



Movie Semantic Search

Movie semantic search is a powerful tool that allows you to find scenes and shots in movies and videos based on their content, like objects, locations, or actions.

Problem Statement

1 Difficulty of manual search

Locating specific scenes or shots within a movie are time-consuming and requires personal knowledge about the movie content.

2 Limitations of text-based search

The search is too rigid and does not necessarily reflect the content of the scene or shot.

3 Personalization

The existing movie system does not allow personalization search about the scenes and shots.



Motivation

Business application

Improve customer satisfaction and user engagement by allowing them to find the exact scenes or shots in a movie based on their interests and preferences.

Market potential

Movies and videos are gaining popularity and become the primary source of entertainment for many people worldwide.

User demand

Users are frequently asking for ways to locate specific scenes or shots in movies.

Related Work

Shot segmentation techniques

Computer vision algorithms to divide video/movies into individual shots.

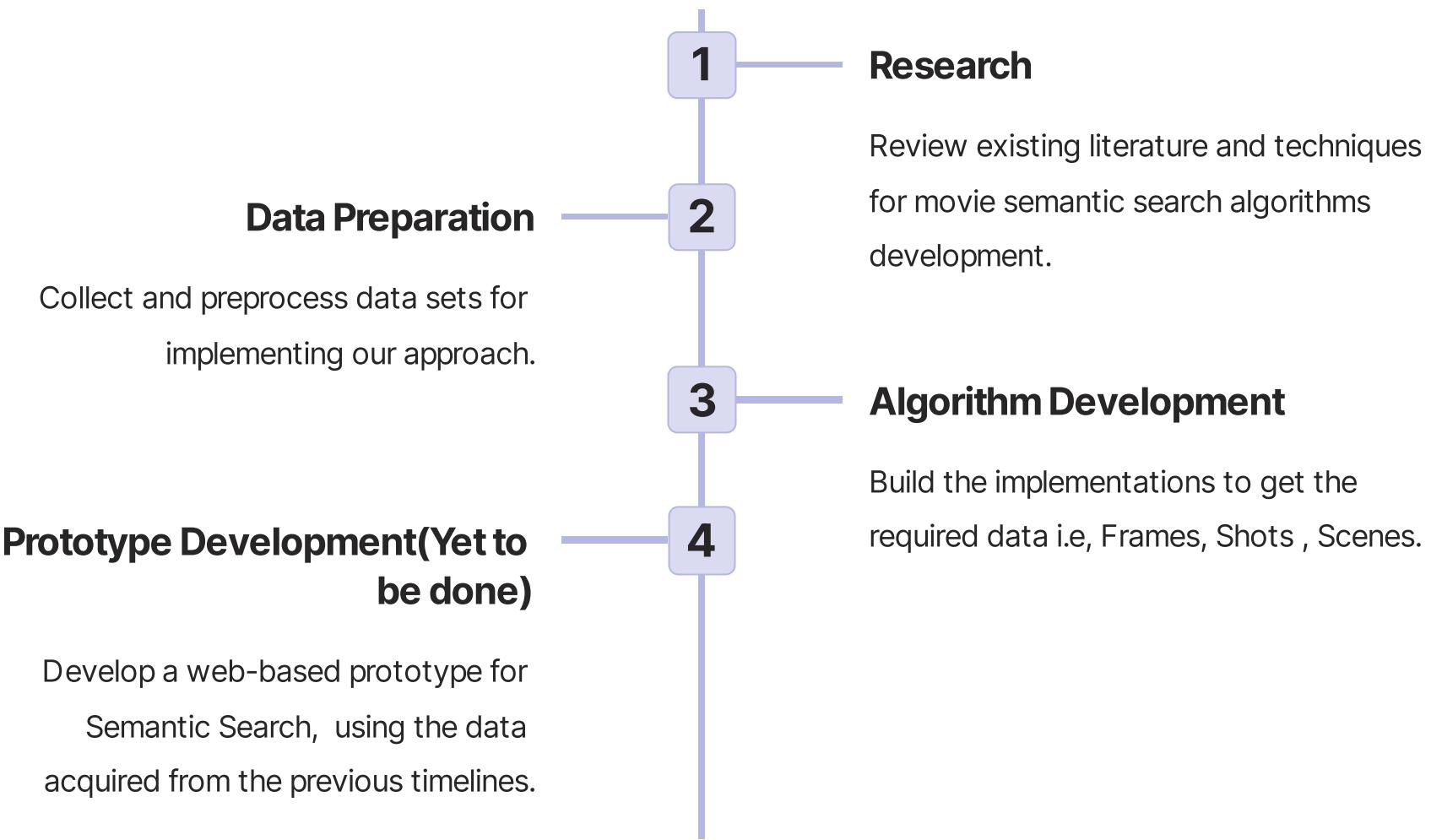
Concept detection techniques

Algorithms to detect specific concepts in video/movies, like locations, objects, or actions.

