



Automating Creative AI for Accessible Image Editing



Group 8

1155228584 Lee Lai Kuen
1155230194 Law King Wa
1155226820 Ho Yuk Lun
1155228484 Wong Ho Yin
1155110585 Cheung Ka Pui
1155120483 Wu Xi
1155226151 Tsang Ka Chun

Background

01





Current Trends

wen_0930 > 美圖秀秀 31/10/25 Ⓛ

+ 最近很紅的Ai雪景照真的好厲害...
沒有會員也有五次免費生成的機會👑



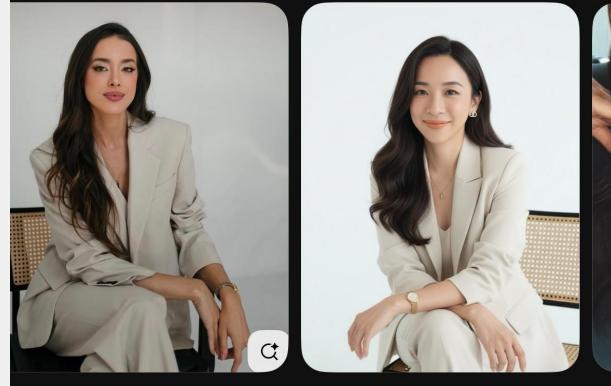
Background and Poses

ekisabelau 30/8/25 Ⓛ

用 Gemini generate LinkedIn portrait 爆 feed.

1. 不過我係 Pinterest 搵左 reference 相
2. 問 Gemini 出 detailed prompt
3. 再 upload 自己張相叫佢跟 prompt 做

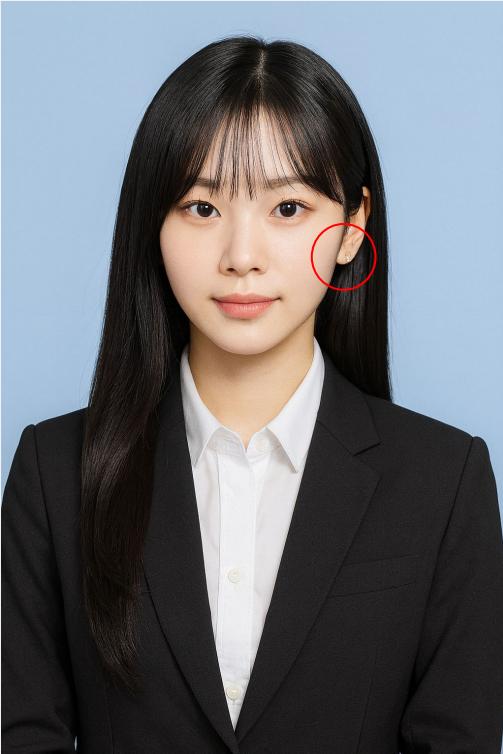
AI都真係太好用.
Prompt 係樓下. 1/5



Professional Photos



Problems of Current AI Tools





Business Value

Leverage AI-driven portrait processing to deliver secure, consistent, and identity-preserving photo generation that meets rising market demand across social, corporate, and verification use cases



Real-Life Constraints

Limitations in recapturing photo:
infants, unavailable,
deceased individuals

Growing User Demand

Driven by rapid
social-media spread and
viral AI trends

High Efficiency

Faster photoshoots,
reduced manual editing
time and quick output

Commercial Consistency

Standardized staff IDs
and corporate profile
photos



Limitations of Current AI Models

	Limitations	Our Approach
	<h3>Safety & Confidentiality Risks</h3> <ul style="list-style-type: none">• Free tools lack data protection• Images may be stored or reused• No confidentiality guarantee	<ul style="list-style-type: none">• Paid, secure processing environment
	<h3>Inconsistent Results</h3> <ul style="list-style-type: none">• Faces, clothing, and objects change unexpectedly• Misgeneration (wrong color, wrong pose, etc.)• Different tools produce different results	<ul style="list-style-type: none">• Label checking for content accuracy• Identity guard to preserve the person
	<h3>Poor User Experience</h3> <ul style="list-style-type: none">• Manual prompting, trial-and-error• Hard for users to describe edits clearly• No structured workflow	<ul style="list-style-type: none">• Auto-prompted editing• Easy & predictable results

02

Demo



03

Methodology & Results





Raw Data

Retrieved from APIs

01



Data Preprocessing

Preprocessing, categorizing, and labelling images

02



Pre-trained Models

Load detection, and face-embedding models

03



Prompt Generation

Automatically generating structured editing prompts

04



Image Generation

Applying AI models to produce edited images

05



Verification

Ensuring identity and content consistency after editing

06



Data Structure



Image Categories



Scenery

- Scenery (Urban, Nature)
- Plain and complex background
- Architecture / Buildings
- Indoor environments

