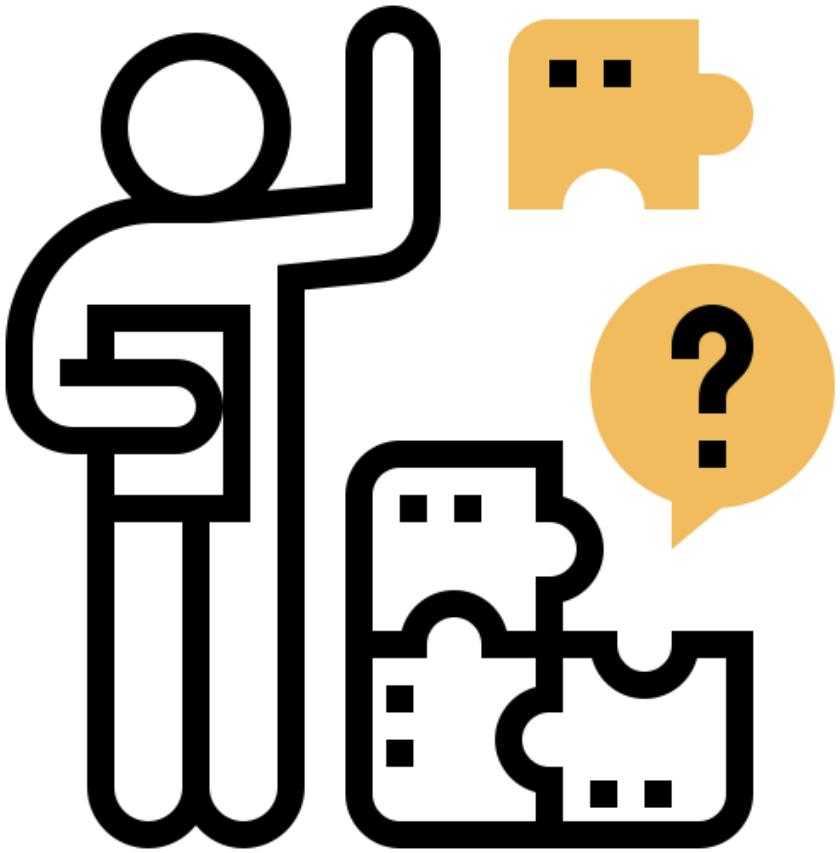
AUDIO - VISUAL EMOTIONAL RECOGNITION

Merve DADAŞ – 201511016 (CENG)
Furkan KARADAŞ – 201511033 (CENG)

Aydın ŞİŞMAN – 201514213 (ECE)
Halil Uğur BAYEZİT – 201514013 (ECE)

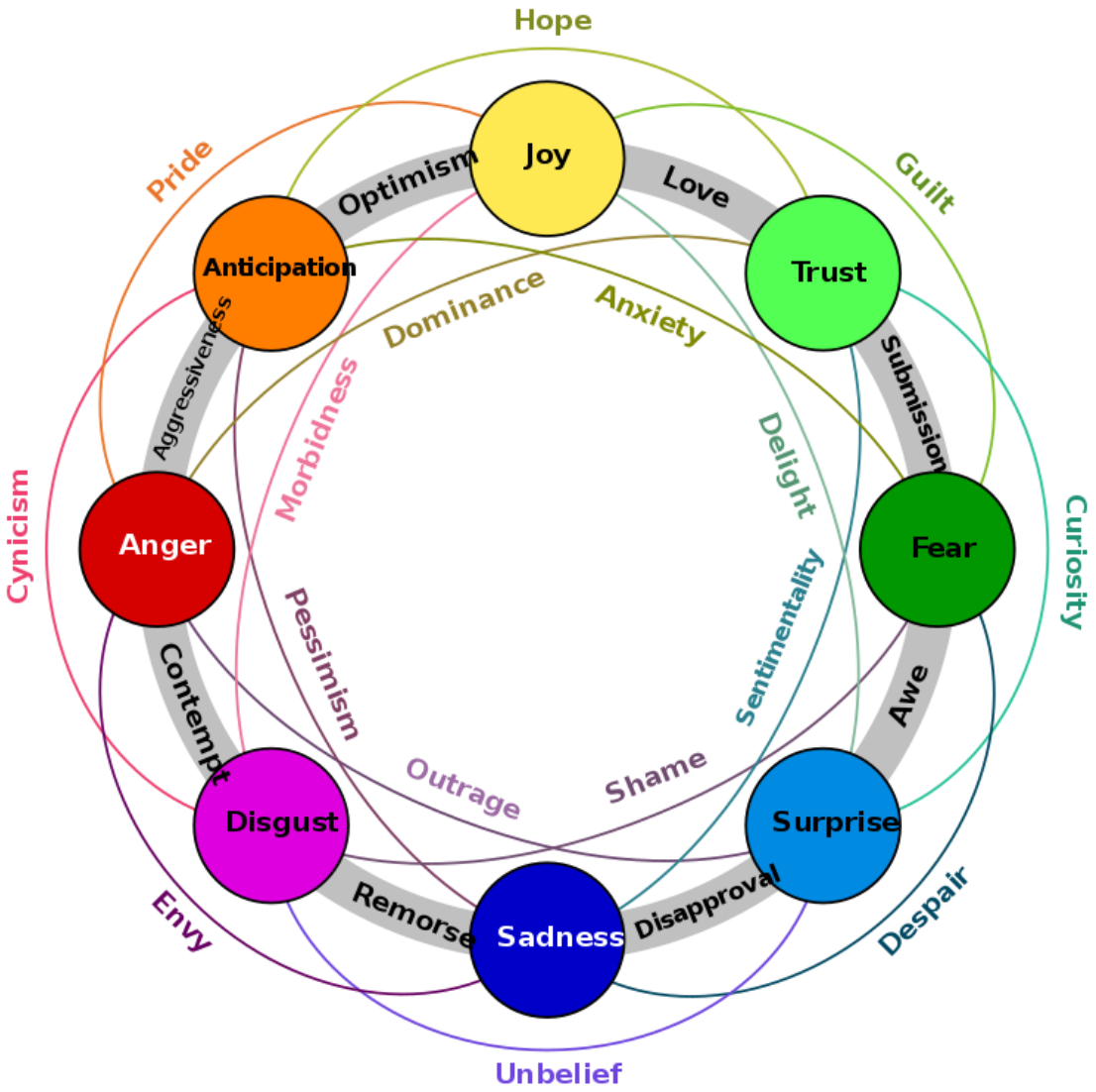
Assoc. Prof. Dr. Hadi Hakan MARAŞ
&
Asist. Prof. Selma ÖZAYDIN

