**Software Architectures**

**Project Report: Video Analysis Platform**

**Halil İbrahim Akça**

**Elif Yıldız**

**Ramazan Ediz**

**Talha Çetinbaş**

**TEAM 8**

**1. Project Summary**

This project is a comprehensive web application that analyzes video files to extract audio emotions, text sentiment, and other key features. It has been developed using a Flask-based web interface and an advanced media analysis pipeline. The project aims to derive valuable insights from video content using artificial intelligence models.

**Instructions on how to run the application are provided in the report.**  
**You can access the source code here:** [github.com/Halilakca17/Software-Architectures](https://github.com/Halilakca17/Software-Architectures)

**2. Technical Infrastructure**

**2.1 Technologies Used**

**• Backend Framework:**

* Flask 2.3.3
* Python 3.10+

**• Frontend:**

* HTML5
* CSS3
* JavaScript
* Bootstrap 5.3.0-alpha1

**• Core Libraries:**

* NumPy 1.26.4+: Numerical operations and data manipulation
* PyTorch 2.0.0+: Deep learning models and GPU support
* Librosa 0.10.0+: Audio processing and analysis
* Transformers 4.30.0+: NLP models
* MoviePy 1.0.3+: Video processing
* Matplotlib 3.7.0+: Data visualization
* Pandas 2.0.0+: Data analysis
* TextBlob 0.17.1+: Text processing
* OpenAI Whisper 20231117: Speech-to-text conversion

**2.2 Project Structure**

metin, ekran görüntüsü, yazı tipi, tasarım içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

**3. Key Features**

**3.1 Video Analysis Pipeline**

**3.1.1 Audio Extraction**

* Extracting audio from video files
* Converting to WAV format
* Audio normalization

**3.1.2 Transcription**

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

* Utilizes Whisper model
* Error correction mechanisms

**3.1.3 Text Summarization**

metin, ekran görüntüsü, yazı tipi içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

* Implementation of the T5 model
* Chunk-based processing
* Summary length optimization
* Fallback summarization strategies

**3.1.4 Sentiment Analysis**

metin, ekran görüntüsü, diyagram, renklilik içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

* Uses the BERTweet model
* Sentiment categories:
  + Joy
  + Optimism
  + Anger
  + Sadness
  + Neutral
* Confidence scores
* Batch processing optimization

**3.1.5 Audio Emotion Analysis**

metin, ekran görüntüsü, yazı tipi, tasarım içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

* Implementation of the HuBERT model
* Emotion transition analysis
* Segmentation
* Visualization

**3.2 Web Interface Features**

metin, ekran görüntüsü, web sayfası, yazılım içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

**3.2.1 File Upload**

* Drag-and-drop support
* Multi-format support (mp4, avi, mov, mkv, webm)
* File size control (max 500MB)
* Progress indicator

**3.2.2 Analysis Tracking**

* Real-time status updates
* Step-by-step process visualization
* Error notifications
* Auto-refresh

**3.2.3 Result Visualization**

* Interactive charts
* Detailed analysis reports

**4. Security and Performance**

**4.1 Security Measures**

* File extension validation
* Secure file name handling (secure\_filename)
* Maximum file size enforcement
* Temporary file cleanup

**4.2 Performance Optimizations**

* Asynchronous processing (threading)
* Batch processing support
* Model caching
* GPU support

**5. User Experience**

**5.1 Interface Features**

* Responsive design
* Modern UI/UX
* Progress bar
* Status notifications
* Error handling

**5.2 Analysis Results**

* Text summary
* Full transcript
* Sentiment analysis charts
* Audio emotion analysis visualizations

**6. Installation and Requirements**

**6.1 System Requirements**

* Python 3.10+
* FFmpeg
* Optional: CUDA-enabled GPU

**6.2 Installation Steps**

1. Creating a virtual environment

metin, yazı tipi, ekran görüntüsü içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

1. Installing dependencies



1. Launching an application



**7. Future Enhancements**

**7.1 Planned Features**

* Multi-language support
* Advanced video analysis
* User accounts
* API integration
* Batch processing support
* Mobile application
* Real-time analysis
* Custom model training

**7.2 Areas for Improvement**

* Model optimization
* Enhanced error handling
* UI/UX improvements
* Performance tuning
* Security updates

**8. Technical Details**

**8.1 Models Used**

* Whisper: Speech-to-text conversion
* BERTweet: Sentiment analysis
* HuBERT: Audio emotion recognition
* T5: Text summarization

**8.2 Data Processing Pipeline**

1. Video upload and validation
2. Audio extraction and normalization
3. Transcription and summarization
4. Sentiment analysis and visualization
5. Saving and presenting the results

**9. Conclusion**

This project offers a comprehensive video analysis solution by successfully integrating modern web technologies and advanced AI models. With its user-friendly interface, powerful analysis capabilities, and scalable architecture, it enables valuable insights to be derived from video content. The project has significant potential for continuous development and improvement.