

AI Video Generator, including intro, goals, features, architecture, tech stack, and roadmap. It's designed to impress collaborators, recruiters, and future investors alike and target to reach.

Project Name: VidAigen (The AI Video Generator for Everyone)

Vision: VidAigen aims to become the world's most versatile and intelligent AI video generator. It turns **text, images, or ideas into stunning videos using cutting-edge AI — including emotion, voice, style, language, and more.

Whether you're:

1. A job seeker trying to impress employers,
2. A content creator looking to scale your work,
3. A company in need of fast, global content creation,

VidAigen = is here to revolutionize how videos are made.

Core Objectives:

- * Build a collaborative open-source AI project.
- * Deploy a full-stack application with video generation, voice synthesis, style editing, and emotion controls.
- * Add real-world features and polish to impress recruiters, users, and possibly investors.

Core Features:

1. Style Control

Allow users to choose visual styles:

- * Cartoon
- * Cinematic

- * Anime

- * Pencil Sketch

- * Minimalist

****Model Ideas****: Using Stable Video Diffusion (SVD) + ControlNet + Textual Inversion

2. Voiceover Generator:

Transform scripts into natural-sounding voiceovers with accent options:

- * Nigerian

- * Indian

- * American

- * British

- * French

- * Arabic

****Model Ideas****:

- * [Bark](<https://github.com/suno-ai/bark>)

- * [Coqui TTS](<https://github.com/coqui-ai/TTS>)

- * \[ElevenLabs API (if licensed)]

3. Emotion-Based Scene Generator

Users choose emotions:

- * Happy

- * Sad

- * Dark

- * Romantic

- * Motivational

These influence:

- * Scene color grading

- * Background music tone

- * Facial expressions (if applicable)

****Tools**:**

- * Emotion classifier on text input

- * Apply conditioning in video generation model

- * Add background music using HuggingFace audio models

4. Face Animation from Photo

Upload a photo → Animate it to match the video or script.

****Tech Options**:**

- * [SadTalker](<https://github.com/OpenTalker/SadTalker>)

- * [Wav2Lip](<https://github.com/Rudrabha/Wav2Lip>)

5. Multi-Language Support

Input in any language → Generate voiceover & subtitles.

****Implementation**:**

- * Use Google Translate API or MarianMT for translation

- * Use multilingual TTS models for voice generation

- * Auto-generate subtitles

6. Timeline Editor (Frontend UI)

Allow users to:

- * Drag and drop generated scenes

- * Insert transitions

- * Edit voiceover timing

- * Add background music

****Tech Options**:**

- * Streamlit components (initial MVP)

- * Advanced: React + Remotion + Flask/FastAPI backend

7. Resume-to-Video

Upload resume (PDF or text) → Turns into a personal job video:

- * Auto-narration

- * Career highlights

- * Emotional tone (e.g., motivational)

Architecture Overview

Frontend (Streamlit/React)

- └— Upload Text/Image/Resume

- └— Select Options (style, emotion, language)

Backend (FastAPI/Flask)

- └── Text-to-Speech (TTS)
- └── Text-to-Video Generation (Stable Diffusion + Motion LoRA)
- └── Photo Animation (SadTalker/Wav2Lip)
- └── Emotion Classifier (BERT/Sentiment Models)
- └── Language Translator (MarianMT API / Google Translate)

Output

→ Final video (with timeline, voice, animation, emotion, subtitles)

Tech Stack

- | Area and Tools |
- | ----- | ----- |
- | Programming | Python, JavaScript |
- | Video Gen | Stable Video Diffusion, ModelScope, Motion LoRA |
- | Audio | Bark, Coqui TTS, ElevenLabs |
- | Voiceover Sync | Wav2Lip, SadTalker |
- | UI/UX | Streamlit, Gradio (MVP), React (timeline editor) |
- | Backend | FastAPI or Flask |
- | Deployment | Hugging Face Spaces, Streamlit Cloud, Vercel |
- | Storage | Firebase, Supabase, mongodb or AWS S3 |
- | NLP | T5, BERT, OpenAI embeddings (optional) |
- | Translations | MarianMT, Google Translate API |
- | Version Control | Git + GitHub |
- | Collab Tools | Notion + Trello + Google Meet |

Roles & Responsibilities

| Roles

| Project Manager

| AI/Model Research

| Voice & TTS Integration

| Streamlit UI / Gradio

| React Timeline Editor

| Backend Dev

| Video Gen + Styling

| Subtitle & Lang