

The Hero in Four Worlds: AI-Powered Visual Storytelling with Consistent Character Design

Alexandru-Bogdan Șerban

Master's Program: Software Engineering

Politehnica University Timisoara

Course: Graphics Processing Systems

Supervisor: As.Drd.Ing. Bianca Gușă

January 2026

Abstract—This project explores the creative potential of modern AI image generation tools through the theme “The Hero in Four Worlds” — placing a consistent character across four dramatically different environments: a tropical jungle, a medieval town, a futuristic city, and a desert landscape. Using a combination of Seedream 4.5 [1], Nano Banana [3], and Flux Kontext [4], we demonstrate a practical workflow for generating visually cohesive narratives while maintaining character consistency. The project evaluates each tool’s strengths and weaknesses in terms of image quality, generation speed, creative flexibility, and consistency preservation. Results show that combining multiple specialized AI tools produces superior outcomes compared to using a single tool, with each platform excelling at different stages of the creative pipeline.

Index Terms—AI image generation, character consistency, visual storytelling, Seedream, Nano Banana, Flux Kontext, generative AI

I. INTRODUCTION

A. Project Theme

This project explores the theme “**The Hero in Four Worlds**” — creating a consistent character placed across four dramatically different environments:

- 1) **Jungle** — Dense tropical vegetation, wildlife, natural lighting
- 2) **Medieval Town** — Historical architecture, cobblestone streets, period-appropriate atmosphere
- 3) **Futuristic City** — Neon lights, advanced technology, cyberpunk aesthetics
- 4) **Desert** — Arid landscape, harsh sunlight, sand dunes and oasis

B. Purpose of the Project

This theme was chosen to test the capabilities of modern AI image generation tools in maintaining character consistency across vastly different environmental contexts. The challenge lies in placing the same hero character in four visually distinct worlds while preserving their identity, clothing, and features. This project demonstrates the practical workflow of combining multiple AI tools for creative visual storytelling.

C. Objectives

The primary objectives of this project are:

- Generate 4 distinct world environment landscapes using AI tools
- Generate a consistent hero character design using AI image generation tools
- Place the character in 4 distinct world environments (7-10 total images)
- Evaluate and compare the performance of Seedream 4.5, Nano Banana, and Flux Kontext
- Document the creative workflow and analyze tool strengths/weaknesses
- Create a compelling visual narrative across all four worlds

II. TOOLS OVERVIEW

This section describes the AI tools used in this project, their capabilities, and their specific roles in the creative workflow.

A. Seedream 4.5 (ByteDance)

Official Website: https://seed.bytedance.com/en/seedream4_5

Access Platform: <https://lovart.ai> (Seedream 4.5 integration) [2]

1) Description: Seedream 4.5 is ByteDance’s advanced AI image generation model, known for producing high-resolution, photorealistic images with excellent detail and composition. It excels at generating complex scenes with multiple elements, dramatic lighting, and cinematic compositions. The model responds well to structured prompts following the formula: Subject + Environment + Style + Technical Specs.

2) Role in Project: Used as the primary tool for generating all four world landscape backgrounds (Jungle, Desert, Medieval Town, Futuristic City). Seedream 4.5 was chosen for its ability to create detailed, high-resolution environments in 16:9 landscape format suitable for desktop wallpaper quality.

B. Nano Banana

Website: <https://gemini.google.com/> (Google Gemini)

1) *Description*: Nano Banana is an AI-powered image editing platform specialized in maintaining character and object consistency across multiple images. It allows users to upload reference images and seamlessly insert consistent elements into new scenes.

2) *Role in Project*: Used for inserting the hero character into all four world landscapes while maintaining consistent appearance. Also tested for in-context editing comparisons alongside Flux-2 Pro.

C. Flux Kontext

Website: <https://flux-context.org/>

1) *Description*: Flux Kontext is a context-aware AI image editing tool that excels at making local, targeted edits to images while preserving the overall composition. It is particularly useful for fine adjustments, background modifications, and adding/removing specific details.

