

Real-World Products Using VLMs

Product	Use Case	Backed by
ChatGPT Vision	VQA, OCR, diagram understanding	OpenAI
Gemini Pro	Video understanding + logic	Google DeepMind
Claude 3	Charts, tables, multimodal dialogue	Anthropic
Perplexity AI	Visual search and info retrieval	VLM Hybrid (CLIP + BLIP)
Adobe Firefly	Prompt-based image generation	Adobe
Google Lens	Product & object recognition	ViT + CLIP
MidJourney	AI art and concept generation	Diffusion + vision guidance