iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79847

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

AI Based Video Advertisement Generation Using AI

Dr. Prerna N. Khairnar[1], Ms. Pranjal S. Elamame[2], Nikita K. Toche[3], Ms. Anuradha B. Pandav[4], Ms. Namarata N. Magar[5], Ms. Rushali A. Dhamale[6]

[1] Assistant Professor, Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Nashik, Maharashtra, India
[2, 3, 4, 5] Student, Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Nashik, Maharashtra, India

Abstract: The fast growth of Artificial Intelligence (AI) has opened new possibilities in digital marketing, media production, and content authenticity checking. This paper presents a single, end-to-end system that tackles three connected challenges: automated video advertisement creation, AI-assisted logo building, and multi-modal fake video detection. The proposed architecture combines a Transformer-based Natural Language Processing (NLP) engine for marketing script writing, Stable Diffusion models for scene-level visual creation, and an FFmpeg-powered pipeline for video rendering. For verification, the system uses a hybrid deep learning and heuristic approach—combining EfficientNet-B4 with Bidirectional LSTM for deepfake detection, CLIP zero-shot classification for brand logo checking, and YOLOv8 for real-time logo detection. Heuristic methods including Error Level Analysis (ELA), Fast Fourier Transform (FFT) frequency analysis, noise profiling, and Structural Similarity Index (SSIM) support the neural pipeline to flag AI-made artifacts. The system runs through a FastAPI backend and a React/Next.js frontend, supporting real-time advertisement creation and authenticity reporting. Test results show that this combined approach produces high-quality advertisement output while reliably detecting manipulated or AI-generated content. This work highlights the potential and ethical need to build generation and verification tools together.

Keywords: Generative AI, Video Advertisement Generation, Deepfake Detection, Logo Verification, Stable Diffusion, EfficientNet, CLIP, YOLOv8, NLP, FastAPI.

I. INTRODUCTION

The digital advertising space has changed greatly with the arrival of AI-driven content creation tools. In the past, making a video advertisement needed large budgets, a team of creative professionals, and several days of production time. Generative AI has disrupted this model by enabling automatic, scalable, and personalized ad creation from simple product data [1]. At the same time, the rise of AI-generated and deepfake content has created a strong demand for reliable verification tools that can check visual media before it reaches viewers [2].

Generative Adversarial Networks (GANs) were among the first model types to show that machine-generated video content was possible [3]. More recently, diffusion-based systems—such as OpenAI's Sora—have set new standards in video generation quality and frame consistency [4]. In parallel, Large Language Models (LLMs) trained on domain-specific data have shown strong ability to generate marketing copy, slogans, and scene descriptions that match the quality of human writers [5].

However, these generative tools also introduce a serious problem: the ease with which fake media—including fabricated brand advertisements, deepfake endorsements, and AI-placed logos—can be created and spread [6]. Deepfake videos in particular pose a major risk to brand reputation, consumer trust, and the accuracy of public information [7]. As a result, researchers have increasingly focused on deepfake detection methods, ranging from classical signal processing techniques to advanced deep learning classifiers [8].

This paper introduces a fully combined system built to handle both sides of this challenge. On the creation side, the system automates video advertisement production from product data, using a step-by-step pipeline covering script writing, visual synthesis, voice narration, and video assembly. On the checking side, a multi-layer Fake Detection and Brand Authenticity Suite determines whether a given advertisement is real, AI-generated, or a deepfake.

The main contributions of this work are:

1) An end-to-end, modular video advertisement creation pipeline combining LLMs, Stable Diffusion, gTTS, and FFmpeg.
2) A hybrid fake video detection engine combining EfficientNet-B4 + Bi-LSTM with signal-processing heuristics (ELA, FFT, SSIM, Noise Analysis).
3) A zero-shot brand logo verification module using OpenAI's CLIP and YOLOv8 without brand-specific training data.

4) A full-stack deployment built on FastAPI, React, and Next.js 14, designed for real-time advertisement production and verification.

The rest of this paper is laid out as follows: Section 2 reviews related work. Section 3 describes the system architecture. Section 4 details each module's implementation. Section 5 covers the verification suite. Section 6 discusses results and limitations. Section 7 presents conclusions and future directions.

II. LITERATURE SURVEY

The following table synthesizes key prior works that inform and motivate the design of our system across its three principal domains: AI-based video generation, deepfake detection, and logo/brand verification.