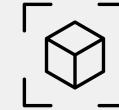
Scenery: 225



Primary Subject

- Human portraits (Gender, Age)
- Pets (Cats, Dogs)
- Ethnic diversity
- Individual vs Group
- Professional portrait, Korean style, ID photo

Human & Pets: 621



Object & Element

- Accessories (Glasses, Earrings)
- Objects (Cups, Phones)
- Furniture (Chairs, Tables)
- Clothing (Formal, Casual)

Object: 495

Image Condition: B&W, Low light, High contrast, Low resolution, Overexposed / Underexposed

Source: Pinterest, Pixabay, KaggleHub



Example

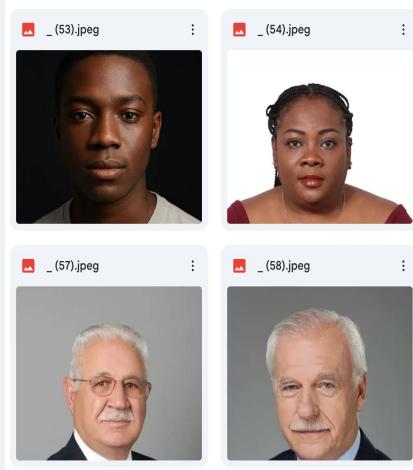
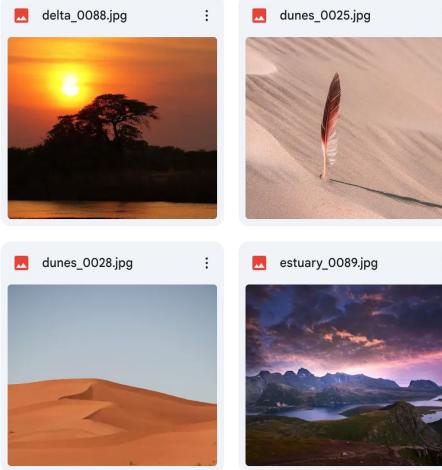
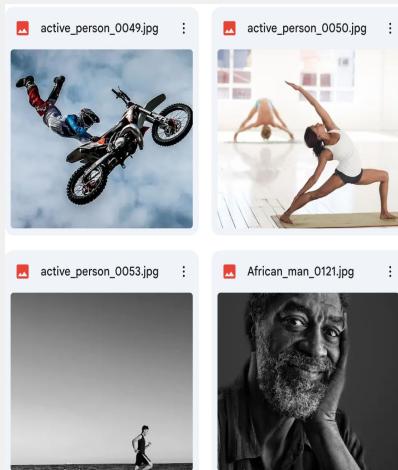
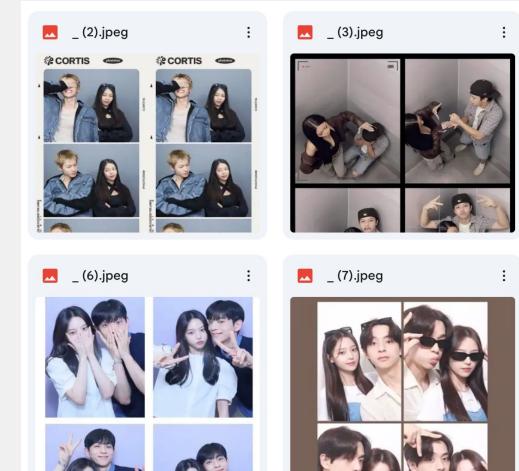