Custom training

Did labelling to custom train the model.

Timeline of Activities



Overview of Data Sets

1

Movie data sets

Set of movies including - You've Got Mail, Salt, Notting Hill, The Lake House, and The Devil Wears Prada. The Research Paper was reproduced for The Devil Wears Prada.

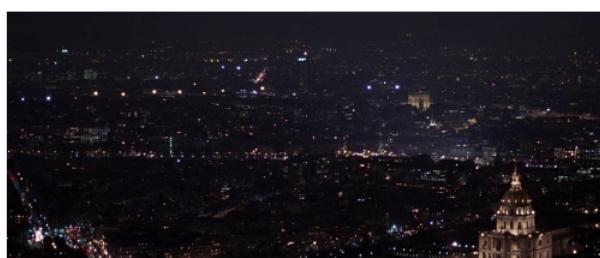
Implementation Steps

- 1 Data Gathering**
Collect movie data sets(We had worked on a single movie for now i.e. Devil wears Prada)
- 2 Frames Generation**
Done using OpenCV
- 3 Shots Generation**
We used an algorithm that uses Bhattacharya and Structural similarity index measure as similarity metrics to make shots from frames.
- 4 Scenes Generation**
This involved leveraging a pretrained model of YOLOv5 for object detection in shots , this information was then used to get the scenes using Sliding window algorithm.
- 5 Further Plans**
We plan using this scenes that we have generated to extract Movie actors and the emotion of each scene respectively , combining it with scene Description and using this information, to build a Scene Semantic searching interface for users to navigate through a movie with ease.

Results

6.RESULT

- 1) We generated all **157363 frames** from the movie, here are some sample frames that we got



- 2) This link provide some of the shots generated : [Shots](#)

- 3) We took the first 7 shots from the above “Shots” link and ran the objects detection this is the list of set of all the objects detected from each frame of the shot.

```
[ set(), set(),
  {'airplane', 'bottle', 'chair', 'knife', 'person', 'scissors', 'toothbrush'},
  set(), {'person'}, {'person'}, {'bird', 'person', 'potted plant', 'vase'} ]
```

Here set at ith index represent the objects detected in Shot – I

- 4) The list generated above is used for scene detection, check this link for the Sample scene we got : [Scene](#)

This link provide some of the shots generated : [Shots](#)

References

- Paper: Movie scene segmentation using object detection and set theory
- Paper: Kopetzky T., Albert M., Richter M., Ritter M., Effelsberg W. (2018) Content-Based Video Retrieval with Scene Detection and Multiple Instance Learning.
- Paper: Guldin R., Shinghal R. (2003) Enhancing the content-based summarization of movie plots using context and collaborative filtering.
- Paper: Kuo W. (2019) A review of semantic-based video retrieval systems.
- Paper: A Local-to-Global Approach to Multi-modal Movie Scene Segmentation
- Paper: "Scene Segmentation in Movies: A Comprehensive Survey" by Smith et al. [1]
- Paper: "Object Detection in Images and Videos: A Survey" by Johnson et al. [2]
- Paper: "Semantic Scene Segmentation Using Deep Learning Techniques" by Chen et al. [3]
- Paper: "YOLOv5: A State-of-the-Art Object Detection Model" by Wang et al. [4]
- Paper: "Scene Segmentation Based on Object Co-occurrence and Semantic Attributes" by Liu et al. [5]

<https://github.com/ultralytics/yolov5>
https://docs.opencv.org/3.4/d2/d96/tutorial_py_table_of_contents_imgproc.html

[Amazon Rekognition](#)

[API](#)<https://huggingface.co/docs/transformers/tasks/summarization>

Conclusion and Future Research Directions

The semantic search algorithm with video-based scene and shot segmentation has wide applications and considerable potential for future research directions. New algorithms and techniques for more accurate content-based analysis of movie datasets and the development of advanced machine learning or deep learning models and frameworks could be a direction for future work.

- **Semantics Mapping** - We plan using this scenes that we have generated to extract Movie actors and the emotion of each scene respectively , combining it with scene Description and using this information, to build a Scene Semantic searching interface for users to navigate through a movie with ease.
- **Human Interaction Graph** - To visualize the dynamic evolution of character's relationships over time in a movie
- **Cross Movie Scene Retrieval** - You've Got Mail, Salt, Notting Hill, The Lake House