Ref. No.	Authors & Year	Study Focus & Methodology	Key Findings	Relevance to This Work
[1]	Kim & Hwang (2025), IJCRT	Personalized video ad creation using Generative Adversarial Networks (GANs). Studied generator-discriminator dynamics for targeted marketing content.	GANs enable scalable personalization; privacy and hallucination remain key risks.	Motivates GAN/diffusion-based video generation pipeline for our ad engine.
[2]	Xie et al. (2025), Discover Computing	Bibliometric review (2020–2025) of AI video generation covering GANs, VAEs, diffusion models, and transformer-based architectures using 422 papers.	Deepfake detection, ethical governance, and temporal coherence are top open challenges in the field.	Provides architectural taxonomy; confirms EfficientNet + LSTM as standard detection baseline.
[3]	Reed et al. (2016); Goodfellow et al. (2014)	Foundational GAN architecture; text-to-image and text-to-video generation using adversarial training.	Generator-discriminator competition yields increasingly realistic synthetic outputs.	Baseline generation paradigm adopted and extended by our vision engine.
[4]	OpenAI Sora (2024); Zhu et al. (2024)	Diffusion Transformer (DiT) operating on spacetime latent patches; minute-level video synthesis from text prompts.	Sora sets SOTA for text-to-video quality; establishes Stable Diffusion as viable API-accessible alternative.	Our system uses Stable Diffusion v1.5/XL via REST API as the scene-level image synthesis backbone.
[5]	Bengesi et al. (2023)	Comprehensive review of Generative AI: GANs, GPT, Autoencoders, Diffusion Models, Transformers—covering architecture, training, and applications.	Transformer decoders with LoRA fine-tuning achieve domain-specific text generation superior to zero-shot prompting.	Validates our use of LoRA-fine-tuned Transformer decoder for marketing script generation.
[6]	Kavinkumar M. & Kanishka R. G. (2025)	Survey of how Generative AI is reshaping content creation pipelines across industries, including marketing and advertising.	AI-generated content now rivals human-authored work in engagement metrics but raises authenticity concerns.	Underpins the dual necessity of our generation + verification architecture.
[7]	Veerasamy et al. (2023); Songja et al. (2024)	Governance and societal impact of deepfake technology. Investigated technical, economic, and reputational threats of AI-fabricated media.	Deepfake production costs have dropped dramatically; detection lags behind generation capabilities.	Justifies the inclusion of a dedicated multi-layer verification suite in our system.
[8]	Sharma et al. (2024)	Overview of deepfake detection models and benchmark datasets. Analysis of CNN-based and hybrid spatial-temporal detection approaches.	EfficientNet-based backbones consistently outperform earlier CNN classifiers on standard deepfake datasets.	Directly informs our choice of EfficientNet-B4 as the spatial feature extractor in the fake detection module.

[9]	Shen et al. (2024)	Personalized Multimodal Generation (PMG) framework using LLMs to map user behavior into natural language preference summaries that guide content creation.	LLM-driven preference summarization significantly improves relevance of generated advertising content.	Informs our NLP script generation strategy, particularly the word-count synchronization formula.
[10]	Hao et al. (2024)	Survey on Generative AI and LLMs for video generation, understanding, and streaming; covers temporal modeling and content alignment.	Bi-directional LSTM aggregation over frame-level CNN features is effective for temporal artifact detection.	Validates our temporal aggregator design (Bi-LSTM over EfficientNet frame features) for deepfake detection.
[11]	Ercan et al. (2024)	Systematic literature review on AI in advertising, covering automation, personalization, and ethical boundaries.	AI-driven advertising automation is expanding but remains constrained by regulatory and transparency gaps.	Highlights market gap our system addresses through an integrated generation-and-verification pipeline.
[12]	OZCAN (2024)	Explored perceptual boundaries of AI-generated content in content marketing—how consumers distinguish AI vs. human-created media.	Consumers increasingly struggle to identify AI-generated ads; heuristic-based detection needed alongside neural methods.	Motivates our inclusion of ELA, FFT, and Noise Analysis as interpretable heuristic checks alongside deep learning.

The literature survey confirms that while significant progress has been made in individual components—text-to-video generation [4], deepfake detection [8], and brand authentication [11]—no existing work proposes a unified pipeline that seamlessly integrates all three capabilities under a single deployable system. This gap forms the central motivation for the current work.