Image Condition



Scenery



Korean Selfie

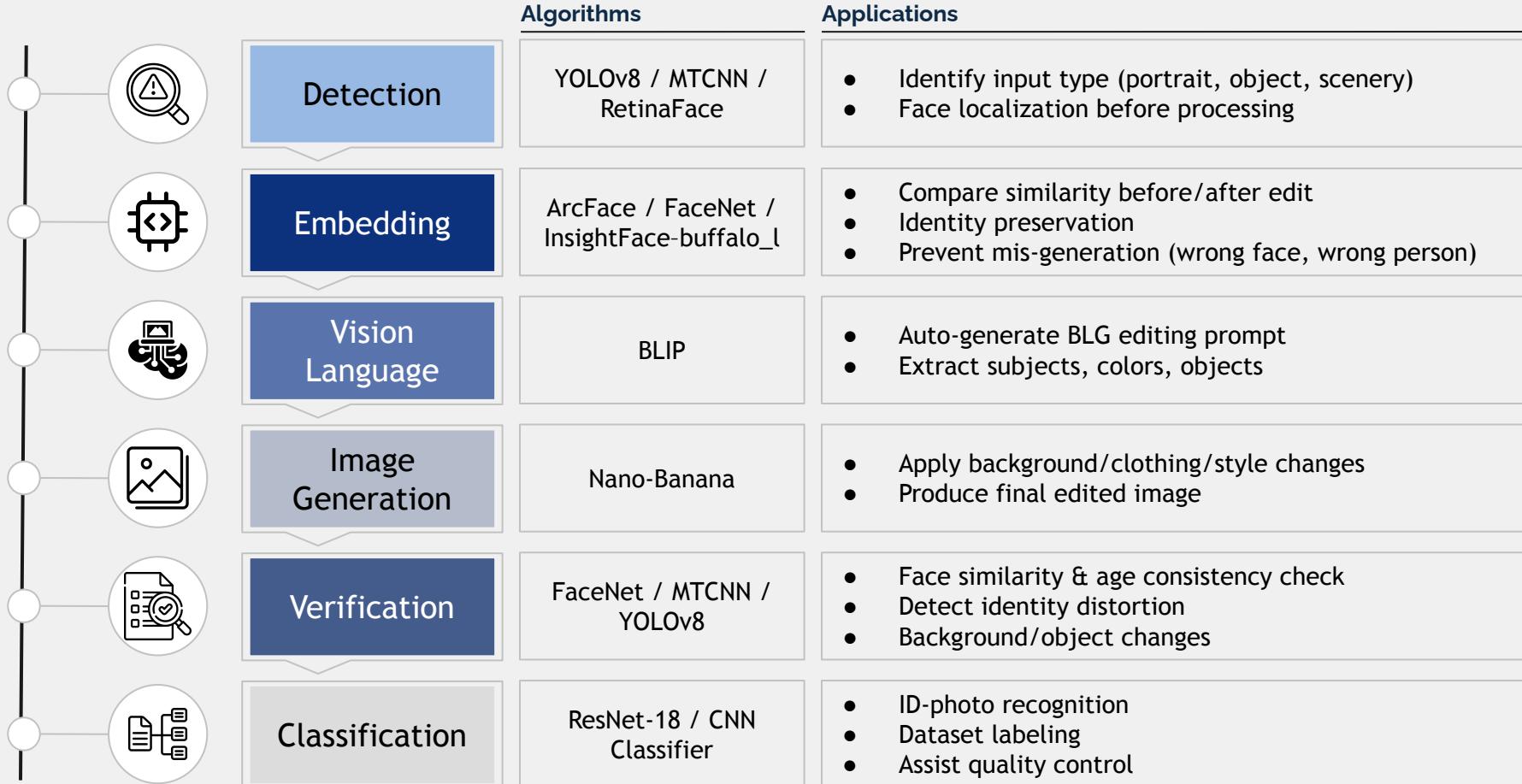
Ethnic & Age



Model Pipeline



Model Pipeline





Data Preprocessing

1 Image Verification

2 Format Standardization

3 Image Resizing

- ✓ Supported formats:
JPEG, PNG, JPG, WEBP, BMP

- ✓ Size range:
[100×100] to [4096×4096] pixels

- ✓ MTCNN requires RGB input

- ✓ Standard format for CV models

- ✓ Resize large images (>1024px)

