

Table 1: CLIP application topics (18 students, 90 ideas)

Topic	# students	Main principle
Satellite / aerial images (disaster damage, land use, exploration)	10	Retrieval, zero-shot classification
Wildlife camera traps	8	Zero-shot classification
Industrial defect detection	7	Prompt-pair scoring
Fashion / e-commerce product search	6	Retrieval
Image-text mismatch detection	5	Image-text agreement
Search in photo collections	4	Retrieval
Medical image retrieval / pathology	4	Retrieval, zero-shot classification
Auto-tagging and categorization	3	Zero-shot classification
Text-to-image generation	3	Generation
Content moderation against rules	3	Prompt-pair scoring
Robots following language commands	3	Grounding
Museum / art collection search	3	Retrieval
Accessibility (alt text, blind users)	3	Grounding, retrieval
Food recognition	3	Zero-shot classification
Hateful memes	2	Linear probe
Plant / crop disease	2	Zero-shot classification
Real estate listing photos	2	Zero-shot classification

Single ideas: waste sorting, road signs, cat breeds, construction safety, font search, geolocation, AI-generated image detection, anime reverse search, moodboards, screenshot clustering, photo editing, recipe-to-video alignment, film mood curves, music search (CLAP), robot reward model, disinformation, prompt-sensitivity testing.

Table 2: CLIP applications grouped by algorithmic principle

Category	Principle	# ideas	Example topics
Text-to-image retrieval	Embed the collection once, rank items by cosine similarity to a text query	30	Satellite search, fashion, museums, medical, fonts
Zero-shot classification	One prompt per class, argmax over similarities	24	Camera traps, waste, food, breeds, room types
Prompt-pair scoring	Normal vs. anomalous prompts, softmax score, threshold flags items	12	Defects, moderation, photo guidelines
Image-text agreement	Similarity of a given image-caption pair; low score flags it	5	Listing mismatches, dataset cleaning
Grounding / localization	Region proposals or patch embeddings matched to text	5	Robot commands, photo editing, assistive tech
Generation	CLIP text encoder conditions an image generator	4	Art, flashcards, disinformation
Trained head on embeddings	Linear probe or fine-tuning on frozen CLIP features	4	Hateful memes, AI-image detection, geolocation
Clustering / exploration	Group embeddings to find structure	2	Moodboards, screenshot sorting
Temporal alignment	Similarity curves over video frames	2	Recipe chapters, film mood
Other	CLIP score as training signal; analysis of CLIP	2	Robot reward, prompt sensitivity