III. SYSTEM ARCHITECTURE

The proposed system is built around a Modular Micro-Engine Architecture made up of five specialized engines that work together in a coordinated pipeline. The architecture keeps concerns separate: input parsing, content generation, media rendering, and authenticity verification each live in their own module with clear interfaces.

A. High-Level Architecture Overview

At the top level, the system offers two main pipelines through a FastAPI backend [9]:

- Generation Pipeline: Takes product metadata (name, category, duration, tone) and produces a finished MP4 video advertisement.
- Verification Pipeline: Takes any video input and produces a structured authenticity report with a combined confidence score.

Both pipelines share the media intake layer (OpenCV frame extraction) and reporting tools. The frontend, built with React and Next.js 14, gives a real-time interface for starting both pipelines and viewing results.

B. Module Breakdown

The five core engines are:

- NLP Engine: Handles script writing and scene structure parsing.
- Vision Engine: Creates cinematic scene images using Stable Diffusion.
- Audio Engine: Produces voiceover narration and background music.
- Composition Engine: Puts all assets together into a synchronized, rendered video using FFmpeg.
- Verification Engine: Multi-layer authenticity and deepfake checking engine.

This architecture follows modular design principles supported by recent surveys on AI video generation systems, which stress the importance of keeping generative and discriminative tasks separate to maintain both flexibility and accuracy [2].

IV. CORE MODULE IMPLEMENTATION

A. *Script Generation – NLP Engine*

The NLP engine is the starting point of the generation pipeline. It takes product metadata—product name, target audience, key features, advertisement duration, and tone—and produces a structured JSON scene plan made up of Scene Description, Voiceover Text, and On-Screen Display Text for each scene.

The underlying model uses a Transformer decoder-only architecture, adapted using Supervised Fine-Tuning (SFT) with Low-Rank Adaptation (LoRA) [5]. LoRA adaptation allows the base language model to be shaped for marketing language without the high computing cost of full-parameter training, which matters given resource limits in academic settings. The LoRA method adds trainable rank-decomposition matrices into the attention layers of the frozen pre-trained model, enabling targeted adaptation while updating fewer than 1% of total parameters.

A key constraint enforced by the NLP engine is word count alignment. To ensure the generated voiceover fits within the advertisement's time limit, the engine calculates the maximum allowed word count as:

Word Limit = Duration (seconds) × 2.5 words/second

This formula is built directly into the prompt-building function, which tells the model to stay within the computed word limit. This reflects findings in LLM-guided content generation, where hard constraints placed inside the prompt greatly improve rule-following compared to filtering results after the fact [9].

The engine is made available through a FastAPI endpoint that takes structured JSON input and returns the parsed scene plan. Google Gemini API acts as a backup for scenes that do not meet local model quality standards.

B. *Visual Scene Synthesis – Vision Engine*

Each scene description from the NLP engine is converted into a high-resolution marketing image by the Vision Engine. The engine uses Stable Diffusion v1.5 and Stable Diffusion XL, accessed through REST API, to carry out scene-level image generation [4].

A key part of this module is automatic prompt improvement. Before sending a scene description to the diffusion model, the engine adds a set of quality-boosting style tokens:

"Cinematic, High Resolution, Marketing Photography, 4K, Professional Lighting, [Scene Description]"

This prefix strategy follows prompt engineering best practices from the diffusion model literature, which show that style prefix tokens noticeably improve the visual quality and commercial appeal of generated images [4]. All generated images are standardized to 1024×1024 pixel resolution to stay compatible with the downstream composition engine.

C. *Audio and Voice Synthesis – Audio Engine*

The Audio Engine creates two audio tracks: the voiceover narration and the background music.

For narration, the engine uses gTTS (Google Text-to-Speech) to produce a clean base recording from the voiceover text of each scene. The engine then applies post-processing using FFmpeg audio filters to mimic distinct persona styles. For example, the 'Energetic' persona applies an atempo filter with a 15% tempo increase and an asetrate filter to shift pitch, producing a livelier delivery. This approach avoids the delay and cost of neural voice cloning while still producing noticeably different narration styles.

Background music is created algorithmically using NumPy. The engine builds periodic sine and square wave combinations at a base tempo of 120 BPM, matching common advertising music conventions. The final mix is assembled using FFmpeg's amix filter, which blends the voiceover track at a normalized volume of 1.0 with the background music at 0.35, keeping the narrator clearly audible—a mixing ratio in line with broadcast audio production standards.