Contents



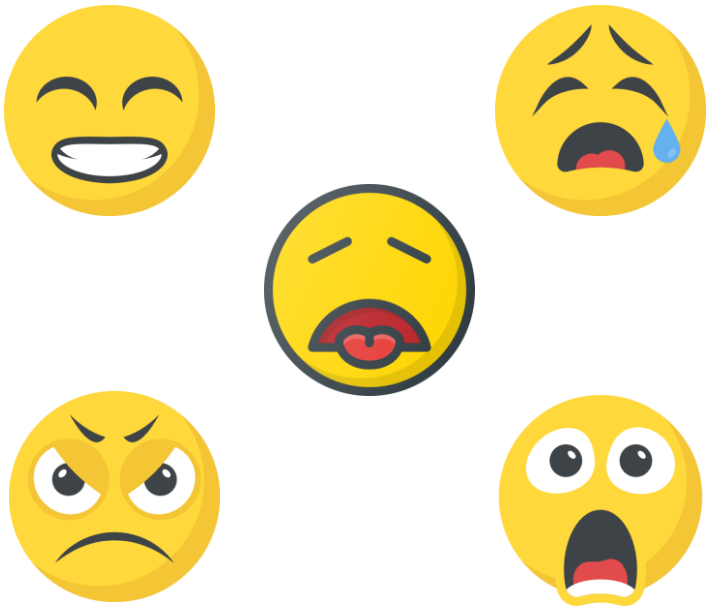
Problem

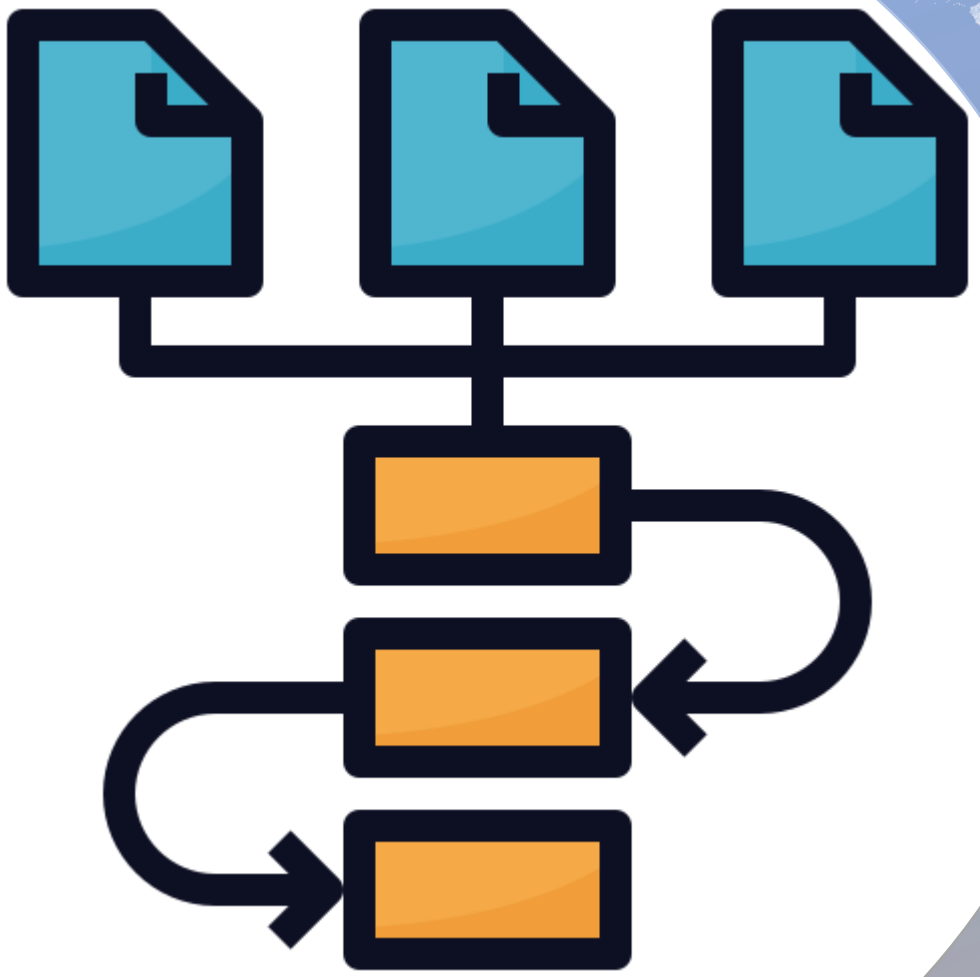
- Detecting the emotion of forensic offenders.
- Analysis of customer satisfaction of any company.
- Satisfaction analysis of students in online/traditional education.
- Analysis of the emotions of candidates in job interviews.