2) *Role in Project*: Used for in-context editing experiments including holographic effects, pose changes, and adding story elements. Demonstrated excellent quality preservation compared to Nano Banana.

D. Additional Tools

lovart.ai — Web platform providing access to Seedream 4.5 with a user-friendly interface. Used as the primary access point for Seedream due to its intuitive controls and aspect ratio settings.

III. METHODOLOGY

A. Workflow Overview

The project follows a three-phase approach: first generating all world landscapes, then creating a consistent hero character, and finally compositing the character into each world with refinements.

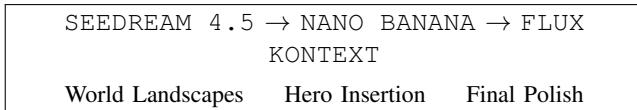


Fig. 1: Three-phase workflow diagram

Phase 1: Generate all four world landscapes with designated “hero spots” — areas where the character will be placed later (stone pedestal in jungle, empty camel saddle in desert, etc.)

Phase 2: Create the hero character reference and use Nano Banana to insert them consistently into each landscape.

Phase 3: Use Flux Kontext for final adjustments, lighting fixes, and detail enhancement.

B. Order of Tools Used

Table I shows the sequential order in which tools were applied throughout the project.

TABLE I: Tool Usage Order

Step	Tool	Purpose
1	Seedream 4.5	Generate 4 world landscape backgrounds
2	Seedream 4.5	Create hero character reference image
3	Nano Banana	Insert hero with consistency across scenes
4	Flux Kontext	Final polish and detail adjustments

C. Prompts Used

This section documents the prompts used for each generation, following the Seedream 4.5 best practices: Subject + Environment + Style + Technical Specs.

1) *Hero Character Reference Prompt*: The hero character was designed as a consistent reference sheet showing multiple angles for use with Nano Banana’s character insertion capabilities. Table II summarizes the distinctive character traits.

TABLE II: Hero Character Traits

Attribute	Description
Inspiration	Timothee Chalamet
Age	20-25 years old
Build	Athletic lean, approx. 180cm
Face	Sharp jawline, high cheekbones
Eyes	Deep-set, hazel-green, expressive
Eyebrows	Thick, dark, natural arch
Nose	Straight, refined
Hair	Dark brown-black, slicked back
Facial Hair	Well-groomed short goatee
Skin	Medium tone, warm undertones
Tattoo	Visible on right forearm
Accessory	Black ring on left index finger
Expression	Intense, confident gaze

Character reference sheet of a handsome young man age 20-25, Timothee Chalamet inspired features , DISTINCTIVE FACE with sharp defined jawline, high prominent cheekbones, deep-set expressive hazel-green eyes, thick dark eyebrows with natural arch, straight refined nose, full lips, medium skin tone with subtle warm undertones, dark brown-black slightly longer hair slicked back, well-groomed short goatee, visible tattoo on his right forearm, athletic lean build not overly muscular approximately 180cm tall, black ring on the left hand's index finger, intense confident gaze, wearing simple neutral gray t-shirt and pants, showing FRONT VIEW, SIDE VIEW, THREE-QUARTER VIEW, and BACK VIEW arranged in a row, character turnaround sheet, full body poses, clean white background, consistent lighting across all angles, photorealistic, 4K resolution, highly detailed facial features for character consistency, professional character design reference sheet



Fig. 2: Hero character reference sheet showing front, side, 3/4, and back views for consistency. Generated with Seedream 4.5 via lovart.ai.