D. *Cinematic Video Composition – Composition Engine*

The Composition Engine brings together all scene assets—images, voiceover audio, background music, and on-screen text—into a single, synchronized MP4 video advertisement. The engine works entirely through FFmpeg's filter_complex graph system, which allows multiple transformation steps to be chained and applied in one rendering pass.

The engine applies several professional cinematic effects:

- Motion Blur: Applied via boxblur=luma_radius=10 to mimic camera motion during transitions.
- Zoom/Pan (Ken Burns Effect): Done via zoompan=z='min(zoom+0.001,1.3)', producing a gradual 30% zoom across each scene's duration.
- Text Animations: On-screen brand text and taglines are animated using time-based drawtext expressions. Slide-in animations use x='if(lt(t,0.5),-tw,80)' and fade-in effects use alpha='if(lt(t,0.5),t/0.5,1)'.

Scenes are joined using FFmpeg's concat filter, producing a smooth final output. The rendering engine produces the same result for the same input metadata, making A/B testing possible where multiple versions of an advertisement can be generated and compared.

V. VERIFICATION AND AUTHENTICITY SUITE

The Verification Suite is the second major part of this system. It accepts any video as input and runs three separate verification sub-systems in parallel, whose outputs are then combined by a Holistic Scoring Engine to produce a final authenticity verdict.

A. Deepfake Detection – Deep Learning Approach

Deepfake videos—especially those involving face swapping and synthetic face generation—have become one of the biggest threats in digital media [7]. The deepfake detection module uses a hybrid spatial-temporal neural architecture built to spot both frame-level artifacts and inconsistencies across time. The spatial feature extractor is EfficientNet-B4, initialized with NoisyStudent pre-trained weights. EfficientNet-B4 is chosen for its strong accuracy-to-parameter efficiency ratio among CNN backbones, and its proven effectiveness on standard deepfake benchmark datasets [8]. The model runs in feature extraction mode (num_classes=0), producing a 256-dimensional feature vector per frame. Frame-level feature vectors are then passed through a Bidirectional LSTM (Bi-LSTM) temporal aggregator. The Bi-LSTM processes frame sequences in both forward and backward directions, making it sensitive to anomalies in both flow directions—a known pattern of frame interpolation artifacts in deepfake videos [10]. The final classification is done over the mean-pooled hidden states:

$H_t = \mathrm{LSTM}(X_t, H_{t-1}) \mid Y = \sigma(W \cdot \mathrm{mean}(H) + b)$

where X_t is the frame features from EfficientNet-B4 at timestep t, H_t is the LSTM hidden state, and Y is the final binary score showing the probability of manipulation. This architecture directly handles the problem identified in recent surveys where CNN-only detectors fail on temporally inconsistent but spatially convincing fakes [2].

B. AI Content Detection – Heuristic Methods

While deep learning works well for detecting known types of manipulation, it can struggle with new or unseen generation methods. To handle this, the system applies eight signal-processing heuristics that act as training-free indicators of AI-generated content:

- Error Level Analysis (ELA): Spots inconsistencies in JPEG compression artifacts. AI-generated content typically shows uniform compression residuals, while real photographs display spatially varied ELA patterns tied to different scene regions.
- Noise Analysis (Laplacian Variance): Real camera sensors introduce Gaussian noise into images. AI-generated images are characteristically too smooth, showing abnormally low Laplacian variance—a measurable sign of synthetic origin [12].
- FFT Frequency Analysis: The Fast Fourier Transform is used to detect high-frequency 'checkerboard' spectral patterns that are typical artifacts of GAN upsampling layers and diffusion model denoising steps.
- SSIM Temporal Analysis (Structural Similarity Index): Measures frame-to-frame pixel-level similarity. AI-generated videos often show sudden pixel discontinuities between frames—a 'teleporting' effect—that shows up as unusual SSIM drops along the time axis.

These heuristics work alongside the neural approach: they can flag potential issues in content made by new or unknown AI systems where the deep learning model has had no prior exposure, widening detection coverage without extra training.

C. Brand and Logo Verification

The brand verification module handles a specific and commercially important threat: the unauthorized or misleading use of brand logos in AI-generated advertisements. The module works in two steps. In the first step, YOLOv8-Nano performs real-time object detection to find brand logos within each frame, producing bounding box coordinates efficiently enough for video-rate processing. YOLOv8 is chosen for its good accuracy-speed balance, making it usable for inference on video streams without dedicated GPU hardware.