Basic Emotions

- Neutral
- Calm
- Happy
- Sad
- Angry
- Fearful
- Surprise
- Disgust



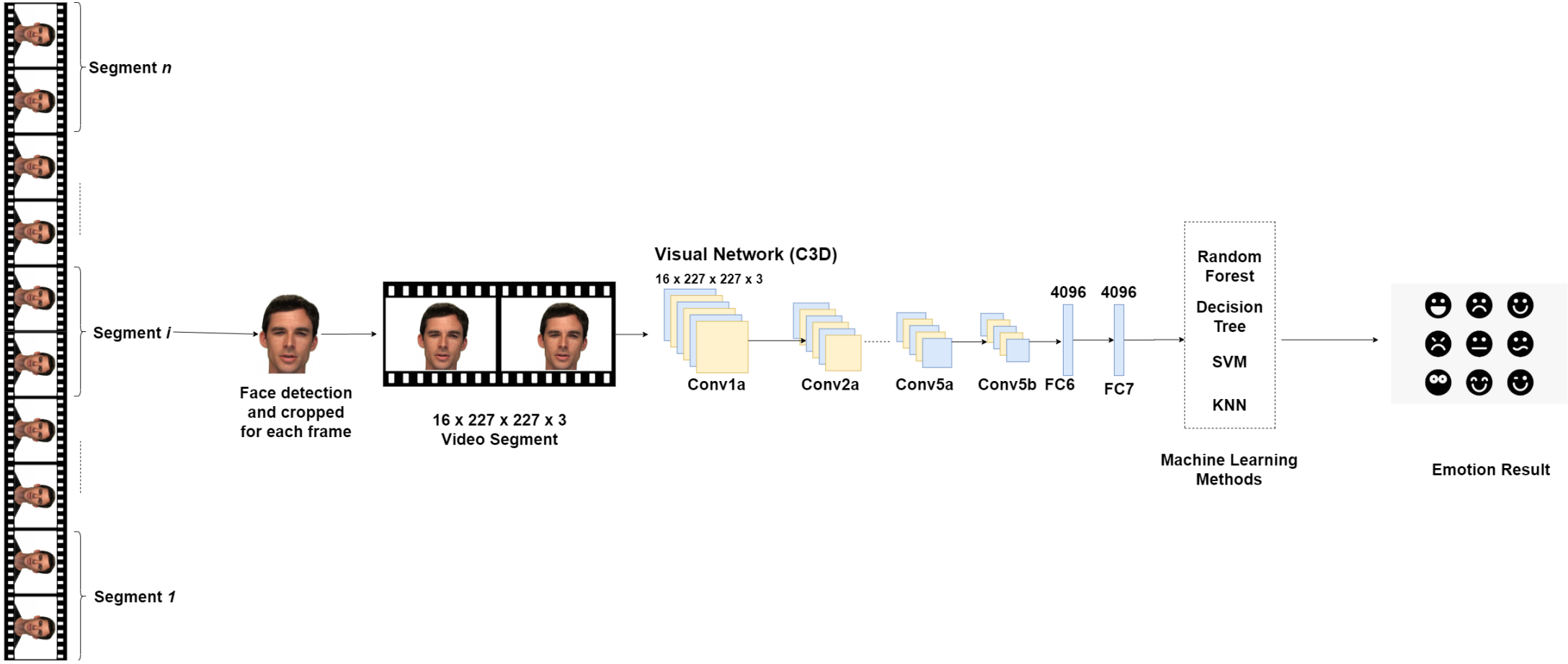


Project Algorithm

Algorithm

- Visual Algorithm
- Audio Algorithm
- Fusion Algorithm

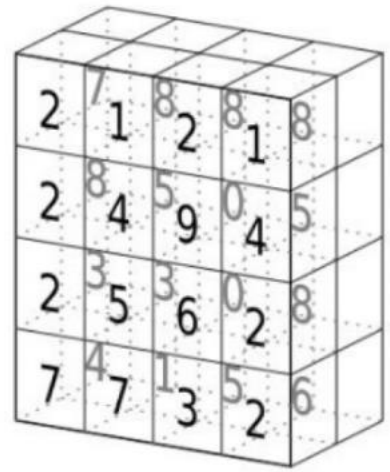
Visual Algorithm



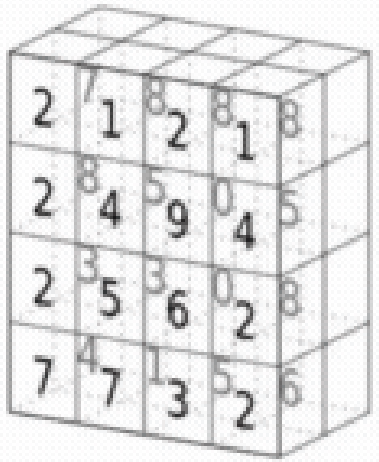
VISUAL PREPROCESSING

Crop & Resize

Converted 16-Frame Segments

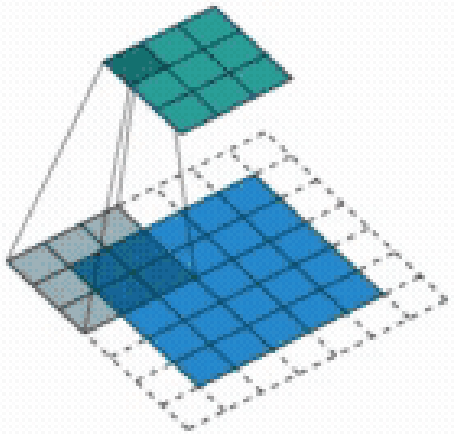


VISUAL TRAINING



Shape: (X, 16, 227, 277, 3)
X: Number of Segment

Input CNN 3D Network



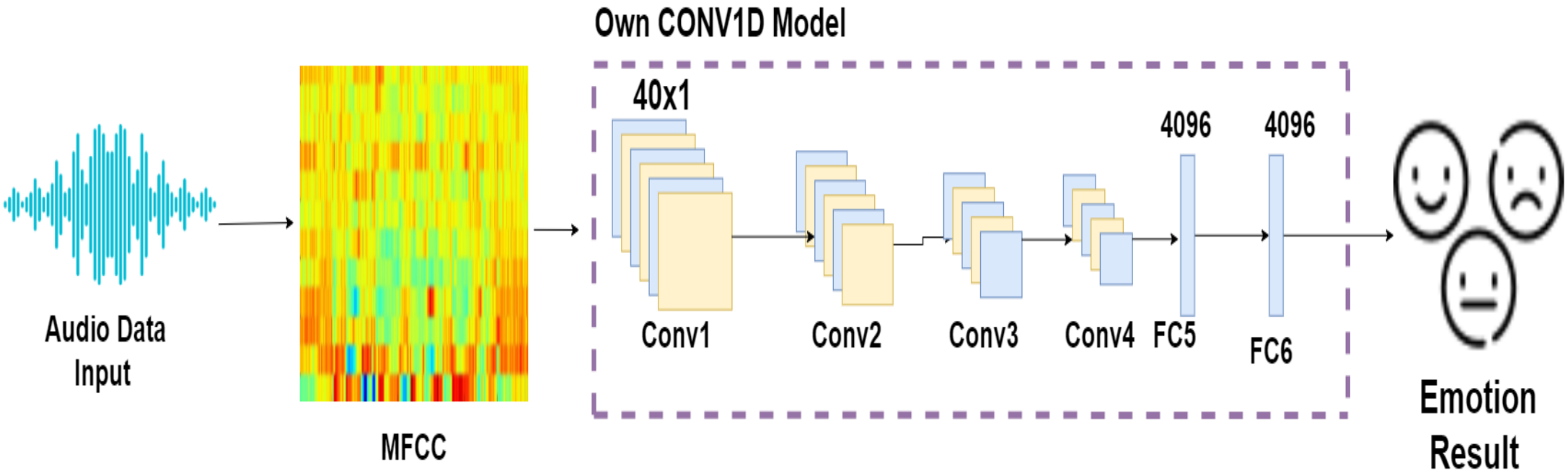
C-3D Model

Output

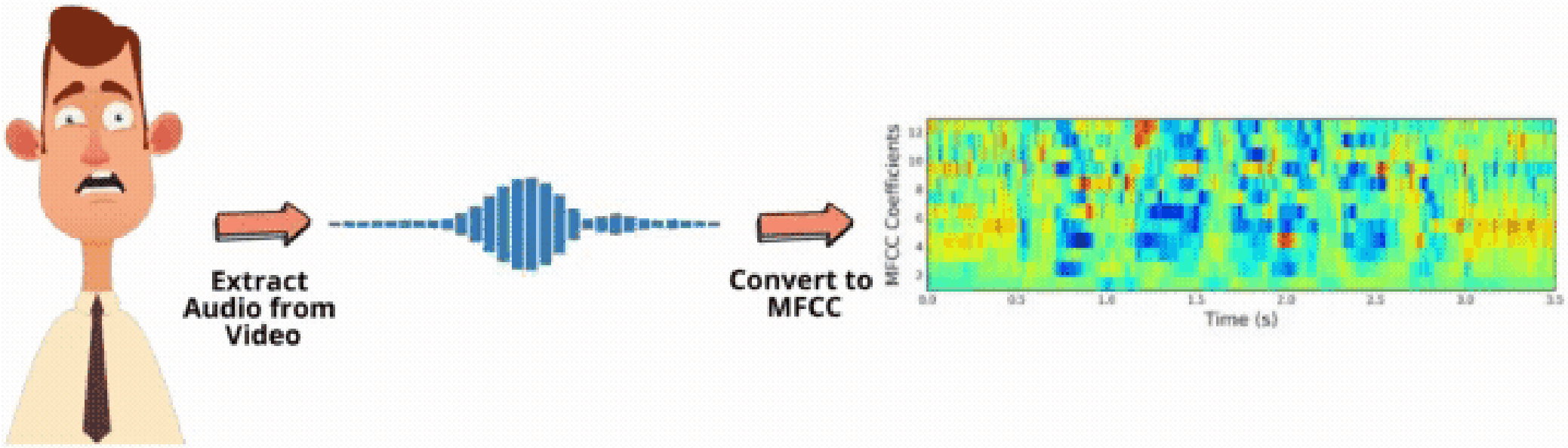
```
[[0.459 0.237 ... 0.0143 0.356]
 [0.417 0.730 ... 0.718 0.462]
 [0.483 0.462 ... 0.155 0.8913]
 ...
 [0.259 0.325 ... 0.922  0.586]
 [0.927 0.071 ... 0.150 0.395]
 [0.974 0.051 ... 0.829 0.603 ]]
```