- ✓ Prevent memory issues

4 Color Space Conversion

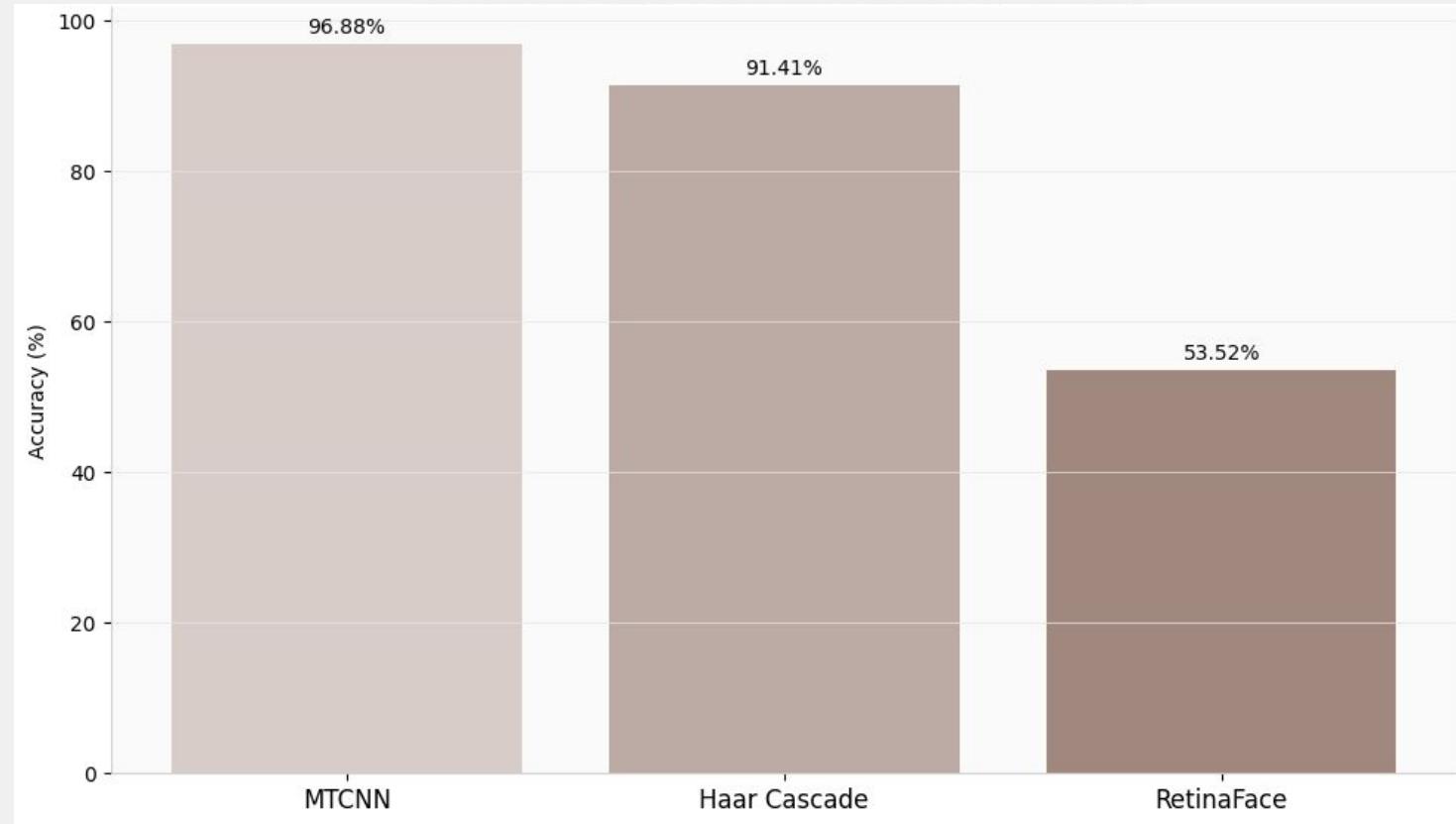
- ✓ BGR→RGB (OpenCV)
- ✓ RGB→Grayscale



Model Comparison

