

Problem Statement

RealAdapt: AI-Based Ad Generation Tool

Objectives

Scalable: Suitable for any product

Accessible: Helps small businesses make ads without a team.

Reduce the average response time

Team Members

Manisha Apchunde

Samiksha More

Firdos Pathan

Rutuja Tikhole

Dataset Link : <https://www.kaggle.com/datasets/paramaggarwal/fashion-product-images-dataset>

This dataset contains over 44,000 fashion product images and metadata such as gender, category, and usage, ideal for training AI models to generate personalized and platform-specific advertisements. Eg, Affectnet, webvid, Fashion dataset.

Benefits and Impact

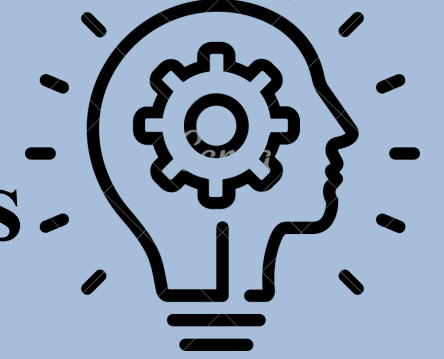
Time & Cost Efficient: Automates content creation process

Accessible: Ads without marketing team

Personalization: Tailored to audience needs

Scalable: Suitable for any product .

Implementation Details



Script generation with LLMs.

Image creation with Stable Diffusion.

Face restoration with GFPGAN (Generative Facial Prior GAN).

Audio with ElevenLabs.

Final composition with MoviePy.

Methodology

Input Layer

- 1.Data Collection:** Gather raw input data (JSON, CSV, text prompts) through manual upload or API integrations.
- 2.Data Preprocessing:** Clean, normalize, and prepare data using Pandas, NumPy, NLTK for consistent quality.
- 3.Script Generation:** Use OpenAI GPT, LLaMA to automatically generate scene scripts and instructions from processed data.

Processing Layer

- 1.Scene Generation:** Apply Stable Diffusion to convert scripts into visual frames or scenes.
- 2.Face Restoration:** Improve facial details using GFPGAN to ensure realistic visuals.
- 3.Upscaling:** Enhance the resolution and clarity of generated frames via Real-ESRGAN.