Shape: (X, 4096)
X: Number of Segment

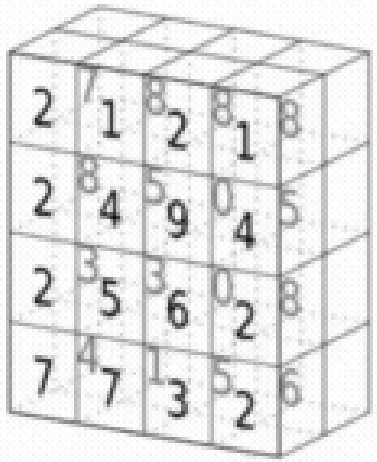
Audio Algorithm



AUDIO PREPROCESSING

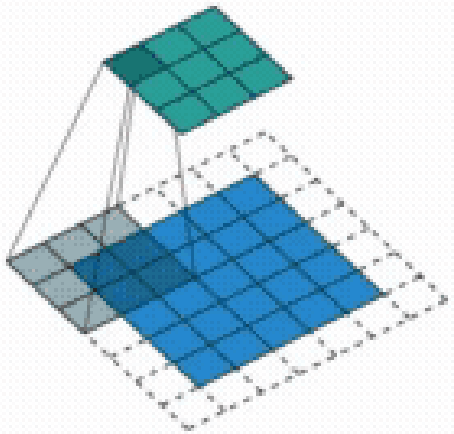


AUDIO TRAINING



Shape: (1, 40)
MFCC Array

Input 1D-CNN Network



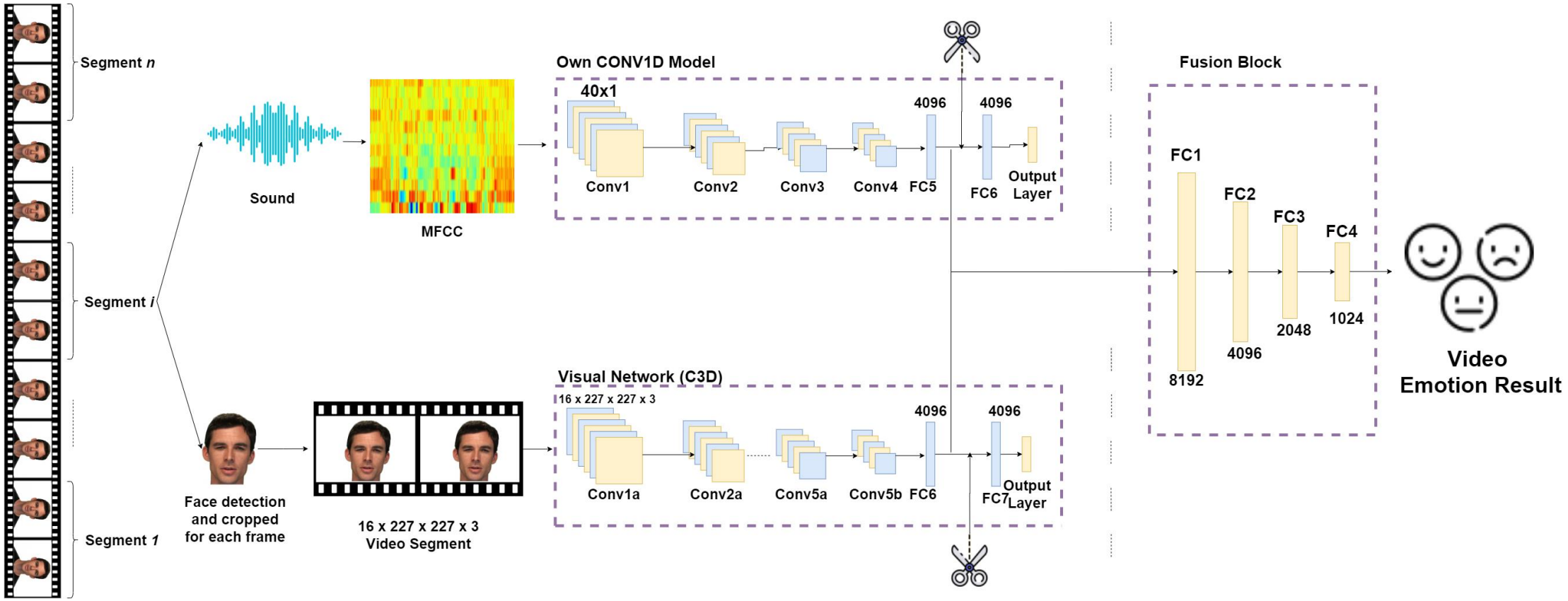
1D-CNN Model

Output

```
[[0.459 0.237 ... 0.0143 0.356]
 [0.417 0.730 ... 0.718 0.462]
 [0.483 0.462 ... 0.155 0.8913]
 ...
 [0.259 0.325 ... 0.922  0.586]
 [0.927 0.071 ... 0.150 0.395]
 [0.974 0.051 ... 0.829 0.603 ]]
```

Shape: (1, 4096)

Fusion Algorithm



FUSION DATA PREPARATION FOR VISUAL

[[0.459 0.237 ... 0.0143 0.356]
[0.417 0.730 ... 0.718 0.462]
[0.483 0.462 ... 0.155 0.8913]
...
[0.259 0.325 ... 0.922 0.586]
[0.927 0.071 ... 0.150 0.395]
[0.974 0.051 ... 0.829 0.603]]

