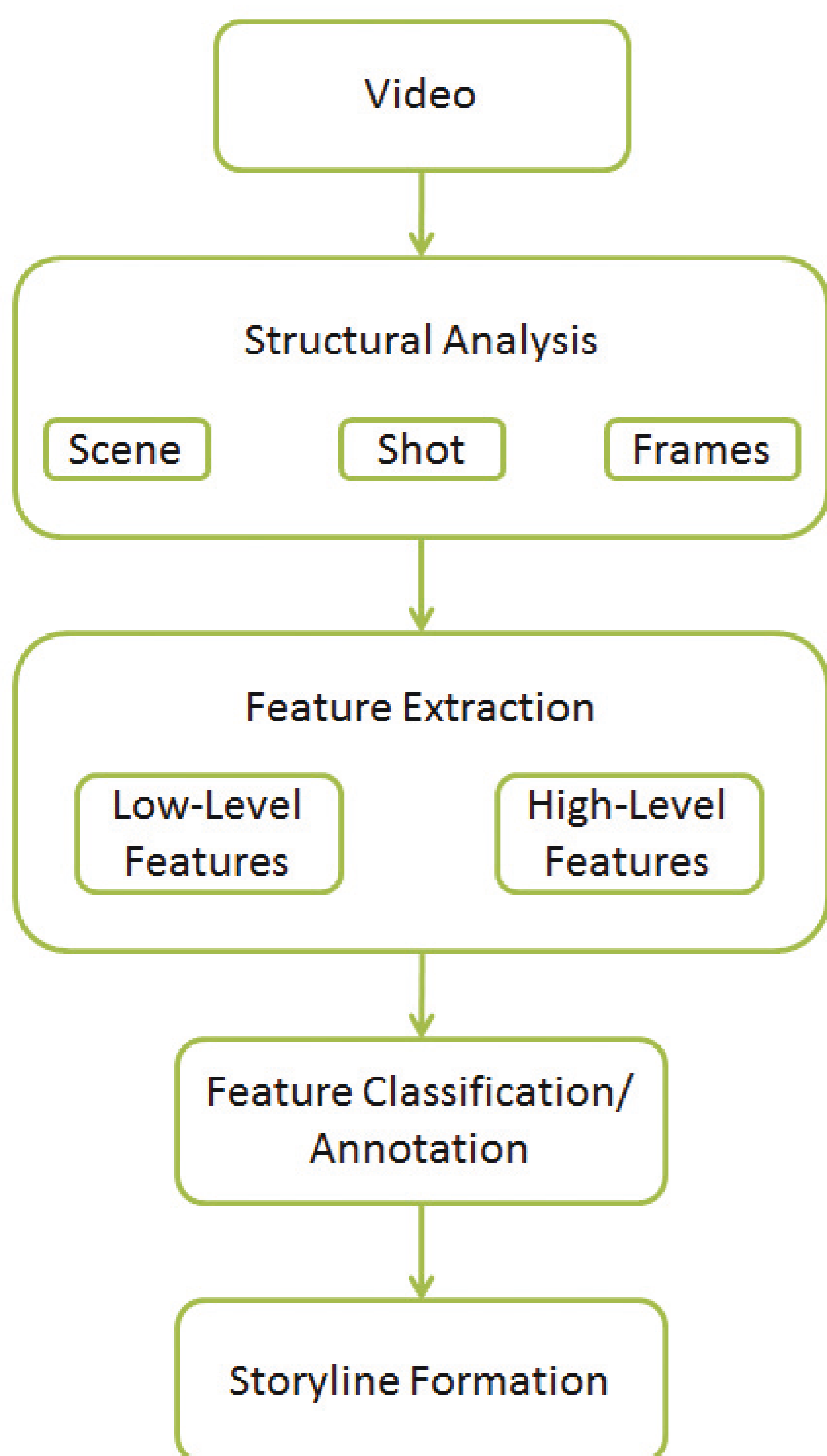


# Automatic Video Annotation

Currently all video search engines are text-based, i.e. they search for the text labels associated with any video to retrieve the desired ones. However, this can lead to incorrect or inaccurate results, as labelling or annotating a video is mainly done manually. Consequently, many false positive results are generated during video searches by mislabelled videos. To solve this problem we need to improve the process of video annotation. This can be achieved by automatic annotation of videos based on their actual content, rather than text labels or tags. To accomplish this we need to enable computers to extract video “storylines”, composed by the events or actions taking place in each video. This has the potential to save time and provide better results for online video searches, as well as improve event detection in real-world surveillance footage. The project aims to facilitate Probabilistic Semantic Search and Query Answering by annotating videos in the way described, through machine learning techniques such as Probabilistic Graphical Models (PGMs) and Discriminative Deformable Part-Based (DDPB) Models able to automatically detect and recognize objects, persons, actions and events within a video.

## Literature Review



### Structural Analysis

First step in understanding semantics of a video is structural analysis by breaking it down into smaller objects like scenes, shots, key frames. The main objective is to divide the video into small meaningful objects for understanding the details of a particular concept .

### Feature Extraction

There are two types of features that can be extracted from the video objects:

1. Low-Level Features including color, shape, texture, etc.
2. High-Level Features like SIFT, PCA-SIFT, HoG, SURF, Bag-of-Features, Bag-of-Video-Words and more

### Feature Classification/Annotation

Features are automatically classified through understanding and differentiating between different feature concepts. Semi-supervised methods for Graphic Modelling are best suited for automatic detection and classification of the semantic concepts. Some of these methods include Latent Semantic Analysis, Latent Dirichlet Analysis, Probabilistic Canonical Correlation Analysis.

### Storyline Formation

Lastly, the storyline can be formed by automatically assigning the classified features to appropriate tags and then creating sentences through human understandable language semantics. The techniques of Natural Language Processing are essential parts of this method. User-centric search query model can help in updating the information through the feedback model.

## Applications

Automatic Annotation and storyline formation can have countless applications in multiple industries including web-based video search, healthcare, surveillance and security.