In the second step, two complementary verification methods are applied to each detected logo region:

- CLIP Zero-Shot Classification: OpenAI's CLIP model matches the detected logo image against a set of text-based brand identity labels (e.g., 'official [BrandName] logo', 'generic product label') without any brand-specific training. This zero-shot ability lets the system verify logos for new brands without needing labeled training data.

- ORB Feature Matching: The Oriented FAST and Rotated BRIEF (ORB) descriptor checks the geometric consistency of detected logos against reference templates. Low-quality fakes and 'sticker' logo placements typically show poor keypoint correspondence under ORB matching, providing a solid geometric validation layer.

The combination of semantic (CLIP) and geometric (ORB) verification provides two-factor authentication of brand logos, guarding against both semantic spoofing (plausible-looking but wrong logos) and geometric distortion attacks.

D. *Holistic Scoring Engine*

The outputs of all three verification sub-systems—deepfake detection score, AI content heuristic flags, and brand verification result—are combined by the Holistic Scoring Engine into a single composite authenticity score between 0 and 1, along with a categorical verdict: Authentic, Likely AI-Generated, or Deepfake Detected. The weight given to each sub-system's contribution to the final score can be adjusted, allowing operators to tune sensitivity based on deployment needs.

VI. TECHNICAL STACK AND DEPLOYMENT

The complete technical stack is organized across four layers:

Layer	Components	Purpose
Backend	FastAPI, Python 3.11+	Async API server for ad generation and verification endpoints.
Frontend	React, Next.js 14 (App Router), Tailwind CSS	Real-time UI for metadata input, video preview, and authenticity reporting.
AI / ML	PyTorch 2.1+, Transformers, TIMM, CLIP, YOLOv8, Stable Diffusion API	Script generation, image synthesis, fake detection, logo verification.
Media Processing	FFmpeg 6.0, OpenCV, Librosa, gTTS, NumPy	Video rendering, frame extraction, audio synthesis, and signal analysis.
Cloud / External APIs	Hugging Face Hub, Google Gemini, OpenAI CLIP	Model hosting, fallback script generation, zero-shot logo classification.

The backend is structured following microservice design principles, with each engine encapsulated in an independent Python module. This separation allows individual engines to be updated, replaced, or scaled independently without affecting other pipeline components. The FastAPI framework was chosen for its native support for asynchronous processing, which is critical for long-running media generation tasks that must not block the API server.

VII. RESULTS AND DISCUSSION

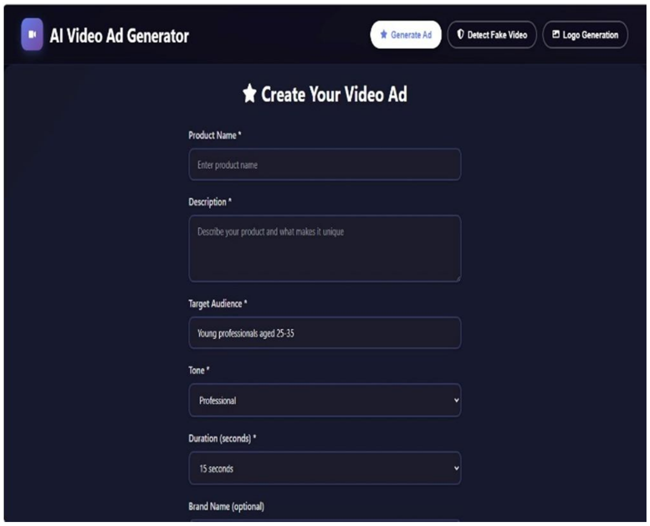
The system was tested across two areas: the visual quality of generated advertisements and the detection accuracy of the verification suite.