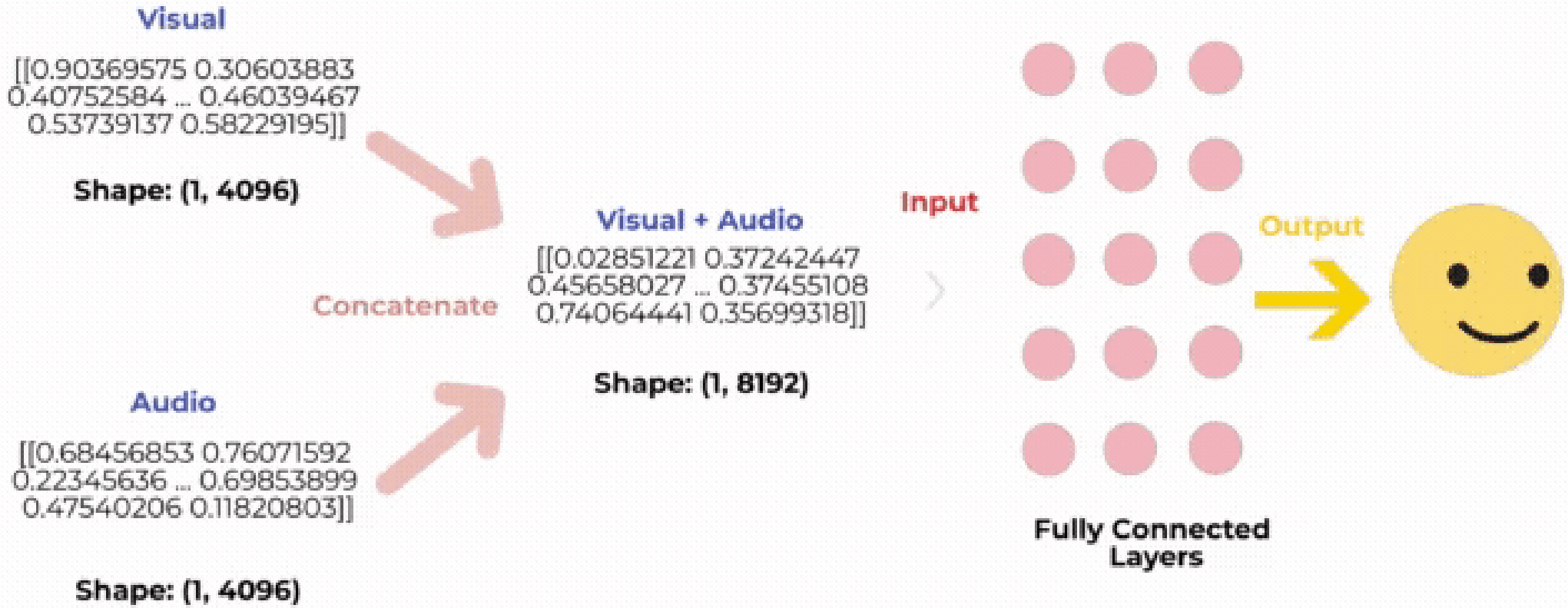
Shape: (X, 4096)
X: Number of Segment

Apply Average Pooling

[[0.90369575 0.30603883
0.40752584 ... 0.46039467
0.53739137 0.58229195]]

Shape: (1, 4096)

FUSION TRAINING



Dataset Contents & Experimental Result

Dataset Contents

RAVDESS (The Ryerson Audio-Visual Database of Emotional Speech and Song)

- **24 actors (12 Female and 12 Male)**
- **Video files are 1280x720 pixels at 30 FPS (HD format, 720p).**
- **Audio files are at 48 kHz sampling frequency (.wav) and 16-bit rate.**
- **Speech Emotions**
  Neutral, Calm, Happy, Sad, Angry, Fearful, Surprise and Disgust.
  Number of videos: 1440
- **Song Emotions**
  Neutral, Calm, Happy, Sad, Angry and Fearful.
  Number of videos: 1012

Experimental Results

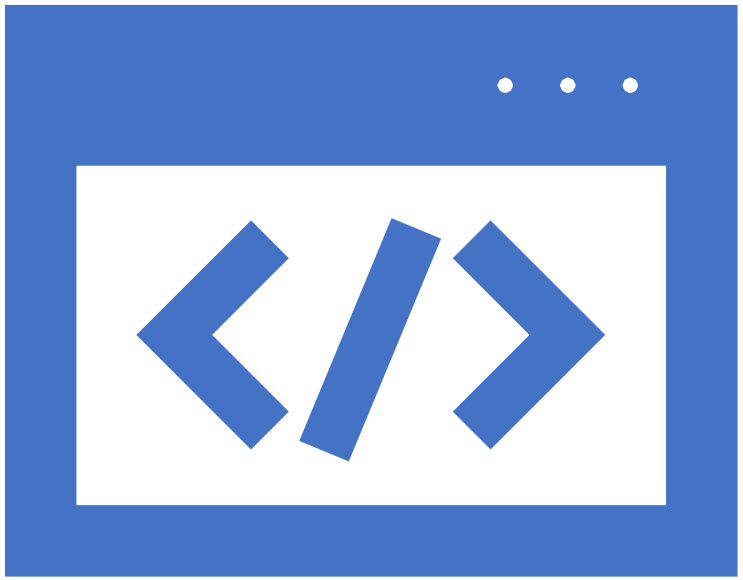
DEEP LEARNING APPLICATION RESULT OF VISUAL			
Dataset		RAVDESS SONG DATASET	RAVDESS SPEECH DATASET
Method	Algorithm	Accuracy (%)	Accuracy (%)
Deep Feature Extraction	SVM	93.30	90.45
	KNN	87.00	86.98
	Random Forest	95.69	95.39
	Decision Tree	73.22	69.51
Transfer Learning	3D CNN	94.12	93.94

DEEP LEARNING APPLICATION RESULT OF AUDIO		
Dataset	RAVDESS SONG DATASET	RAVDESS SPEECH DATASET
Algorithm	Accuracy (%)	Accuracy (%)
1D CNN	71.42	72.56
AlexNet	52.47	53.72
VGG16	51.68	54.12

