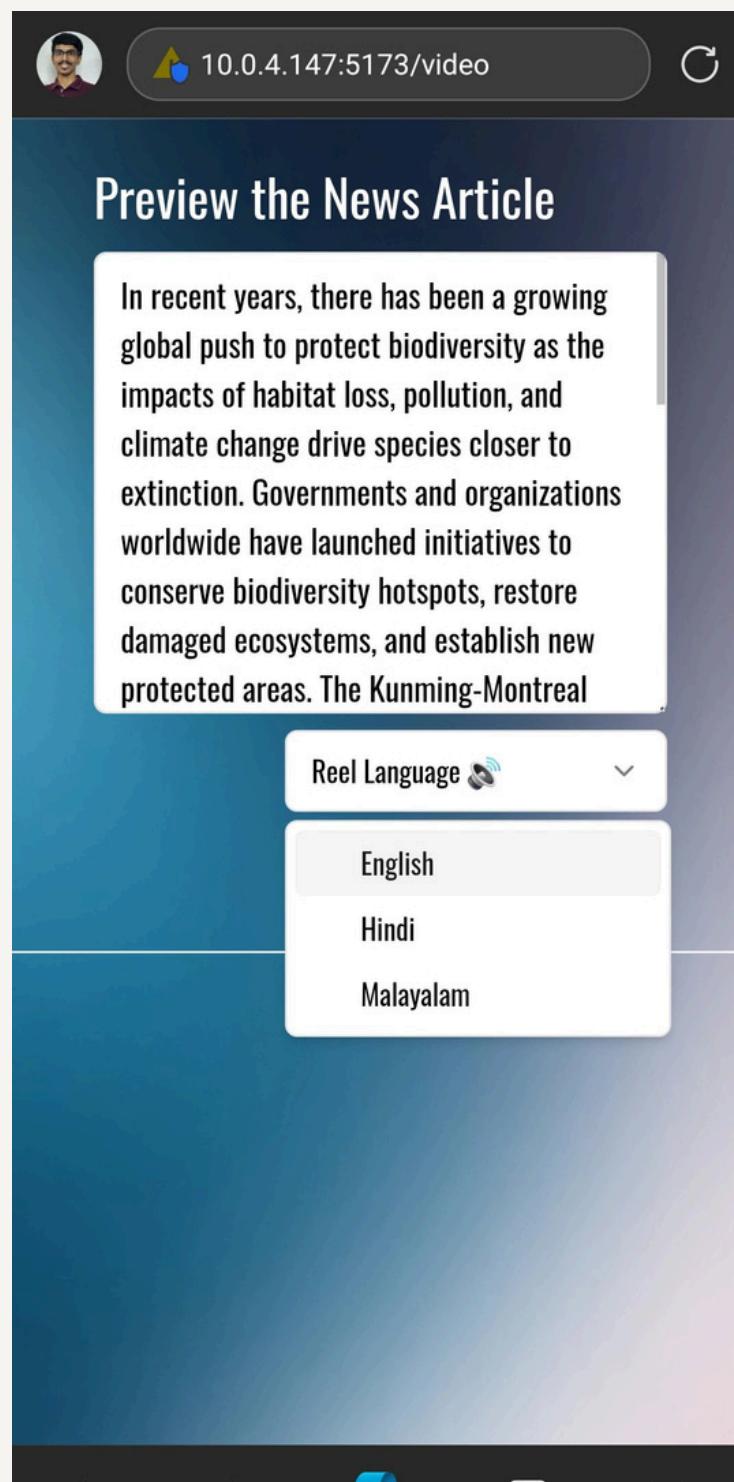
Meanwhile, to address this problem, PMPML has decided to acquire 600 new CNG buses. Among these, 400 buses will be inducted into the PMPML fleet in April, with the remaining 200 self-owned buses to be inducted by the end of June. Recently, over 250 buses have been removed from the fleet due to the expiry of their lifespan.