2) Jungle World Prompt:

Dense tropical jungle, lush green canopy with towering ancient trees, golden sunbeams piercing through the foliage creating god rays, volumetric lighting, a prominent mossy stone plateau in the center of a clearing bathed in sunlight, a majestic lion sitting beside the stone looking upward toward the top of the stone with curious attention, colorful parrots perched on branches, monkeys in the trees, a toucan with vibrant beak, butterflies floating in the light beams, misty atmosphere, rich vegetation with ferns and exotic flowers, wide panoramic view, landscape orientation, cinematic widescreen composition, photorealistic, 16:9 aspect ratio, 4K resolution, highly detailed

3) Desert World Prompt:

Vast golden desert landscape at warm afternoon light, endless rolling sand dunes stretching to the horizon, gentle wind lifting wisps of sand into the air, delicate sand particles drifting across the scene, a shimmering oasis with palm trees visible in the far distance, a caravan of camels with riders traveling in the background slightly out of focus, in the foreground center a single majestic camel with ornate saddle and decorative blankets standing ready with no rider, the camel's mane and saddle fabric gently blowing in the breeze, soft desert haze, heat shimmer effect, dramatic shadows on dunes, wide panoramic view, landscape orientation, cinematic composition, 16:9 aspect ratio, 4K resolution, photorealistic, National Geographic style

4) Medieval Town Prompt:

Epic medieval jousting tournament at golden hour, grand arena with colorful heraldic banners, in the foreground center a magnificent white warhorse with ornate saddle FACING DIRECTLY TOWARD THE CAMERA with NO RIDER ready for its champion, the horse standing proud and battle-ready, behind it at the far end of the arena an opponent knight in dark armor mounted on a black horse FACING AWAY with back toward camera preparing to charge, a beautiful princess in elegant gown watching from an ornate royal balcony waving a silk favor, crowds cheering from wooden stands, king and nobles in the royal box, dramatic tension before the joust, wide panoramic view, landscape orientation, photorealistic, 16:9 aspect ratio, 4K resolution, cinematic medieval fantasy

5) Futuristic City Prompt:

Breathtaking cyberpunk rooftop vista at night, standing at the edge of a rain-soaked skyscraper overlooking Night City, thousands of neon lights illuminating the urban canyon below, aerodynes and flying cars zooming between massive holographic billboards, cherry blossom petals drifting in the wind mixing with rain, empty rooftop ledge in foreground center, towering corporate arcologies piercing the smog layer, flickering advertisements, steam rising from vents, Cyberpunk 2077 and Blade Runner 2049 aesthetic, ultra-wide cinematic composition, landscape orientation, 16:9 aspect ratio, 4K resolution, photorealistic, dramatic mood lighting, neon noir atmosphere

D. Intermediate Results

Table III shows the current progress of landscape generation.

TABLE III: Landscape Generation Progress

World	Status	File	Hero Spot
Jungle	Complete	jungle_1.png	Stone plateau
Desert	Complete	desert_1.png	Empty camel saddle
Medieval	Complete	medieval_3.png	Riderless warhorse (facing camera)
Futuristic	Complete	cyberpunk_1.png	Rooftop edge

IV. RESULTS

A. Final Images



Fig. 3: Jungle world landscape with stone plateau and lion. Generated with Seedream 4.5 via lovart.ai.

1) Jungle Landscape: World: Jungle

Tool(s) Used: Seedream 4.5 (lovart.ai)

Description: Dense tropical jungle scene featuring golden sunbeams piercing through the canopy, creating dramatic god rays. A mossy stone plateau in the center serves as the designated hero spot. A lion sits beside the stone, gazing upward, creating a natural focal point for where the hero will stand.



Fig. 4: Desert world landscape with camel and distant oasis. Generated with Seedream 4.5 via lovart.ai.

2) Desert Landscape: World: Desert

Tool(s) Used: Seedream 4.5 (lovart.ai)

Description: Vast golden desert with rolling sand dunes, an oasis visible in the distance, and a caravan in the background. A single camel with ornate saddle stands ready in the foreground — the empty saddle serves as the hero spot where the character will be placed.



Fig. 5: Medieval jousting tournament with riderless warhorse facing camera, opponent knight with back to viewer, and princess on balcony. Generated with Seedream 4.5 via lovart.ai.