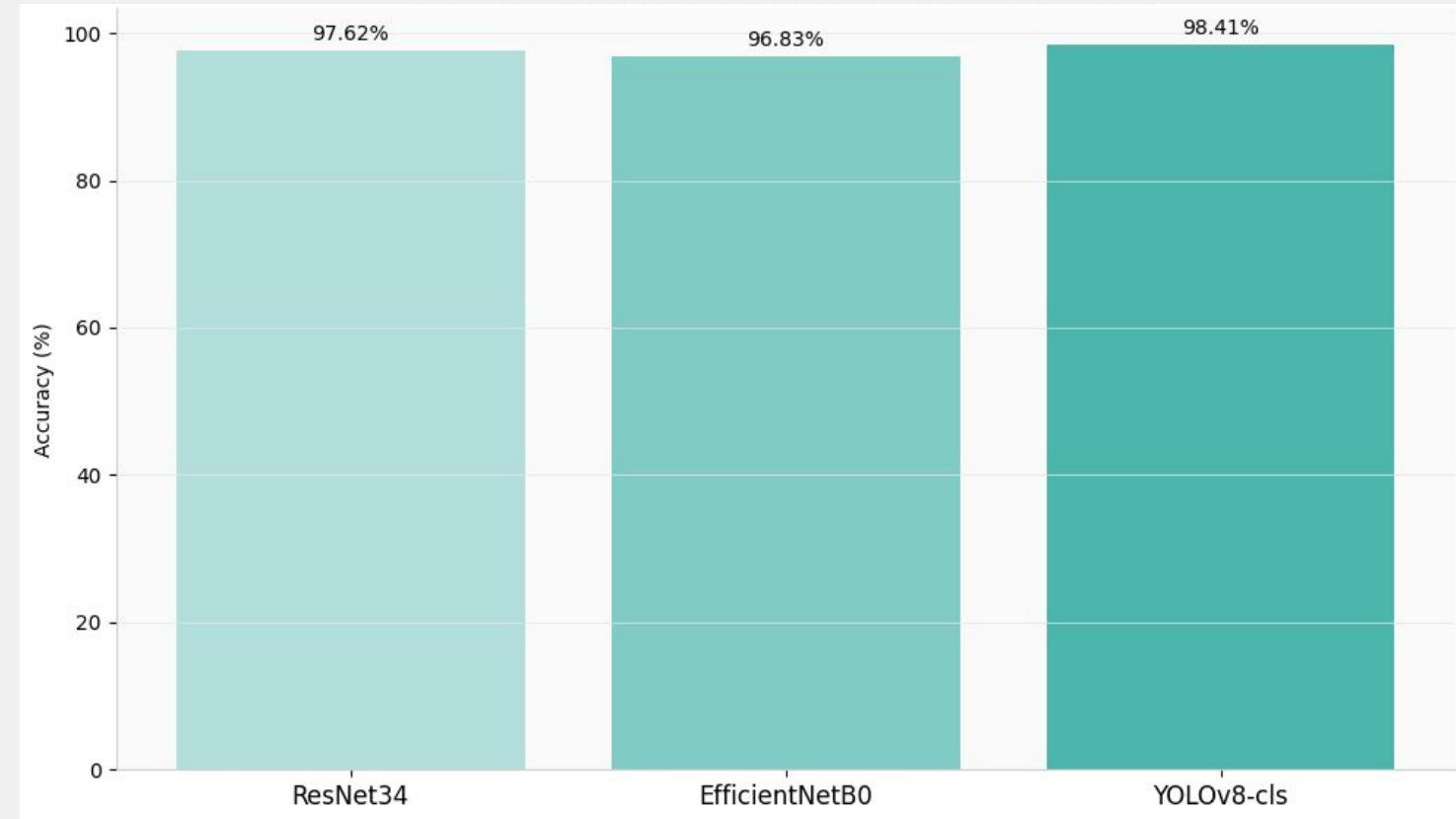


A. Face Detection



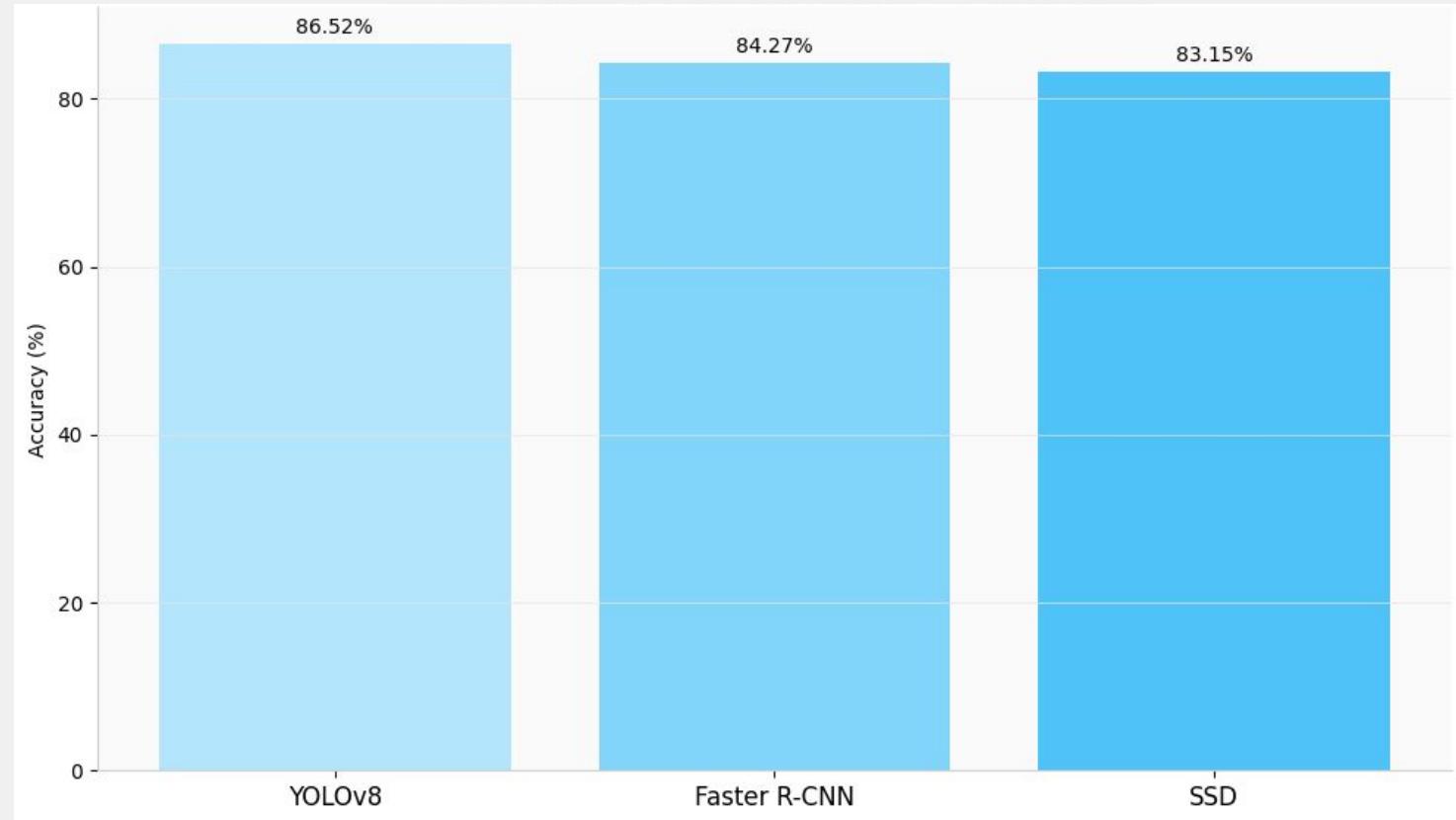


B. Object Detection



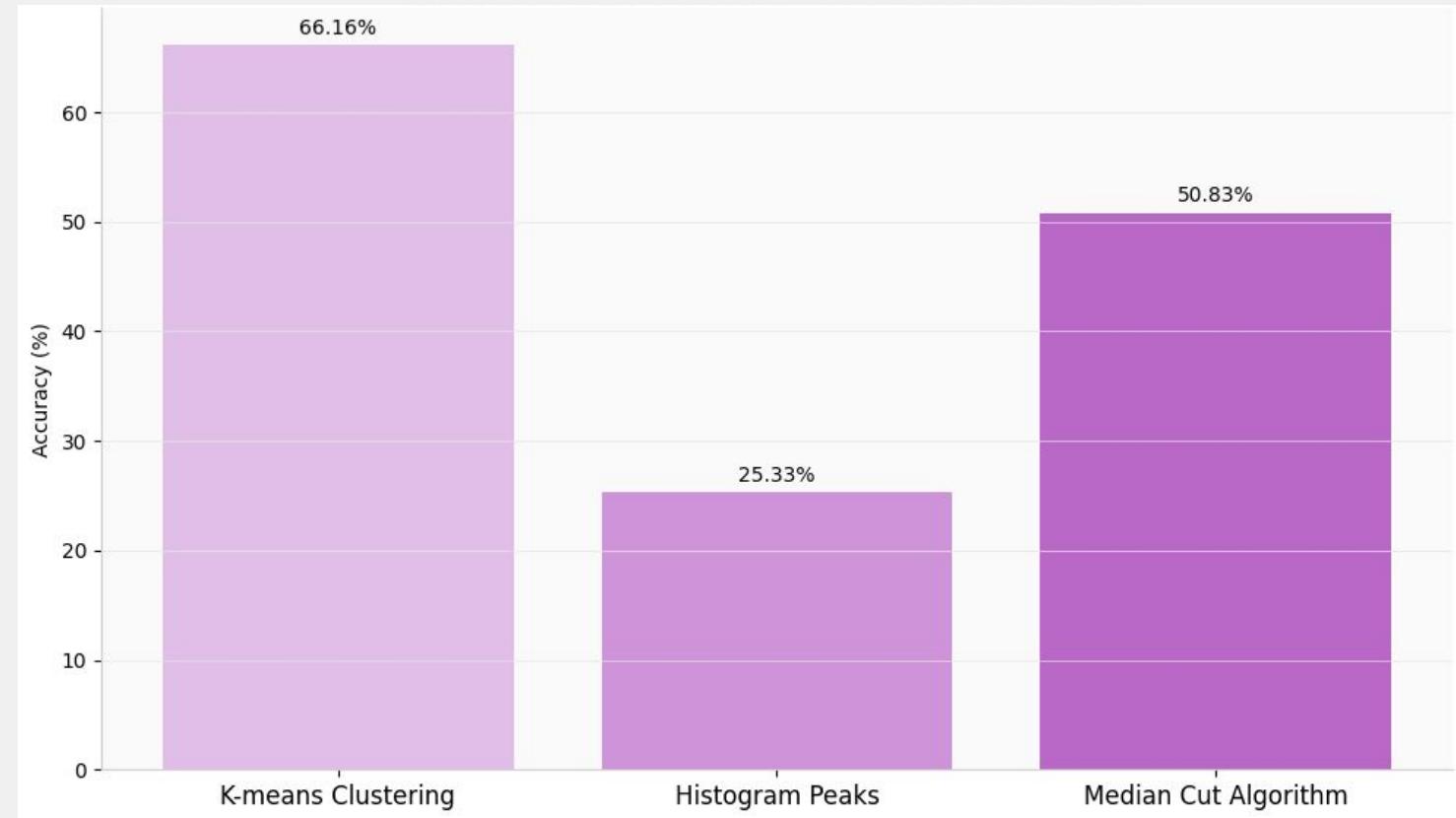


C. Person Detection





D. Color Detection



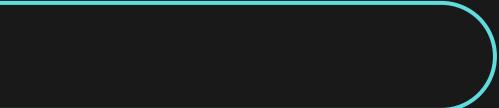


Image Analysis



Image Analysis

Image Analysis

Images Uploaded
1

Status
 Analysis complete!