**PMPML records 53k breakdowns in 3 years**



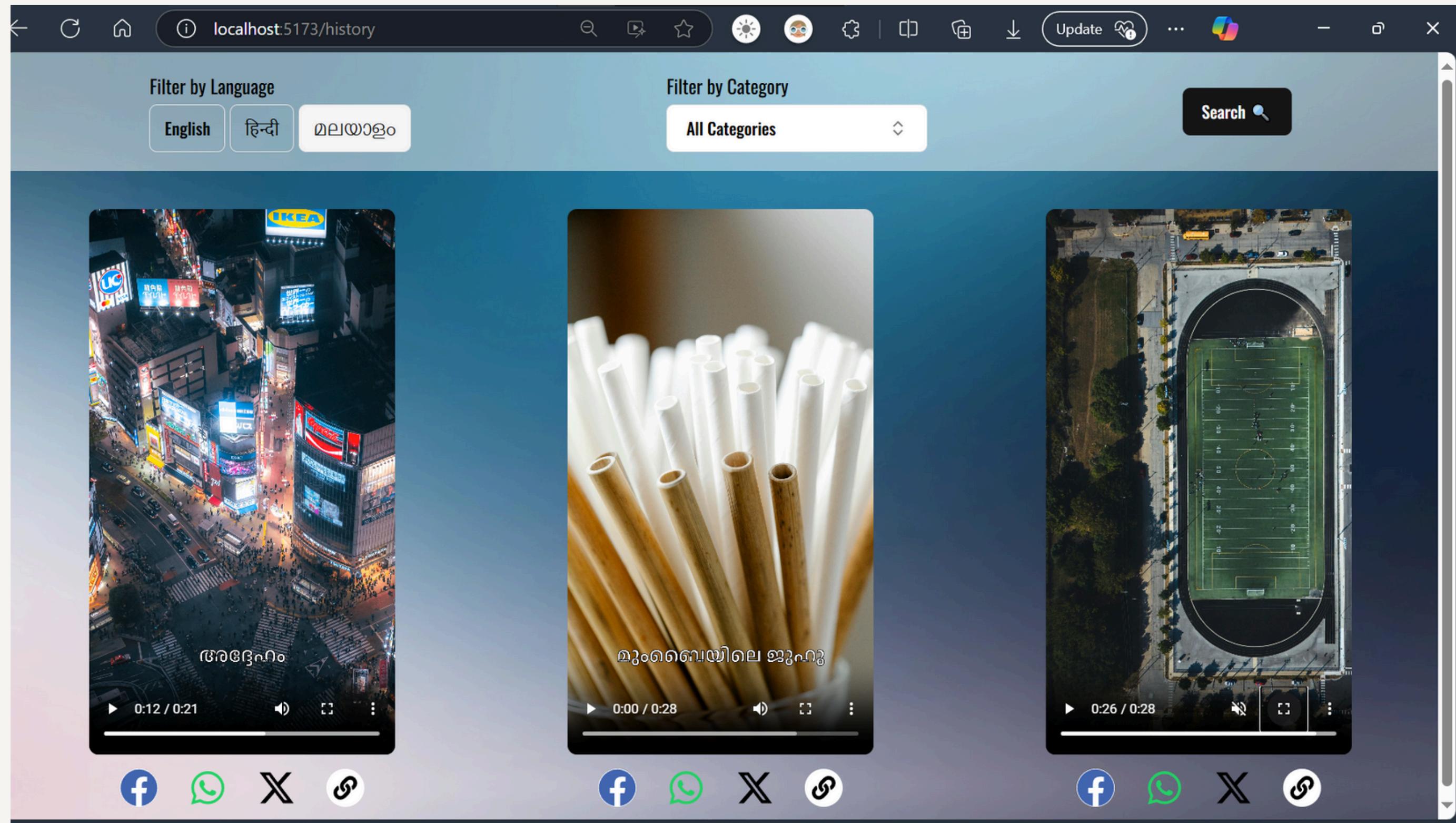

# OUTPUTS

## Reel generation page



# Reel history page

# OUTPUTS



# EVALUATION RESULTS

## OCR

<b>Model / System</b>	<b>PCR %</b>	<b>CER %</b>	<b>WER %</b>
Tesseract 3.04	90.08	9.92	17.36
Tesseract 4.00	91.13	8.87	15.52
FineReader Engine 11	88.40	11.60	20.30
Gemini-2.0-Flash	96.21	3.79	6.63

TABLE I: OCR Performance Comparison

# EVALUATION RESULTS

## Summarizer

<b>ROUGE</b>	<b>Precision</b>	<b>Recall</b>	<b>F1-score</b>
ROUGE-1	0.9651	0.1593	0.2735
ROUGE-2	0.8353	0.1365	0.2347
ROUGE-L	0.9419	0.1555	0.2669

# EVALUATION RESULTS

## Classifier

	Micro-F1	Macro-F1	Accuracy	No. of Instances
All (combined)	0.734	0.746	0.894	1129

	precision	recall	f1-score	support
arts, culture, entertainment and media	0.602151	0.875	0.713376	64
conflict, war and peace	0.611111	0.916667	0.733333	36
crime, law and justice	0.861538	0.811594	0.835821	69
disaster, accident and emergency incident	0.691176	0.886792	0.77686	53
economy, business and finance	0.779221	0.508475	0.615385	118
education	0.847458	0.735294	0.787402	68
environment	0.589041	0.754386	0.661538	57
health	0.79661	0.79661	0.79661	59
human interest	0.552239	0.672727	0.606557	55
labour	0.855072	0.830986	0.842857	71
lifestyle and leisure	0.773585	0.476744	0.589928	86
politics	0.568182	0.735294	0.641026	68
religion	0.842105	0.941176	0.888889	51
science and technology	0.637681	0.8	0.709677	55
society	0.918033	0.5	0.647399	112
sport	0.824324	0.968254	0.890511	63
weather	0.953488	0.931818	0.942529	44

# WORK DONE

Article Segmentation - Basil Eldho Joseph

News Summarizer - Daniel Robin

News Classifier - Abhinav Sobi

Video Generation - Alan Joseph

User Interface - Alan Joseph, Basil Eldho Joseph

# ASSUMPTIONS

- **Clear Input Image:** Distorted or low-resolution images may lead to poor results.
- **Genuineness of Provided News:** Authenticity of the input news content is not validated by the system.
- **News Language:** Input news article is assumed to be in English language.