3) Medieval Town: World: Medieval Town

Tool(s) Used: Seedream 4.5 (lovart.ai)

Description: Epic jousting tournament with reversed composition — the hero’s white warhorse faces directly toward the camera while the opponent knight has his back to the viewer. This setup allows for front-facing hero insertion.



Fig. 6: Cyberpunk rooftop overlooking neon-lit megacity with flying vehicles. Generated with Seedream 4.5 via lovart.ai.

4) Futuristic City: World: Futuristic City

Tool(s) Used: Seedream 4.5 (lovart.ai)

Description: Breathtaking cyberpunk rooftop vista at night overlooking a neon-lit megacity. Flying vehicles zoom between holographic billboards while rain and cherry blossoms drift through the air. The empty rooftop ledge in the foreground serves as the hero spot where the character will stand surveying the city.

B. Hero Composites

This section presents the final composited images where the hero character has been inserted into each landscape using Nano Banana.

1) Desert — Hero Composite: See Figure 7.

Outfit: Traditional white desert robes with headscarf (keffiyeh style)

Pose: Sitting on camel saddle, looking at camera

Tool Used: Nano Banana

Observation: Slight quality reduction observed during compositing compared to original Seedream 4.5 output.



Fig. 7: Hero character inserted into desert scene, wearing white desert robes and riding the camel. Composited using Nano Banana.

2) Jungle — Hero Composite: See Figure 8.

Outfit: Mowgli/Tarzan jungle style — minimal clothing, bare chest, loincloth

Pose: Standing heroically on stone plateau, commanding presence

Tool Used: Nano Banana

Note: May require Flux Kontext refinement for seamless integration.



Fig. 8: Hero character as king of the jungle, standing on stone plateau in Mowgli/Tarzan style. Composited using Nano Banana.

3) Medieval — Hero Composite: See Figure 9.

Outfit: Silver knight armor, no helmet (face and hair visible)

Pose: Mounted on warhorse facing camera, holding jousting lance, fierce expression

Tool Used: Nano Banana



Fig. 9: Hero as knight mounted on warhorse facing camera, silver armor without helmet, ready for joust. Composed using Nano Banana.

4) Cyberpunk — Hero Composite: See Figure 10.

Outfit: Sleek black futuristic jacket with tech details, dark pants, combat boots

Pose: Standing at rooftop edge, looking over shoulder or gazing at city

Tool Used: Nano Banana

Note: Enhanced prompt used to maintain photorealistic quality and avoid stylized/drawn appearance.



Fig. 10: Hero in futuristic tech outfit standing at rooftop edge overlooking neon-lit city. Composed using Nano Banana.

C. In-Context Editing Comparison

This section compares in-context editing capabilities between Flux-2 Pro and Nano Banana using the same prompt on the cyberpunk scene.

I) Edit Prompt Used:

CRITICAL: DO NOT MODIFY THE CHARACTER'S APPEARANCE, FACE, BODY, CLOTHING, OR POSITION. Keep all character features EXACTLY as they are.
DO NOT change the background or city elements.

ONLY ADD: A holographic lion projection from the character's ring.

Character's left hand raised in front of him, palm facing up.

From the BLACK RING on his left index finger, a glowing holographic LION is projected - a majestic blue-cyan hologram of a lion resembling the one from the jungle, made of translucent glowing light particles and geometric wireframe lines. The hologram lion is visible and prominent,

floating above the ring, looking powerful and alert.

The hologram casts soft blue light on character's face and hand.

Futuristic AR tech visualization, cyberpunk aesthetic.

ONLY ADD the holographic lion element.
Photorealistic integration.



Fig. 11: Cyberpunk scene with holographic lion edit using Flux-2 Pro.

2) Flux-2 Pro Result: **Observations:** Resolution maintained at high quality. Hologram placement less intuitive but output remains crisp and detailed.



Fig. 12: Cyberpunk scene with holographic lion edit using Nano Banana.

3) Nano Banana Result: **Observations:** Hologram positioning more logical and natural. However, some resolution/quality reduction observed compared to Flux-2 Pro output.



Fig. 13: Medieval scene with rearing horse edit using Flux-2 Pro.

4) Medieval Rearing Horse — Flux-2 Pro: Observations:
Very crisp output with excellent lighting. Details remain sharp and on point. Overall superior quality preservation.



Fig. 14: Medieval scene with rearing horse edit using Nano Banana.

5) Medieval Rearing Horse — Nano Banana: Observations:
Slight cartoon-like stylization introduced. Quality dropped compared to Flux-2 Pro. Edit interpretation reasonable but less photorealistic.



Fig. 15: Jungle scene with mystical ring origin edit using Flux-2 Pro.

6) Jungle Ring Origin — Flux-2 Pro: Observations:
More creative interpretation with additional creatures and excellent lighting effects. However, character has an extra arm artifact. With polishing, could work well. Overall more imaginative but less accurate.



Fig. 16: Jungle scene with mystical ring origin edit using Nano Banana.

7) Jungle Ring Origin — Nano Banana: Observations:
Character position preserved correctly. However, cartoonish stylization applied. Effects do not integrate well with the original photorealistic design. Less creative but more stable character preservation.



Fig. 17: Desert scene with ghost lion companion edit using Flux-2 Pro.

8) Desert Ghost Lion — Flux-2 Pro: Observations:
Crisp resolution maintained. However, character appearance significantly altered — no longer looks like the original hero. Quality preserved but identity lost.



Fig. 18: Desert scene with ghost lion companion edit using Nano Banana.

9) Desert Ghost Lion — Nano Banana: Observations:
Slightly dimmed quality compared to original. Character also altered and no longer resembles the original hero. Both tools struggle with character consistency during complex edits.

TABLE IV: In-Context Editing Comparison

Criteria	Flux-2 Pro	Nano Banana
Resolution Quality	Excellent	Reduced
Lighting/Details	Crisp, sharp	Slightly degraded
Style Preservation	Photorealistic	Slight cartoonization
Edit Positioning	Less intuitive	More logical
Character Consistency	Variable (may alter)	Variable (may alter)
Overall Winner	Recommended	Good alternative

10) In-Context Editing Summary:

D. Visual Story Summary

The hero's journey begins in the primordial jungle, where atop a sunlit stone plateau, he receives a mystical ring from ancient spirits while a lion watches with knowing eyes. Empowered by this artifact, he travels to a medieval kingdom where he becomes a knight champion, facing opponents in jousting tournaments with fierce determination. His adventures continue through the scorching desert, riding camelback with the spirit of his lion companion manifesting from the ring as a ghostly guide. Finally, in the distant future, he stands on a cyberpunk rooftop overlooking a neon-lit megacity, the lion now appearing as a holographic projection from his ring — a luminous memory of where his journey began.

V. ANALYSIS & COMPARISON

A. Tool Comparison Matrix

Table V presents a comparative analysis of the three AI tools used in this project.

TABLE V: Tool Comparison Matrix (Rating: 1-5)

Criteria	Seedream	Nano Banana	Flux
Quality	5	3	5
Speed	5	3	5
Consistency	5	3	4
Creativity	5	3	5
Ease of Use	4	3	4

B. Strengths and Weaknesses

Seedream 4.5:

- Excellent photorealistic quality for landscape generation
- Very consistent outputs matching prompt specifications
- Best tool for initial background/scene creation
- Fast generation times

Nano Banana:

- Good for character compositing and insertion tasks
- More logical and intuitive element positioning
- Quality and resolution reduction observed in outputs
- Slight cartoonization of photorealistic source images
- Often changes angles/positions, requiring multiple prompt retries

Flux-2 Pro:

- Excellent resolution and quality preservation
- Creative interpretation of in-context editing prompts
- Best overall in-context editing quality
- Sometimes alters character identity (extra limbs, facial changes)
- Edit positioning can be less intuitive than expected

C. Issues Encountered

- 1) **Resolution drop during Nano Banana compositing:** The first hero composite (desert scene) exhibited noticeably lower resolution compared to the original Seedream-generated background.
- 2) **Character identity changes in Flux-2 edits:** Desert and jungle in-context edits altered the hero's appearance significantly, including an extra arm artifact in the jungle scene and facial changes in the desert scene.
- 3) **Angle/position changes in Nano Banana:** The tool frequently modified camera angles and element positions despite explicit instructions to preserve the background exactly, requiring multiple prompt retries.
- 4) **Quality stylization in Nano Banana:** Photorealistic source images became slightly cartoonish after processing, losing some of the original photorealism.

Solutions Applied:

- Enhanced prompts with explicit quality and resolution preservation instructions
- Used Flux-2 Pro for quality-critical in-context edits
- Multiple iterations to achieve acceptable results

D. Character Consistency Analysis

Character consistency proved to be an interesting challenge in this project. While Seedream 4.5 generated the initial hero reference sheet with highly distinctive features (sharp jawline, hazel-green eyes, goatee, slicked-back hair, tattoo, black ring), maintaining these features varied based on edit complexity.

For simpler in-context edits, both tools performed well. The medieval rearing horse edit successfully preserved the hero's identity while changing the horse's pose — demonstrating that these tools can maintain character consistency when the requested changes are straightforward.

However, more complex edits involving additional elements (holographic lions, mystical effects, spirit companions) caused both Nano Banana and Flux-2 Pro to alter the character's appearance more significantly. The desert ghost lion edit resulted in noticeable identity changes in both tools.

Key Finding: Current AI in-context editing tools work well for simple pose/element changes but struggle with complex multi-element edits while preserving character identity. Simpler is better for consistency.

VI. CONCLUSIONS

A. Lessons Learned

- **Start with the best generation tool:** Seedream 4.5 produces the highest quality base images — invest time in getting the initial landscapes perfect before attempting compositing or editing.
- **Design “hero spots” intentionally:** Empty saddles, stone plateaus, and rooftop edges made character insertion much easier than trying to add characters to busy scenes. Plan for insertion during background generation.
- **Simpler edits = better consistency:** Complex in-context edits (adding multiple elements, special effects) often

alter character identity. Keep edits focused and simple for best results.

- **Use explicit constraints in prompts:** Phrases like “DO NOT MODIFY the character’s appearance” and “KEEP BACKGROUND EXACTLY AS IS” significantly improved results. The more specific and emphatic the constraints, the better the tools respected them.

B. Best Practices Identified

- Create detailed character reference sheets with multiple angles before starting any compositing work
- Use explicit “DO NOT MODIFY” instructions in prompts to protect background and character elements from unwanted changes
- Test complex edits in multiple tools (both Flux-2 and Nano Banana) and compare results before finalizing
- Build scenes with intentional insertion points (empty spaces, waiting mounts) for easier character compositing

C. Ideas for Extensions

- Add more world environments (underwater, space, post-apocalyptic, etc.)
- Create animation sequences transitioning between worlds
- Develop hybrid workflows combining multiple tools for optimal results
- Explore video generation tools for animated storytelling
- Test newer AI models as they release for improved character consistency

REFERENCES

- [1] ByteDance, “Seedream 4.5 - AI Image Generation,” 2025. [Online]. Available: https://seed.bytedance.com/en/seedream4_5
- [2] lovart.ai, “Seedream 4.5 and Flux Access Platform,” 2025. [Online]. Available: <https://lovart.ai>
- [3] Google, “Gemini - AI Image Editing (Nano Banana),” 2025. [Online]. Available: <https://gemini.google.com/>
- [4] Flux Kontext, “Context-Aware Image Editing,” 2025. [Online]. Available: <https://flux-context.org/>
- [5] “AI Image Generation Tutorial - Lovart, Seedream & Flux,” YouTube, 2025. [Online]. Available: https://www.youtube.com/watch?v=J_UfnhCCYfo