Research Task 06 — AI-Generated Deep Fake Interview

**Siddharth Savla**

# 1. Goal of the Task

The goal of this task was to transform insights from Task 05 (Syracuse Men’s Soccer 2024 dataset) into an AI-generated deep fake interview. The objective was to simulate an ESPN-style interview with the Syracuse coach, summarizing season performance, defensive strategies, and key players. Additionally, this task aimed to explore AI tools, understand their capabilities and limitations, and document the complete workflow.

# 2. How I Approached the Task

I started by reviewing the Task 06 instructions carefully. Initially, I thought I would create one AI video, but I realized the professor’s focus was on documenting process, tool exploration, and learning outcomes.

## Step 1 - Understanding the Requirements

Reviewed the instructions thoroughly and identified that the emphasis was on workflow and tool usage rather than just producing a single final video.

## Step 2 - Exploring Tools and Facing Limitations

After reviewing the Task 06 instructions, I started brainstorming options for generating an AI-based interview.  
Since I wasn’t sure which tools would work best, I discussed multiple possibilities with ChatGPT to get recommendations.

Together, we explored several AI video generation tools:

* **D-ID & HeyGen** → ChatGPT suggested these as popular tools for creating realistic talking-head AI videos. I tested both, but they required paid subscriptions to export HD clips without watermarks, so I decided to skip them.
* **Synthesia** → Looked promising for generating professional studio-style interviews, but the free version was too limited, and the student plan was unavailable.
* **Gemini Veo 3** → Based on ChatGPT’s advice, I finalized Gemini Veo 3 Fast because I had a Gemini Pro subscription and saw the YouTube link provided in the task documents, also using Google Veo, which produces high-quality videos, and supports dynamic backgrounds and lip-syncing.

However, while using Gemini Veo 3 Fast, I faced some limitations:

* **Timeout restrictions**: I had to wait 3 hours between generating new videos due to Google’s usage limits.
* **8-second limit per clip**: I couldn’t generate one long interview. ChatGPT helped me design a workaround by splitting the interview into three prompts and later merging them with the tools.
* **File uploads not supported**: Initially, I wanted Gemini to analyze my Task 05 dataset automatically, but Veo 3 doesn’t allow .xlsx or .py uploads. With the analysis I did for Task 05 and with the help of ChatGPT, I suggested to summarize my Task 05 analysis manually and embed those insights directly into the prompts.

This approach ensured that the generated interview was fact-based, accurate, and aligned with my previous dataset analysis.

## Step 3 - Building the Script and Prompts

Once Gemini Veo 3 Fast was finalized, I worked with ChatGPT to design an ESPN-style interview script.  
The goal was to make the coach’s responses sound realistic and data-driven, so I incorporated my Task 05 analysis directly into the prompts:

* Syracuse scored and conceded the same number of goals last season.
* The shot conversion rate was lower than that of opponents.
* Recommendation: strengthen the defense to improve the win rate by ~30%.
* Highlight key players: Gabe Threadgold (scoring) and Michael Acquah (assists).

Initially, my first few prompts failed because the generated video didn’t show an interviewer and felt unnatural.  
After discussing this with ChatGPT, we realized that I needed to add more scene details to make the responses context-rich and authentic:

* Camera angles (e.g., side-by-side interview shots).
* Background settings (stadium, locker room, press conference).
* Crowd effects and lighting details for realism.
* Add a coach image to make it look realistic

Through these iterations, I refined the prompts and decided to split the interview into three 8-second parts due to Veo’s limitations:

1. Season Recap: Coach summarizes performance and stats.
2. Defensive Strategy & Key Players: Coach explains plans and highlights top performers.
3. Closing Motivational Message: Coach shares goals and confidence for next season.

## Step 4 - Generating the Clips

After finalizing the prompts with ChatGPT, I moved to Gemini Veo 3 Fast to generate the interview videos.  
This stage involved a lot of testing, refining, and iteration to get a realistic output.

* Uploaded the Syracuse coach’s photo. This helped Veo create a realistic avatar with accurate lip-sync and expressions.
* Generated three separate clips instead of one long video due to Veo’s 8-second limit per video.
* The first few generations looked too static and lacked interview realism, the background didn’t match the tone.
* To fix this, I started experimenting with different scene setups in each prompt, which improved the visual flow significantly:
  + Clip 1: Outdoor stadium with Syracuse banners, crowd noise, and natural lighting.
  + Clip 2: Indoor press conference setup with microphones and branding.
  + Clip 3: Locker room environment with players celebrating slightly blurred in the background.

By changing scenes across the clips, I made the final interview look dynamic and visually engaging, which gave it a professional ESPN-style vibe.

## Step 5 - Merging the Final Interview

Once I had all three clips generated from **Gemini Veo 3 Fast**, the next step was to **combine them into a seamless ESPN-style interview video**.  
For this, I used **Cap Cut**, as it’s free, intuitive, and offers a lot of customization options.

**Editing Workflow: Imported the three generated clips** into CapCut and arranged them in sequence to maintain a natural interview flow **(Did it for my interest!)**

* **Added captions** to make it look like a **professional ESPN interview with subtitles**.
* **Applied smooth transitions** like fade-ins.
* **Final Output:** Exported the merged video with a total runtime of **~25 seconds**.

By merging the clips thoughtfully, the final video looks **polished and dynamic** while still showcasing the insights from **Task 05**.

# 3. Prompts Used in Gemini Veo 3

See attached document for full prompt details provided in the submission.

# 4. Tools Explored and Used

|  |  |  |  |
| --- | --- | --- | --- |
| Tool | Purpose | Final Use | Outcome |
| Gemini Veo 3 Fast | AI video generation | ✅ Yes | Generated 3 clips |
| Coach Image | Improve realism | ✅ Yes | Realistic lip sync |
| CapCut / Canva | Merging + editing | ✅ Yes | Final 25-sec video |
| D-ID / HeyGen | Alternative tools | ❌ Explored only | Skipped due to limits |
| Python | Task 05 insights | ✅ Yes | Provided accurate stats |

# 5. Challenges & Learnings

Challenges Faced:

• Explored multiple tools like D-ID and HeyGen but finalized Gemini Veo 3 Fast due to free access and better features.  
• Faced 3-hour generation limits and 8-second clip restrictions, requiring the interview to be split into three parts.  
• Could not upload Task 05 dataset directly to Gemini Veo, so I manually summarized insights into the prompts.  
• Early outputs lacked realism, requiring several to achieve a natural ESPN-style interview.

Key Learnings  
• Accurate stats and context-rich prompts from Task 05 led to better, fact-based outputs.  
• Using the Syracuse coach’s image improved lip-sync, expressions, and overall realism.  
• Combining multiple short clips in CapCut with transitions, overlays, and branding created a professional ESPN-style output.  
• Learned how scene details, lighting, and audience effects can significantly enhance video quality.

# 6. Final Output Reflection

This task demonstrated how powerful AI tools like Gemini Veo 3 can be, but also showed that achieving high-quality results requires well-crafted prompts, accurate analysis, iterative refinement, and thoughtful post-production editing.