# REQUIREMENTS

## Software Requirements

- Visual Studio Code
- PgAdmin 4
- Postman
- WebBrowser

## Hardware Requirements

- Intel Core i5 or higher
- Nvidia GTX 1080
- 8 GB RAM
- Screen with a resolution of 1280x720
- Network connectivity

# GANTT CHART

# BUDGET

## CLOUD INFRASTRUCTURE

- For video storage, backend hosting, and user data storage.

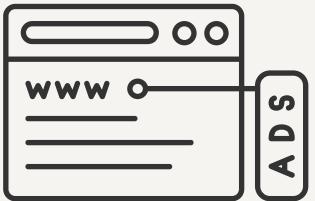
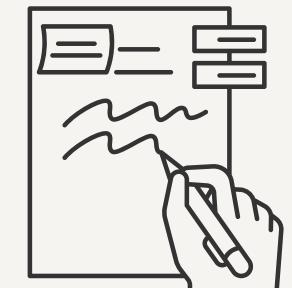
RS.5000

## 3<sup>rd</sup> PARTY APIs

- Gemini API
- Pexels API

RS.4500

# RISKS & CHALLENGES



- Processing complex article layouts
- Extracting paragraph texts
- Handling multiple languages
- Video Generation complexity

# OUTCOMES

- Segmenter outputs sub-image containing individual news article.
- OCR module extracts news article text from the sub-image.
- Summarizer produces summarized news fit for reel content.
- Classifier assigns news category like sports, politics, etc. to the article text.
- Video assembler uses the outputs so far to generate the final reel.
- Final outcome: Users can view and share the generated short & engaging news reels via a mobile friendly web app.

# CONCLUSION

---

A software solution that converts newspaper articles into engaging short videos, making news more accessible to a wider audience.

Utilizes advanced NLP and Image Processing techniques to automatically generate audio summaries, select relevant visuals, and produce high-quality, captivating video content that brings news to life.

# REFERENCES

---

- T. Ma, Q. Pan, H. Rong, Y. Qian, Y. Tian and N. Al-Nabhan, "T-BERTSum: Topic-Aware Text Summarization Based on BERT," in IEEE Transactions on Computational Social Systems, vol. 9, no. 3, pp. 879-890, June 2022
- Y. Liu, Z. Xie and H. Liu, "An Adaptive and Robust Edge Detection Method Based on Edge Proportion Statistics," in IEEE Transactions on Image Processing, vol. 29, pp. 5206-5215, March 2020
- Y. S. Chernyshova, A. V. Sheshkus and V. V. Arlazarov, "*Two-Step CNN Framework for Text Line Recognition in Camera-Captured Images*," in IEEE Access, vol. 8, pp. 32587-32600, February 2020
- J. Memon, M. Sami, R. A. Khan and M. Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)," in IEEE Access, vol. 8, pp. 142642-142668, July 2020
- Q. Chen, Q. Wu, J. Chen, Q. Wu, A. van den Hengel and M. Tan, "Scripted Video Generation With a Bottom-Up Generative Adversarial Network," in IEEE Transactions on Image Processing, vol. 29, pp. 7454-7467, May 2020

# REFERENCES

---

- J. Zheng and L. Zheng, "A Hybrid Bidirectional Recurrent Convolutional Neural Network Attention-Based Model for Text Classification," in IEEE Access, vol. 7, pp. 106673-106685, March 2019
- K. M. Hasib et al., "MCNN-LSTM: Combining CNN and LSTM to Classify Multi-Class Text in Imbalanced News Data," in IEEE Access, vol. 11, pp. 93048-93063, June 2023
- J. Park, E. Lee, Y. Kim, I. Kang, H. I. Koo and N. I. Cho, "Multi-Lingual Optical Character Recognition System Using the Reinforcement Learning of Character Segmenter," in IEEE Access, vol. 8, pp. 174437-174448, Jan 2020
- J. Memon, M. Sami, R. A. Khan and M. Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)," in IEEE Access, vol. 8, pp. 142642-142668, July 2020
- Q. Chen, Q. Wu, J. Chen, Q. Wu, A. van den Hengel and M. Tan, "Scripted Video Generation With a Bottom-Up Generative Adversarial Network," in IEEE Transactions on Image Processing, vol. 29, pp. 7454-7467, May 2020

**THANK YOU**