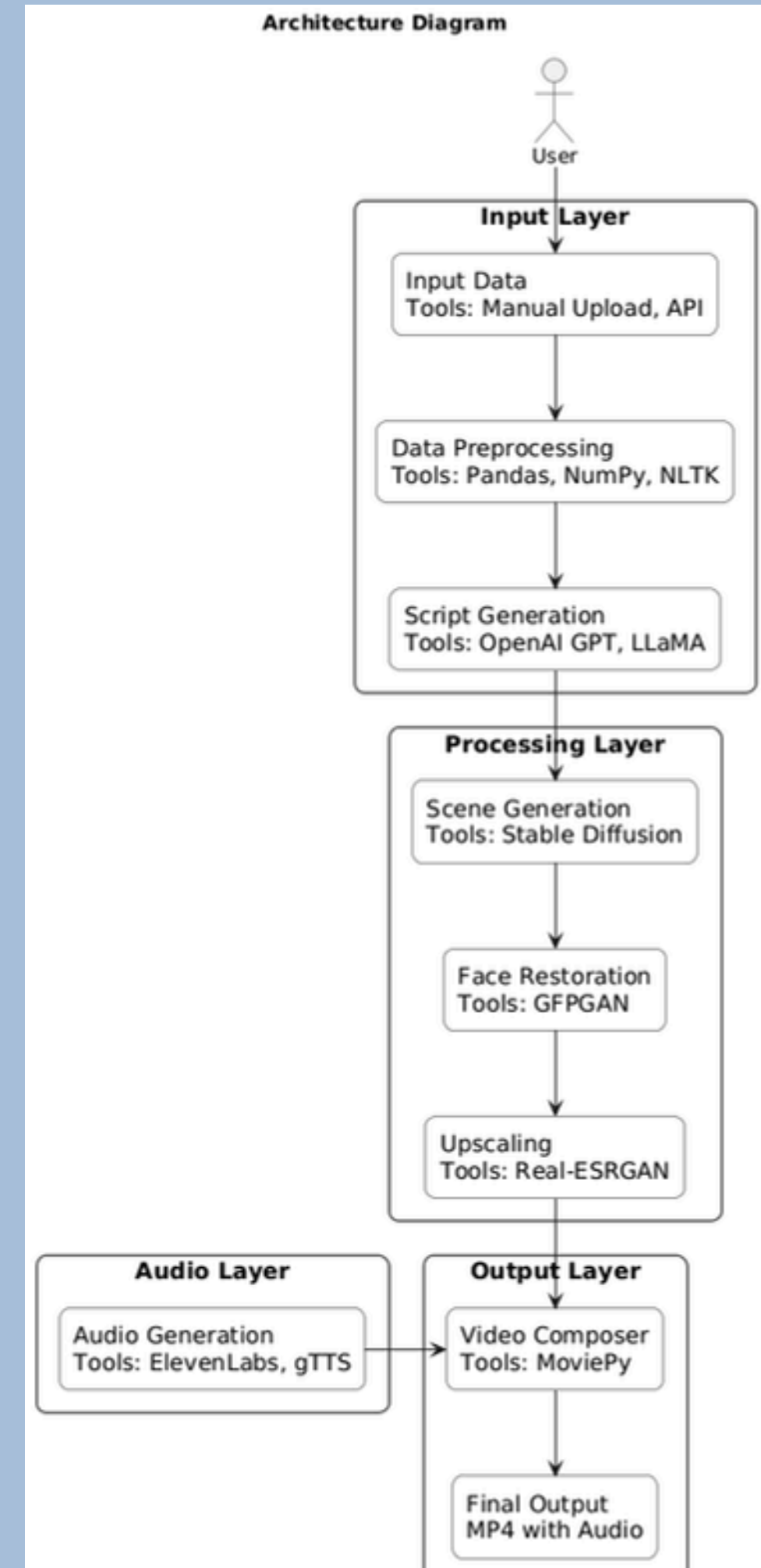
Audio Layer

- 1.Audio Generation:** Utilize ElevenLabs, gTTS for converting scripts into voiceovers or narration.
- 2.Video Composition:** Merge frames and audio tracks using MoviePy to produce a cohesive video.
- 3.Final Output:** Export the finished video in MP4 format with synchronized audio.

Workflow Flow

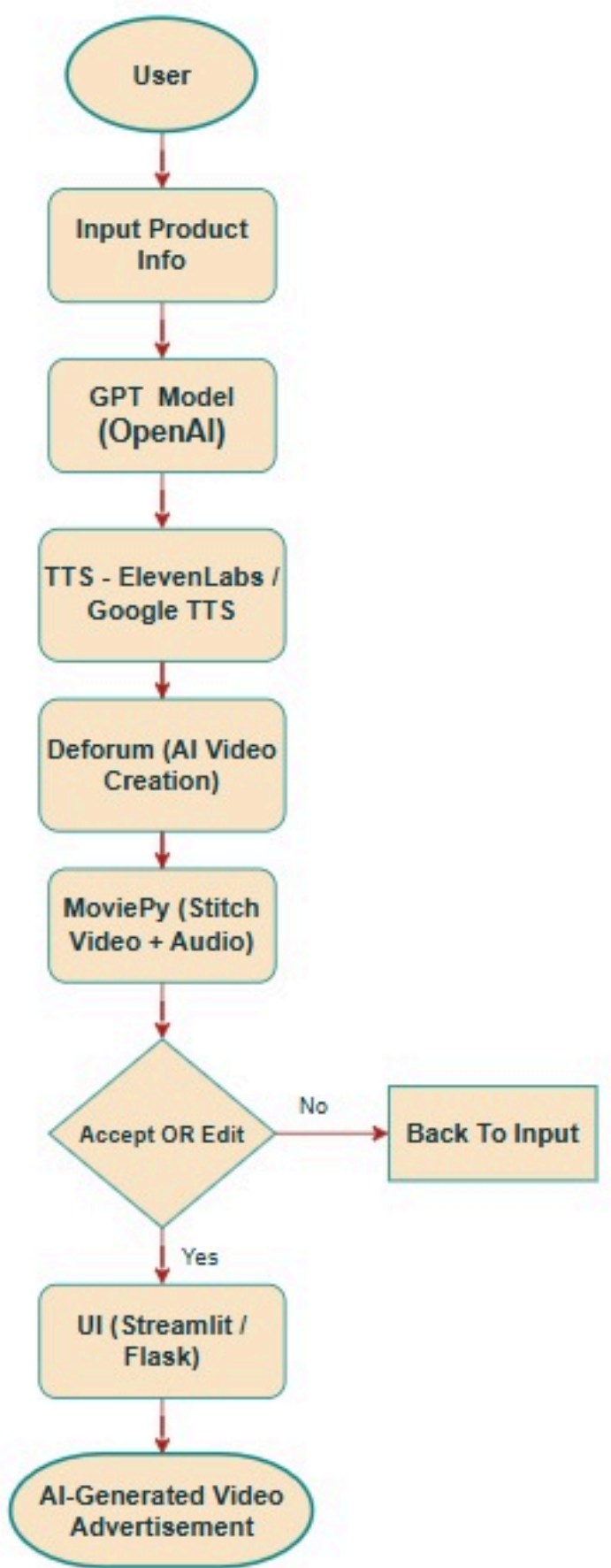
• User → Input Layer → Processing Layer → Audio Layer → Output Layer → Final Video