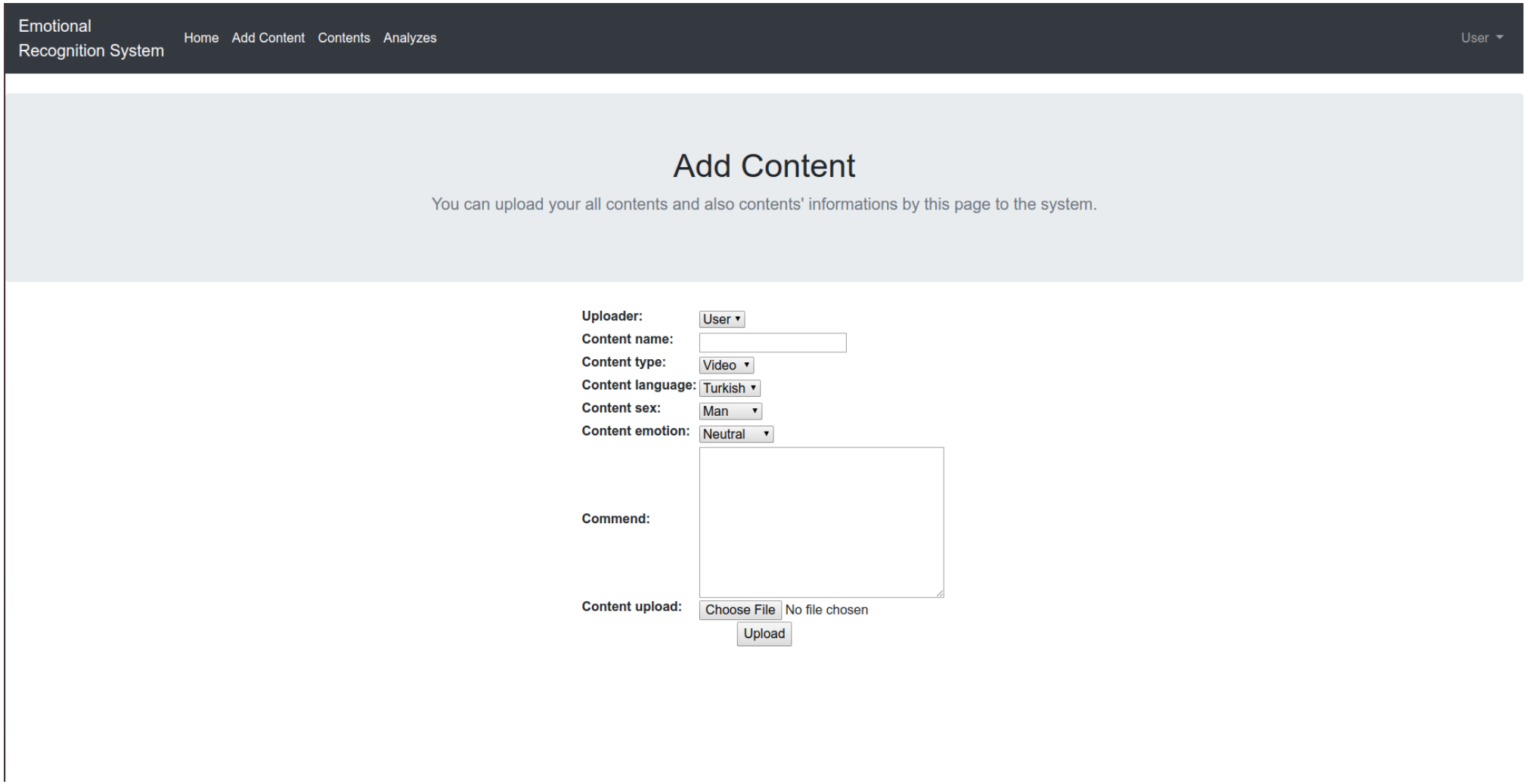
Experimental Results

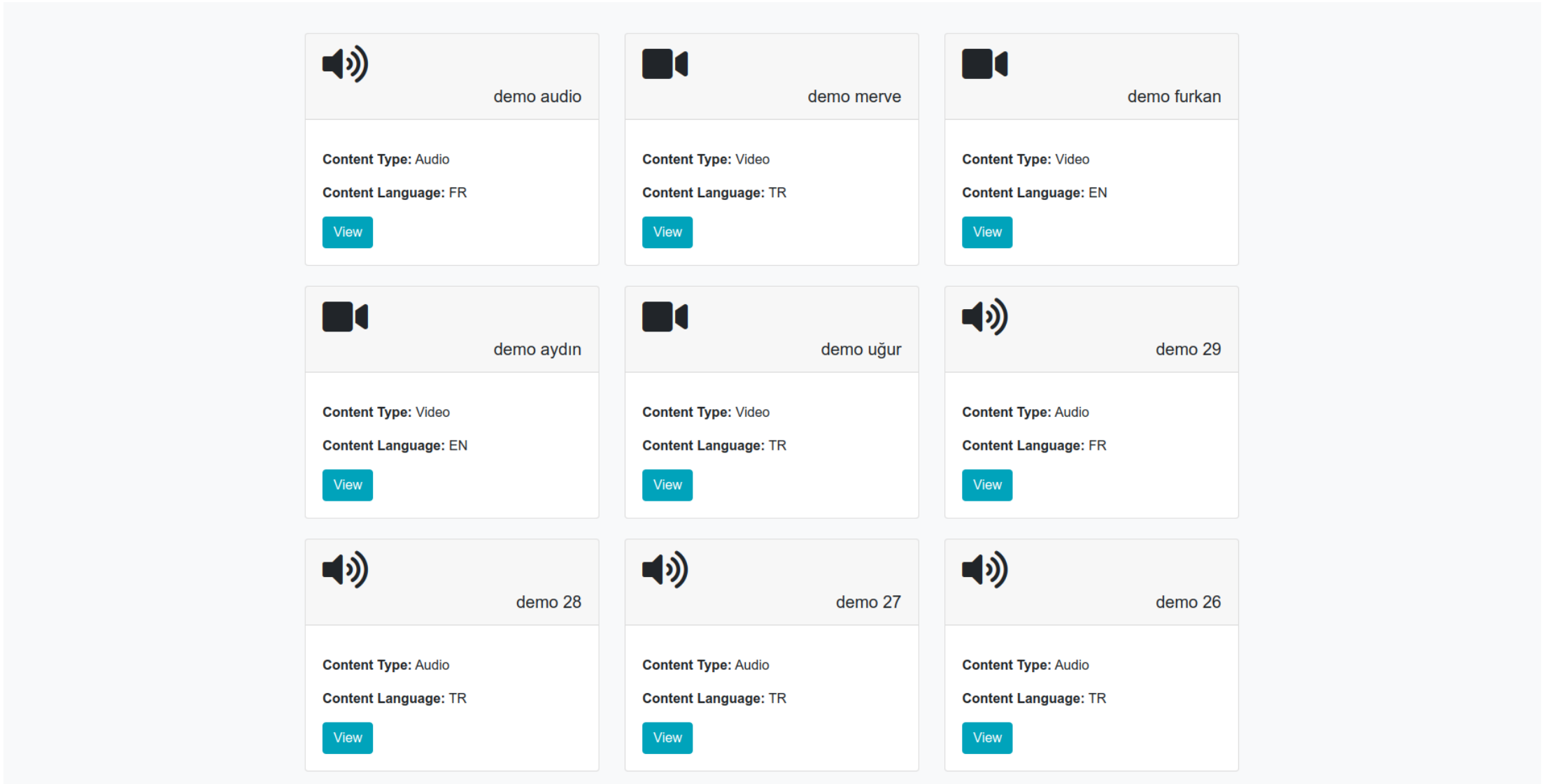
DEEP LEARNING APPLICATION RESULT WITH FUSION		
Dataset	**RAVDESS SONG DATASET**	**RAVDESS SPEECH DATASET**
Algorithm	**Accuracy (%)**	**Accuracy (%)**
Fully Connected Layer	96.80	96.73

NOTE: 80% of dataset have used to train the model and 20% to test it.



GUI DESIGN


Emotional Recognition System
Home
Add Content
Contents
Analyzes
User
Add Content
You can upload your all contents and also contents' informations by this page to the system.
Uploader:
User
Content name:
Content type:
Video
Content language:
Turkish
Content sex:
Man
Content emotion:
Neutral
Commend:
Content upload:
Choose File
No file chosen
Upload


demo audio
Content Type: Audio
Content Language: FR
View
demo merve
Content Type: Video
Content Language: TR
View
demo furkan
Content Type: Video
Content Language: EN
View
demo aydın
Content Type: Video
Content Language: EN
View
demo uğur
Content Type: Video
Content Language: TR
View
demo 29
Content Type: Audio
Content Language: FR
View
demo 28
Content Type: Audio
Content Language: TR
View
demo 27
Content Type: Audio
Content Language: TR
View
demo 26
Content Type: Audio
Content Language: TR
View

demo furkan

Content Uploader: User

Content Type: Video

Content Language: EN

Content Sex: Man

Content Emotion:calm

Commend:demo

Find Emotion (using visual)

Find Emotion (with fusion)

Find Emotion (using visual)

Find Emotion (with fusion)

You can see the emotional state of your content below.