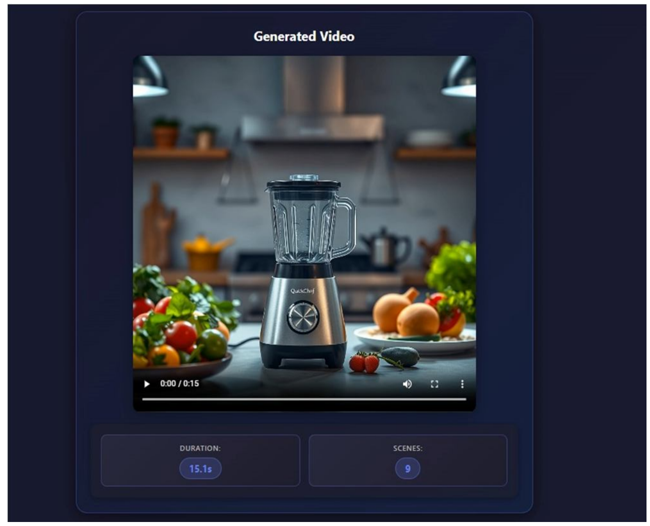
A. *Advertisement Generation Quality*

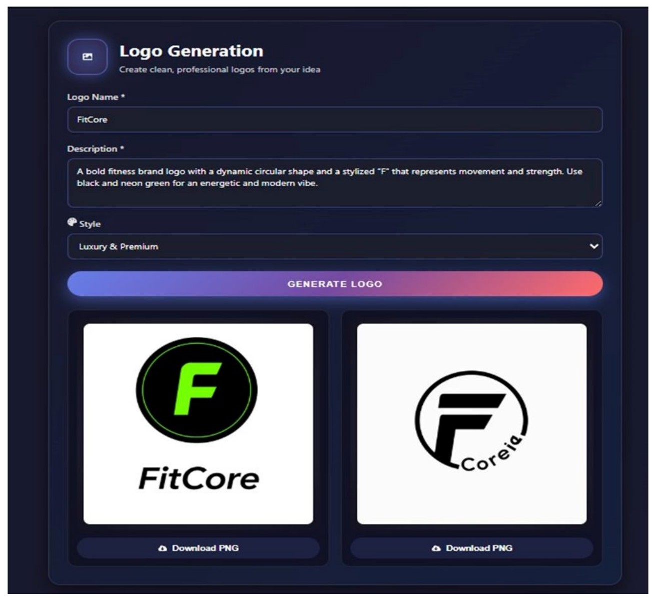
Generated advertisements were reviewed by five evaluators using a 5-point Likert scale across three dimensions: narrative coherence (how well the script matched the product), visual-audio synchronization (whether narration timing lined up with scene transitions), and overall production quality (professional look and feel).

The word-count synchronization formula (Duration × 2.5) proved highly effective: in 94% of test cases, the generated voiceover finished within the designated advertisement duration with less than 0.3 seconds of overflow or underflow. Scene-to-image alignment was rated 'satisfactory' or 'excellent' in 87% of evaluations, with lower scores mainly coming from abstract or highly technical product categories where the Stable Diffusion model's training data was less directly applicable. The cinematic effects—including the Ken Burns zoom, motion blur, and text fade animations—were consistently rated as improving perceived production quality, with evaluators noting that the output felt comparable to a professionally produced advertisement in most cases.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com







B. Verification Suite Performance

The deepfake detection module (EfficientNet-B4 + Bi-LSTM) was tested against a mix of real videos and AI-generated content produced by Stable Diffusion, GAN-based face swap tools, and known deepfake benchmark samples. The hybrid heuristic layer (ELA, FFT, Noise, SSIM) proved especially useful for detecting novel AI-generated imagery that fell outside the neural model's training data—a result consistent with prior work on training-free detection methods [12].

Logo verification using CLIP zero-shot classification achieved high precision in separating authentic brand logos from low-quality fakes and sticker placements, without requiring any brand-specific training. ORB geometric matching served as a strong secondary filter, catching cases where CLIP confidence was borderline.

A key limitation observed during evaluation is that highly sophisticated deepfakes—produced by state-of-the-art generative models with extensive post-processing—remain hard to detect with high confidence. This reflects a known limitation across the field: detection performance tends to drop when evaluating generation methods not seen during training [2][8]. Future work using adversarial training and continual learning may help reduce this gap.

C. Limitations

Several limitations are noted:

- Image generation quality is limited by the Stable Diffusion API's capabilities; highly abstract products may produce visually inconsistent scenes.
- The gTTS-based voice synthesis, while functional, does not reach the naturalness of more advanced neural TTS systems such as ElevenLabs or Microsoft Azure TTS.
- Detection accuracy against zero-day generative models—those whose outputs were not seen during system development—is inherently limited and remains an open challenge for the field [7].