Ready for Prompt Generation

Per-Image Analysis

Image 1



1 Face(s) Detected

- Face 1: 100.0% confidence (mtcnn)

Objects (2 total)

- person: 81.2%
- cell phone: 31.9%

More Details

Type: **Portrait**

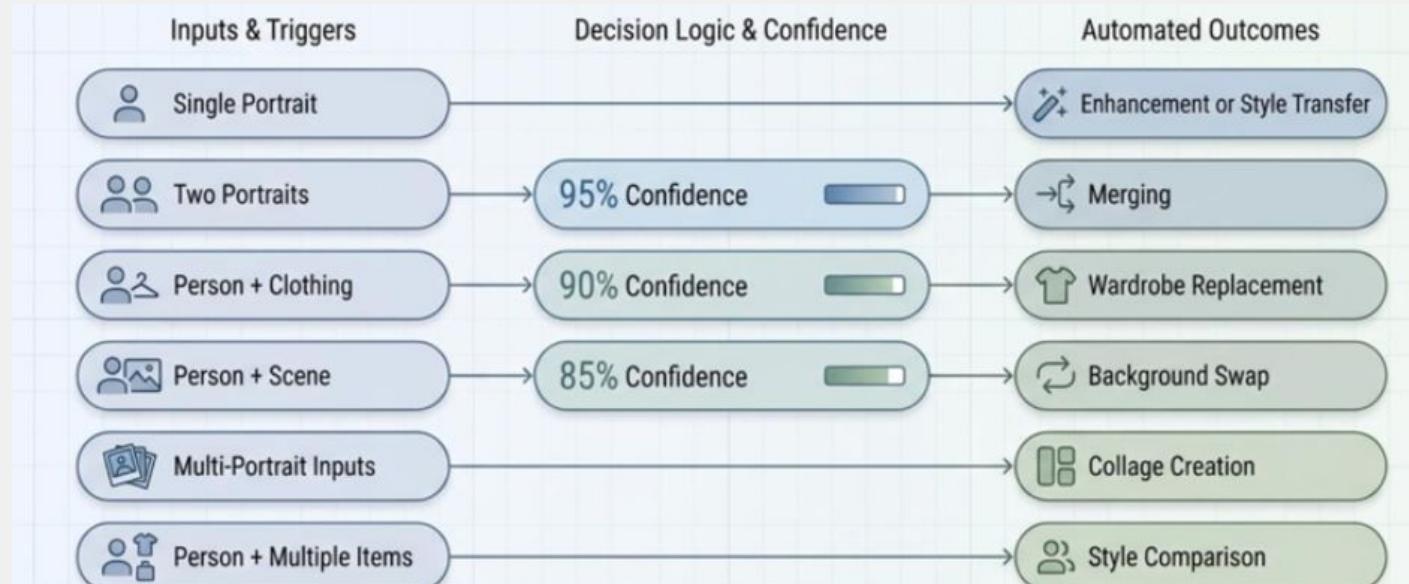
1 2

Suitable

Model	Status	Purpose	Accuracy
MTCNN	Active	Face detection	100%
Haar Cascade	Fallback	Face detection	~70%
YOLO	Optional	Object detection	~90%
K-means	Active	Color analysis	High



Smart Intent Detection (9 Scenarios)



Detected Intent: Enhance Or Style

Reason: Single portrait detected - suitable for enhancement, background change, or style effects

Confidence
90%

Recommended Actions:

- Change Background
- Apply Style Effect
- Add Custom Frame
- General Enhancement





Prompt Generation



Sample Prompts

Cartoon Effect

"Portrait of 1 person with a simple, clean background and calm, relaxed vibe. Add humor with exaggerated expressions or cartoon effects. Preserve facial features and identity."

ID Photo

"Transform into a professional ID photo. **PRESERVE THE EXACT SAME PERSON**, same face, identity, and features. DO NOT alter: face structure, eyes, nose, mouth, skin tone, hair. Change ONLY background to pure white (RGB 255,255,255). **Output: exactly 300x450 pixels.**"

Change Background

"Change background to tropical beach with palm trees and ocean. Keep person, lighting, and perspective unchanged."

Change Clothing

"Change clothing to formal business suit. Keep face, body proportions, and natural look unchanged."



Prompt Customization

A. General Enhancement

Additional Style Effects

Funny, Professional,
Artistic, Vintage, Modern

B. Change Background & Clothing

i) Background

Beach, City, Nature, Office, Party

ii) Clothing

Custom outfits and instructions

C. ID Photo

Background and Color

White, Blue (RGB) Output for
Passport/ VISA/ ID / LINKEDIN

D. Custom

i) Editing

ii) Frame

Birthday, New Year, Christmas

E. Merge People

Group photos

F. Style Effects

Funny, Professional, Artistic,
Vintage, Modern



Image Generation



Format

Funny



Beach



ID Photo



Graduation



Business Formal



Elegant



Generating image with Poe Nano Banana... This may take a moment...