Results:

😐 Calm

demo audio

Content Uploader: User

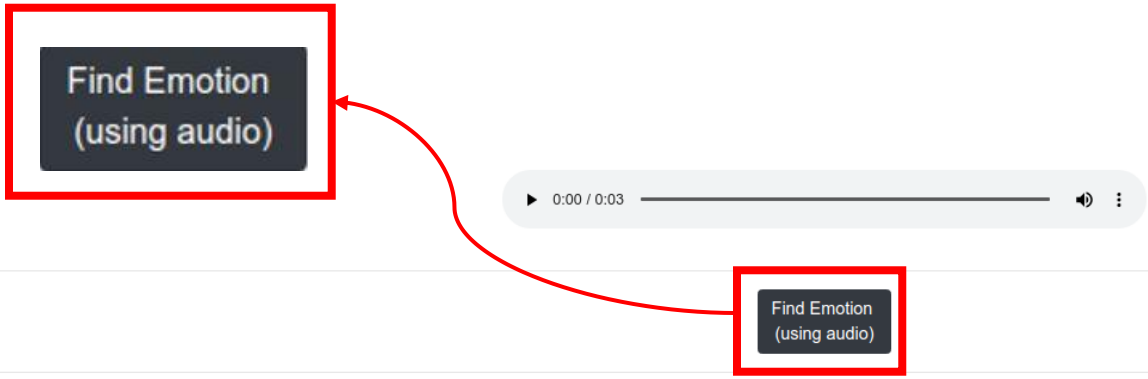
Content Type: Audio

Content Language: FR

Content Sex: Man

Content Emotion:neutral

Commend:demo



You can see the emotional state of your content below.

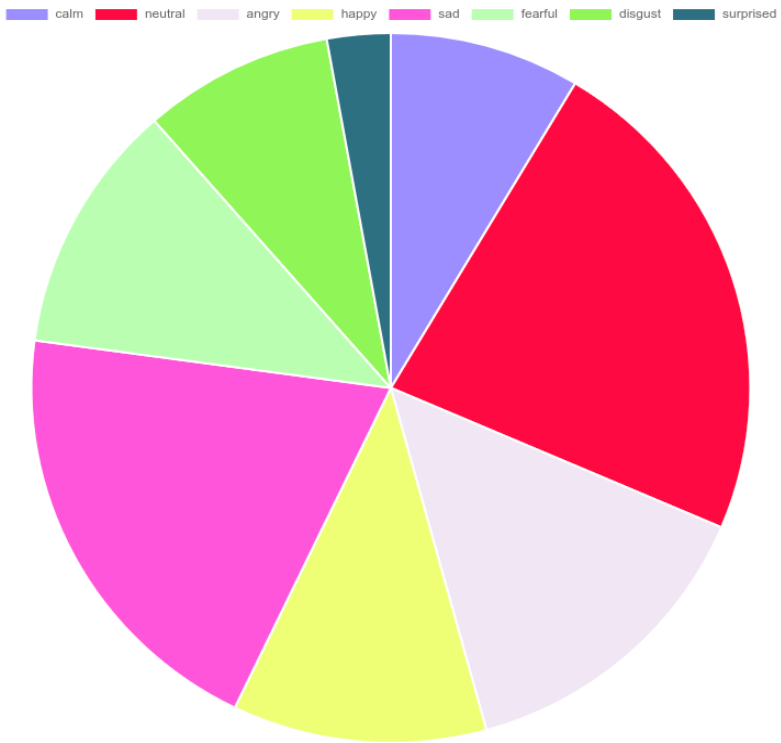
Results:

MLPC Result: sad
LDC Result: sad
KNN Result: neutral
SVC Result: sad

Analyze of contents

You can see all contents analyze.

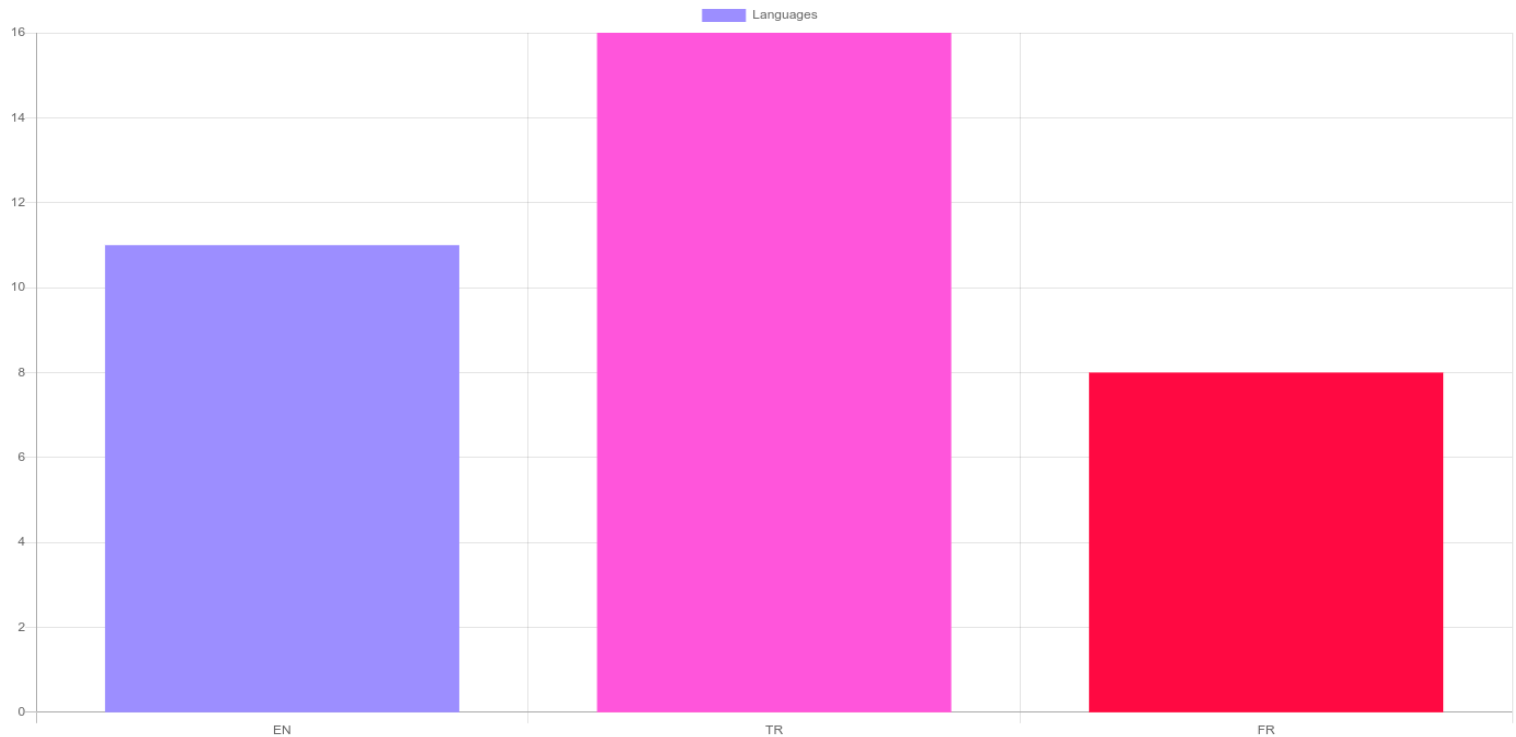
Emotion Analyze | Language Analyze | Gender Analyze



Analyze of contents

You can see all contents analyze.