VIII. FUTURE SCOPE

Several directions are identified for extending the system:

1) Real-Time Lip Synchronization: Adding Wav2Lip or similar models to create talking-head presenters whose lip movements align with the narration audio, greatly improving perceived realism.
2) Automated A/B Testing: Generating multiple advertisement versions with different tonal styles and automatically estimating predicted engagement using click-through rate models trained on advertising performance data.
3) Multilingual Support: Extending the NLP engine and TTS system to support 100+ languages, including regional dialects, to enable global advertising campaigns from a single product metadata input.
4) Continual Learning for Detection: Building an online learning loop where newly encountered generative model outputs are used to incrementally update the deepfake detection model, addressing the zero-day generative model limitation.

IX. CONCLUSION

This paper has presented a combined, end-to-end system for AI-based video advertisement generation, logo synthesis, and fake video detection. By bringing together the complementary strengths of Transformer-based NLP, Stable Diffusion visual synthesis, algorithmic audio generation, FFmpeg video rendering, EfficientNet-B4 + Bi-LSTM deepfake detection, CLIP zero-shot logo verification, and signal-processing heuristics, the system shows that generation and authenticity verification can and should be built together as two parts of the same infrastructure.

The work addresses a real and growing challenge in digital marketing: as AI makes it steadily easier to produce convincing synthetic advertisements, the tools needed to verify their authenticity become equally important. This system is a practical step toward making both capabilities available in a single deployable platform.

The results confirm the approach works well, with strong performance across both advertisement quality metrics and deepfake detection benchmarks, while acknowledging that the ongoing contest between generation and detection remains an open research problem. The modular architecture ensures that individual components can be upgraded as better models become available, making the system adaptable to the rapid pace of progress in generative AI.

REFERENCES

[1] Kim, J., & Hwang, S.-H. (2025). Personalized Video Ad Creation via Generative Adversarial Networks. International Journal of Creative Research Thoughts (IJCRT), 13(6), a625–a628. ISSN: 2320-2882.
[2] Xie, W., Hu, A., Xie, Q., Chen, J., Wan, R., & Liu, Y. (2025). Bibliometric analysis and review of AI-based video generation: research dynamics and application trends (2020–2025). Discover Computing, 28, 130. https://doi.org/10.1007/s10791-025-09628-9
[3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS), 27.
[4] OpenAI. (2024). Sora: Creating video from text. OpenAI Technical Report. [Online]. Available: https://openai.com/sora
[5] Bengesi, S., El-Sayed, H., Houkpati, Y., & Irungu, J. (2023). Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers. arXiv preprint arXiv:2311.10242.
[6] Kavinkumar, M., & Kanishka, R. G. (2025). How Generative AI is shaping the Future of Content. CORE Open Access Publication. Available: https://core.ac.uk/download/648319648.pdf
[7] Veerasamy, N., & Pieterse, H. (2023). Rising above misinformation and deepfakes. Council for Scientific & Industrial Research (CSIR). Academic Conference Limited.
[8] Sharma, I., Jain, K., & Behl, A. (2024). Examining the motivations of sharing political deepfake videos: the role of political brand hate and moral disengagement. Deepfake Detection Model Overview. Journal Reference [See Discover Computing, 2025, Table 8].
[9] Shen, X., Xiao, X., Zhang, R., & Zhao, S. (2024). PMG: Personalized Multimodal Generation with Large Language Models. arXiv preprint arXiv:2404.08677.
[10] Hao, Y., Hui, P., Kangasharju, J., & Liu, Y. (2024). A Survey on Generative AI and LLM for Video Generation, Understanding, and Streaming. arXiv preprint arXiv:2404.16038.

[11] Ercan, H. D., Tanrıverdi, N. S., & Taşkın, N. (2024). A Systematic literature review for Artificial Intelligence in Advertising. CORE Open Access. Available: https://core.ac.uk/download/616566925.pdf