Figure 2: Architecture Diagram




Literature Survey			
Title	Author(s)	Conclusion	Comparison with Our Project
Make-A-Video: Text-to-Video Generation without Paired Data	Yossi Singer et al., Meta AI, 2022	Introduces video generation from text without paired video data. Does not support emotional tone, branding, or ad-specific storytelling.	Our system adds product personalization, emotional tone scripting, and full ad structuring including voice and UI.
Text-to-Video-Zero: Text-to-Image Diffusion Models are Zero-Shot Video Generators	Feng Yang et al., Shanghai AI Lab & Tencent AI, 2023	Zero-shot video generation from text using image models. Lacks personalization and emotional depth.	We enhance personalization based on user input and add emotion-aware voiceover + visual alignment.
Emotion-aware Advertising Text Generation	Yue Wang et al., Peking University, 2021	Generates ad text using emotional cues, but doesn't include visuals or audio/video.	We generate full ad videos—script, visuals, voice—with emotion-targeted scenes, not just text.
Deforum – AI Video Generation using Stable Diffusion	Community (Open Source), 2022	Deforum creates AI videos from text prompts using Stable Diffusion, but it needs manual input and lacks automatic scene flow or voice support.	We automate scene-wise script → visual generation without user prompt effort.

Flow Chart



Tech Stack



Streamlite



HuggingFace /Gemini

Deforum, GAN

ElevenLabs/Google TTS

moviepy / ffmpeg

Google Colab



Future Scope

- 1-Cross-Platform Adaptability** – Ads can be auto-optimized for YouTube, Instagram, TikTok, OTT platforms, and even AR/VR.
- 2-Data-Driven Iteration** – Integration with analytics allows performance tracking and automatic regeneration of improved ad versions.
- 3-Cost-Effective Creativity** – Small and medium businesses gain access to ad production quality .
- 4-Future-Proof Expansion** – Include AI avatars, 3D scene generation, real-time adaptive ads, and immersive metaverse-ready experiences.

OUTPUT

Advertisement
For electric
vehicle
Named
“Yello”
Electric



© 2014 Yello Electric Vehicle Co. All rights reserved.

OUTPUT

**We have
just given
1 scene
scooter
advertisement
Requirement
and
4-5 second
advertisement
timeline**



Marketing- tools where suscription to remove ads need

Conclusion

- Supports scalable, personalized digital marketing. Also save human efforts.
- Makes professional ads easy, fast, and affordable as well as saves time.
- Enables businesses to create targeted ads without tech skills.
- Ready-to-use MP4 video ad, optimized for Instagram like platforms.

References

- US20250156898A1- AI-driven tailored ad generation via symbolic + connectionist AI . Link- US20250156898A1 - System and method for ai based tailored advertising content generation - Google Patents
- WO2024112129A1- Automated video ad generation with personalization/localization.Link- WO2024112129A1 - Automated video advertisement generation method, and system therefor - Google Patents
- US20200279303A1- Job-based remote compute pipeline for ad video creation.Link- US20200279303A1 - Automated video advertisement creation - Google Patents

Thank You !