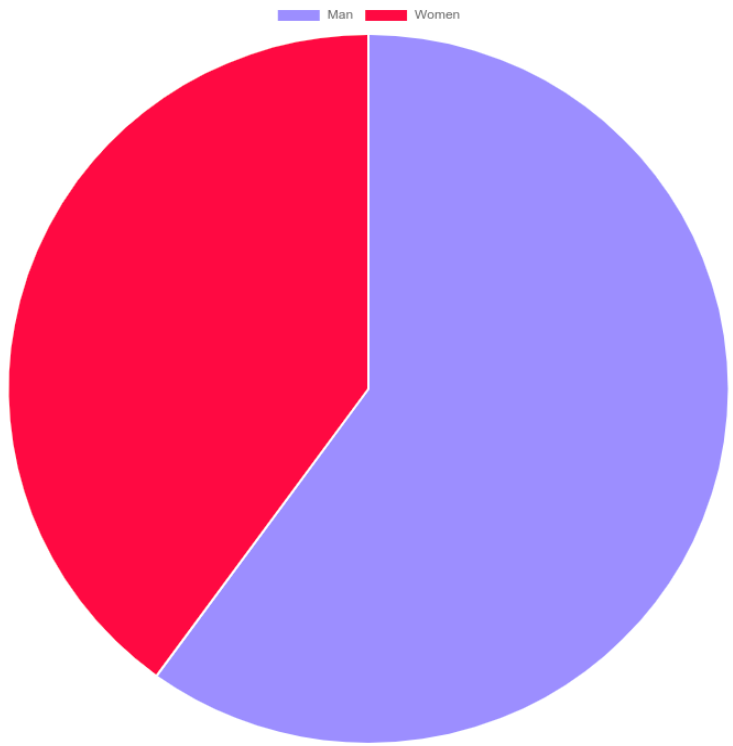
Emotion Analyze | Language Analyze | Gender Analyze



Analyze of contents

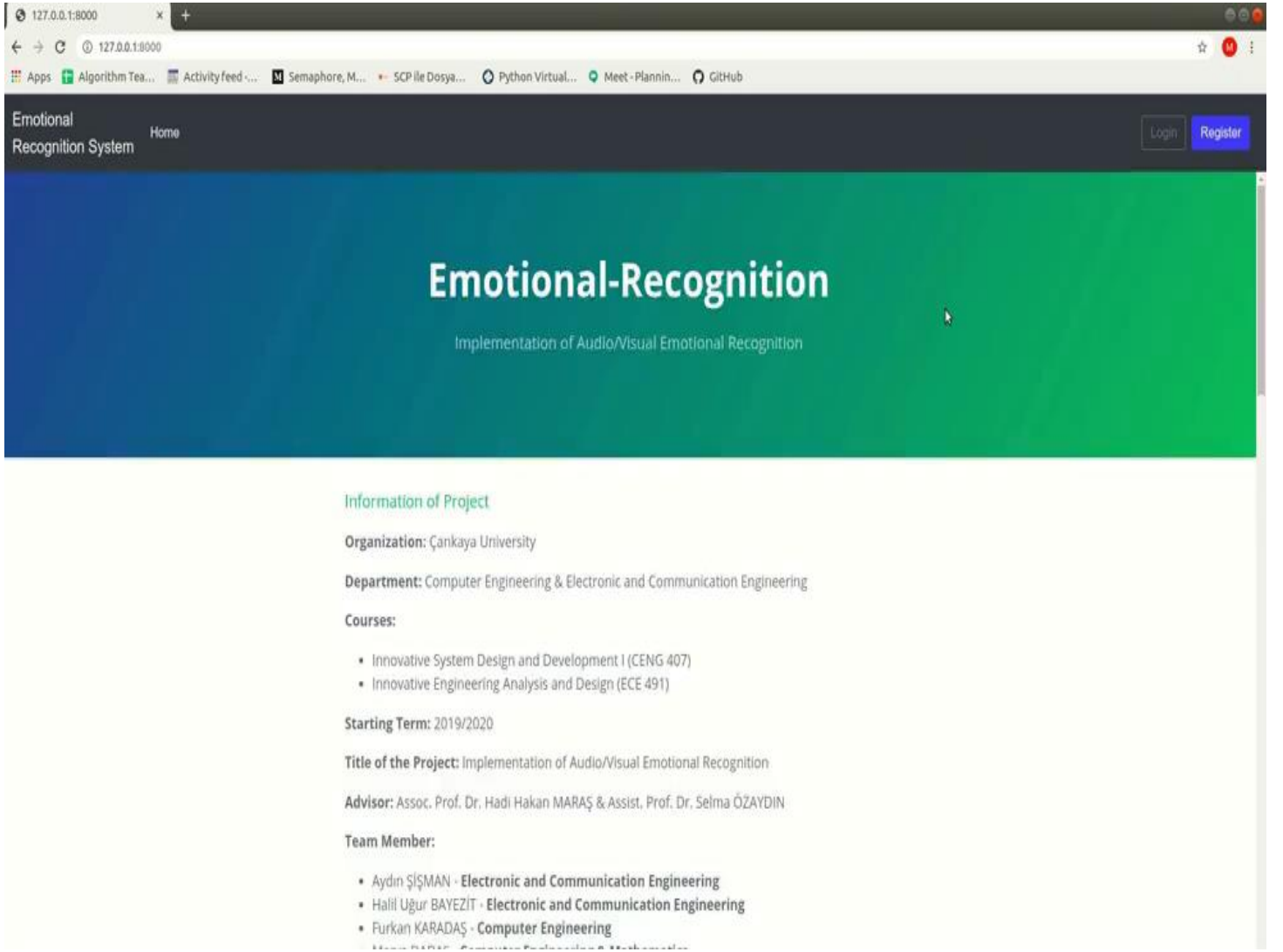
You can see all contents analyze.

Emotion Analyze | Language Analyze | Gender Analyze



- Audio Algorithm
- Visual Algorithm
- Fusion Algorithm
- Emotion Analysis
- GUI Design

Demo


127.0.0.1:8000
Apps
Algorithm Tea...
Activity feed ·...
Semaphore, M...
SCP ile Dosya...
Python Virtual...
Meet - Plannin...
GitHub
Emotional Recognition System
Home
Login
Register
Emotional-Recognition
Implementation of Audio/Visual Emotional Recognition
Information of Project
Organization: Çankaya University
Department: Computer Engineering & Electronic and Communication Engineering
Courses:
Innovative System Design and Development I (CENG 407)
Innovative Engineering Analysis and Design (ECE 491)
Starting Term: 2019/2020
Title of the Project: Implementation of Audio/Visual Emotional Recognition
Advisor: Assoc. Prof. Dr. Hadi Hakan MARAŞ & Assist. Prof. Dr. Selma ÖZAYDIN
Team Member:
Aydın ŞİŞMAN - Electronic and Communication Engineering
Halil Uğur BAYEZİT - Electronic and Communication Engineering
Furkan KARADAŞ - Computer Engineering

Thank you, best team

❤️

QUESTION