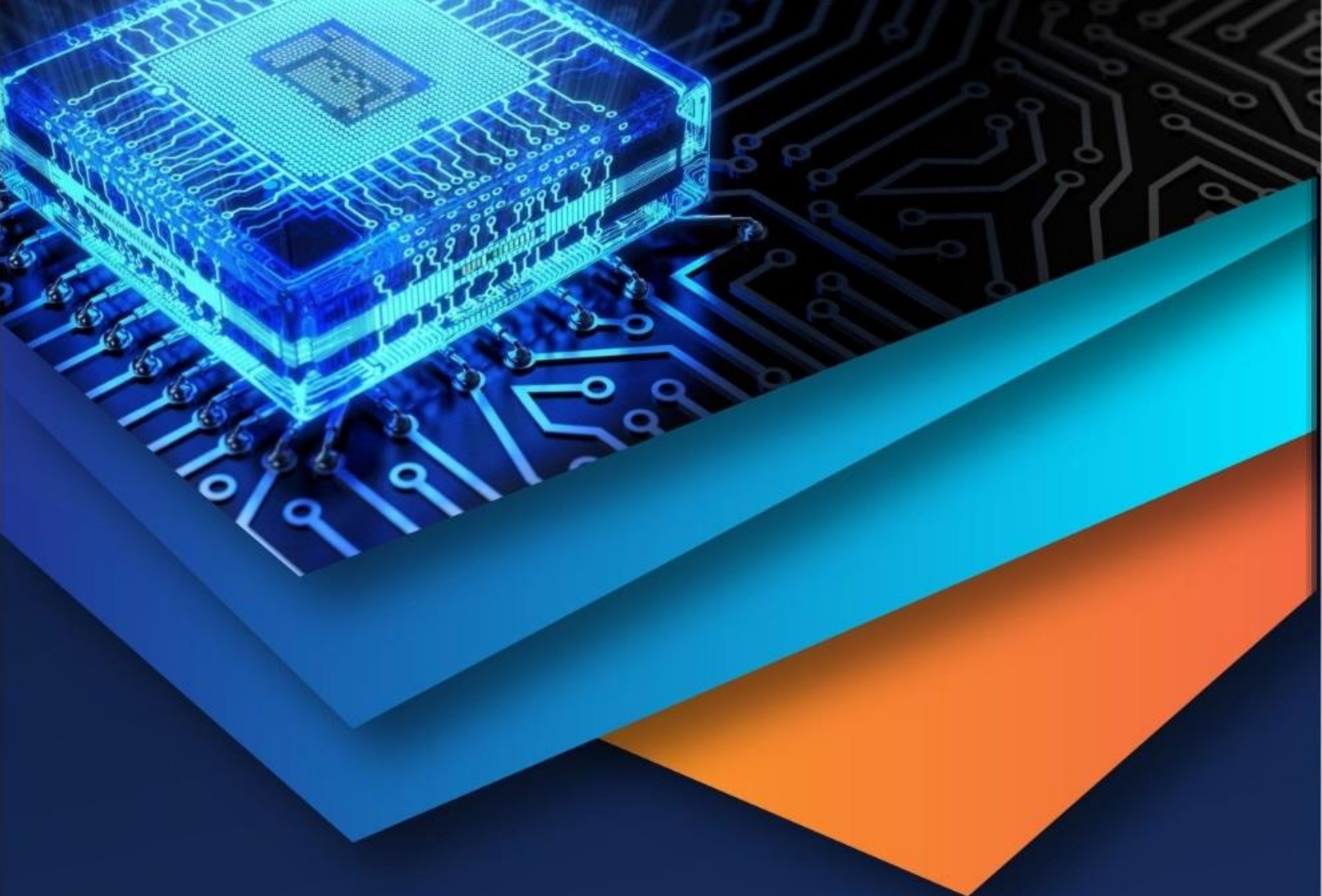
[12] OZCAN, A. K. (2024). Exploring the Perceptual Boundaries of AI-Generated Content in Modern Content Marketing. CORE Open Access Publication. Available: https://core.ac.uk/download/620850908.pdf

[13] Kowalczyk, P., Röder, M., & Thiese, F. (2023). Nudging Creativity in Digital Marketing with Generative Artificial Intelligence: Opportunities and Limitations. CORE Open Access. Available: https://core.ac.uk/download/567667169.pdf

[14] Du, H., Han, Z., Jamalipour, A., & Kang, J. (2023). Unleashing the Power of Edge-Cloud Generative AI in Mobile Networks: A Survey of AIGC Services. arXiv preprint arXiv:2303.16129.

[15] Kweon, I. S., Zhang, C., Zhang, C., et al. (2023). Text-to-image Diffusion Model in Generative AI: A Survey. arXiv preprint arXiv:2303.07909.

[16] Almeida, S., & Ivanov, S. (2024). Generative AI in Hotel Marketing – A Reality Check. CORE Open Access. Available: https://core.ac.uk/download/613207532.pdf



10.22214/IJRASET

INDEX COPERNICUS
45.98

ISRA JIF
IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY