

# Visual Generative AI

Master AI Image & Video Generation: Prompting, Branding & Tools

---

A complete guide to text-to-image prompting, creating consistent brand assets and characters, basic text-to-video generation, and the essential visual AI tool stack.

A Study Guide for AI Learners

# Table of Contents

- 1. Text-to-Image Prompting Strategies (Styles, Lighting, Angles)
- 2. Creating Logos, Social Media Assets & Consistent Characters
- 3. Basic Video Generation (Text-to-Video)
- 4. Tool Stack: Midjourney, Leonardo.ai, RunwayML
- 5. Practice Exercises
- 6. Quick-Reference Cheat Sheet

# 1. Text-to-Image Prompting Strategies

Text-to-image models (Midjourney, DALL-E, Leonardo.ai, Stable Diffusion) turn written descriptions into images. The quality of the result depends heavily on how precisely you describe the subject, style, lighting, and composition.

## The Prompt Formula

A strong image prompt typically combines: **Subject + Style + Lighting + Angle/Composition + Extra details** (color palette, mood, medium).

"A cozy mountain cabin at sunset, watercolor illustration style, warm golden-hour lighting, wide-angle shot, muted earthy color palette, peaceful mood"

## Styles

- Photographic: 'DSLR photo, 85mm lens, shallow depth of field'
- Illustration: 'flat vector illustration', 'watercolor painting', 'anime style'
- 3D/Render: '3D render, octane render, unreal engine, hyperrealistic'
- Art movements: 'in the style of impressionism', 'art deco poster'

## Lighting

- 'Golden hour lighting' — warm, soft, cinematic
- 'Studio lighting' — clean, even, professional product look
- 'Dramatic rim lighting' — high contrast, moody
- 'Soft diffused light' — gentle shadows, natural feel

## Angles & Composition

- 'Close-up shot' / 'macro shot' — focus on detail
- 'Wide-angle shot' / 'aerial view' — show full scene or environment
- 'Low-angle shot' — makes subject look powerful or large
- 'Rule of thirds composition' — balanced, professional framing

Tip: Use negative prompts (in tools that support them) to exclude unwanted elements, e.g. 'no text, no watermark, no extra limbs'.

## 2. Creating Logos, Social Assets & Consistent Characters

### Logos

Describe the brand's industry, personality, and preferred style (minimalist, geometric, mascot-based). Ask for a transparent or plain background so it's easy to use later.

"Minimalist geometric logo for a coffee brand called 'Brew&Co', simple line-art icon of a coffee cup, black and white, flat vector style, plain white background"

### Social Media Assets

Specify the exact platform format (square post, story, banner) and keep branding elements (colors, fonts, logo) consistent by describing them the same way in every prompt.

### Consistent Characters

Getting the same character across multiple images is one of the hardest parts of AI image generation. Key techniques:

- Write a very detailed, fixed character description and reuse it word-for-word in every prompt.
- Use reference images/image-prompting features (available in Midjourney, Leonardo.ai) to lock in a face or style.
- Use 'character reference' or 'style reference' parameters where the tool supports them.
- Generate the character in a neutral pose first, then vary only the action/background in later prompts.

### 3. Basic Video Generation (Text-to-Video)

Text-to-video tools (RunwayML, Pika, Luma) generate short video clips from text or image prompts. They work best with simple, clear motion descriptions rather than complex multi-scene stories.

#### Tips for Better Video Prompts

- Describe one clear motion or action: 'camera slowly zooms in on a steaming cup of coffee'.
- Mention camera movement explicitly: pan, zoom, tracking shot, static shot.
- Keep scenes short and simple — current tools handle a few seconds of coherent motion best.
- Start from a still image (image-to-video) for more control over the starting composition.

"A paper airplane gliding slowly through a sunlit office, camera tracking alongside it, soft cinematic lighting, slow motion"

Text-to-video is still evolving fast — always check a tool's latest documentation for its current capabilities and limitations before relying on it for client work.

## 4. Tool Stack: Midjourney, Leonardo.ai, RunwayML

| Tool        | Type                           | Highlights                                                 | Best For                                                         |
|-------------|--------------------------------|------------------------------------------------------------|------------------------------------------------------------------|
| Midjourney  | Text-to-Image                  | Highly artistic, painterly quality, strong style control   | Concept art, illustrations, branding visuals                     |
| Leonardo.ai | Text-to-Image                  | Fine-tuned models, character consistency, easy asset reuse | Game assets, UI elements, consistent characters, product mockups |
| RunwayML    | Text-to-Video / Image-to-Video | AI video generation, editing tools, motion control         | Short video clips, video editing, creative experiments           |

Common workflow: design characters/assets in Midjourney or Leonardo.ai, then animate key shots in RunwayML for a short promotional video.

## 5. Practice Exercises

- 1. Write a full image prompt using the Subject + Style + Lighting + Angle formula for a product photo.
- 2. Design a logo prompt for a fictional brand of your choice.
- 3. Write a fixed character description you could reuse across 3 different scene prompts.
- 4. Write a simple text-to-video prompt describing one clear camera movement and action.
- 5. Compare results for the same prompt across Midjourney and Leonardo.ai (if available to you).

## 6. Quick-Reference Cheat Sheet

| Element         | Example Keywords                                                |
|-----------------|-----------------------------------------------------------------|
| Style           | watercolor, 3D render, flat vector, photographic, anime         |
| Lighting        | golden hour, studio lighting, dramatic rim light, soft diffused |
| Angle           | close-up, wide-angle, low-angle, aerial view, rule of thirds    |
| Consistency     | fixed character description, reference image, style reference   |
| Video Motion    | pan, zoom, tracking shot, slow motion, static shot              |
| Negative Prompt | no text, no watermark, no blur, no extra limbs                  |

---

Visual generative AI rewards specificity — the more precisely you describe what you see in your mind, the closer the output will match it.