Verification



Result Verification Checkpoints



Face Detection



Age Consistency



Object Changes



Image Quality



Prompt & Background changes

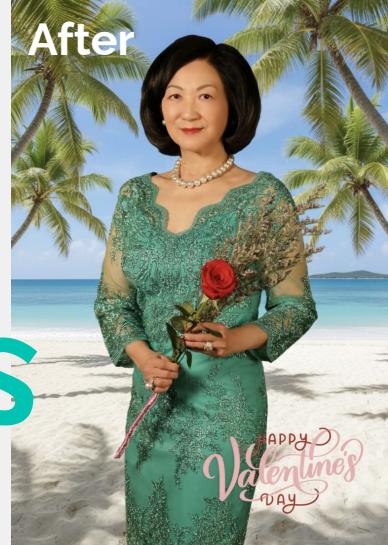
+



ID Photo Verification



Before



After

VS



Step 1: Face Preservation

Function	Models Applied	Passing Condition
Detect <u>unwanted facial changes</u>	MTCNN: Face detection FaceNet: Generates embeddings for faces	> 0.6 face similarity

✓ ✅ FACE Verification

Status: success

Message: All 1 faces preserved

Original Faces	Edited Faces	Matched Faces
1	2	1

📊 Similarity Scores:

Average	Minimum	Maximum	Threshold
0.985	0.985	0.985	0.600

Good similarity score (0.985). Faces appear to match well.



Step 2: Age Consistency

Function	Models Applied	Passing condition	AGE Verification
Detect <u>unwanted age changes</u>	MTCNN: Face detection Deepface: Generates embeddings for faces	> 0.75 age similarity Maximum allowable difference in estimated age = 10	Status: success Message: Age appears consistent (similarity: 0.9852) Details: <ul style="list-style-type: none">• age_consistent: True• estimated_consistency_score: 0.985• tolerance_threshold: 0.750



Step 3: Object Changes

Function	Models Applied	Passing condition
1.Detect whether <u>person class removed</u> 2.Show users <u>new and deleted class</u> in the image	Yolo	No person class removed

✓ OBJECTS Verification

Status: success

Message: No protected objects removed

Object Detection Results:

Original Objects	Edited Objects
1	1

Detected Classes:

Original Classes:	Edited Classes:
• person	• person



Step 4: Image Quality

Function	Models Applied	Passing condition
<u>Detect image quality</u>	PIL python library	min_resolution: (512, 512) Quality_score > 0.85

▼ QUALITY Verification

Status: success

Message: Quality acceptable

Details:

- **quality_acceptable:** True
- **resolution:** 832x1248
- **meets_min_resolution:** True
- **min_resolution_required:** 512x512
- **is_grayscale:** False
- **quality_score:** 0.850



Step 5 & 6: Prompt & Background Changes

Model	Max Token Limit	Model Name	Notable Strengths	Typical Use Cases
CLIP	77	penai/clip-vit-large-patch14-336	Pioneering, Robust	Image-text matching, Retrieval
OpenCLIP	512 (T5)	ViT-B-32	Fast, Easy, Popular	Simple matching, Retrieval
BLIP-2	512-2048	Salesforce/blip-itm-base-coco	Flexible, Long prompts	Captioning, Similarity, VQA
ViLT	512	dandelin/vilt-b32-finetuned-coco	Efficient, Transformer-only	Retrieval, Matching



Step 5 & 6: Prompt & Background Changes

Function	Models Applied	Passing condition
Detect matchness of prompt	BLIP-2 Salesforce/blip-itm-base-coco	>0.25 similarity improvement

```
from transformers import BlipProcessor, BlipForImageTextRetrieval
# Use BLIP for Image-Text Retrieval (better for similarity scoring)
model_name = "Salesforce/blip-itm-base-coco"
self.blip_processor = BlipProcessor.from_pretrained(model_name)
self.blip_model = BlipForImageTextRetrieval.from_pretrained(model_name)
print("✓ BLIP model loaded (pretrained, ITM for retrieval)")
```

✓ PROMPT Verification

Status: success

Message: Edited image matches prompt better (score: 0.2964, improvement: +0.2962)

Details:

- prompt_adherence: True
- edited_matches_better: True
- similarity_original_prompt: 0.000
- similarity_edited_prompt: 0.296
- similarity_original_edited: 0.900
- improvement: 0.296
- threshold: 0.250
- prompt: Change the background to festive party environment with decorations, keep 8 p



Step 7: ID Photo Verification

Function	Models Applied
Check whether generated image is <u>ID photo</u>	Trained Convolutional Neural Network (CNN) classifier Model validation accuracy: 92.31%

✓ ID PHOTO Verification

Original Image Edited Image

Not ID Photo (expected)

ID Photo

Confidence
0.00%

Confidence
100.00%

Class Probabilities

- id_image: 0.00%
- not_id_image: 100.00%

Class Probabilities

- id_image: 100.00%
- not_id_image: 0.00%





One more thing...



Q & A

04





Thanks!