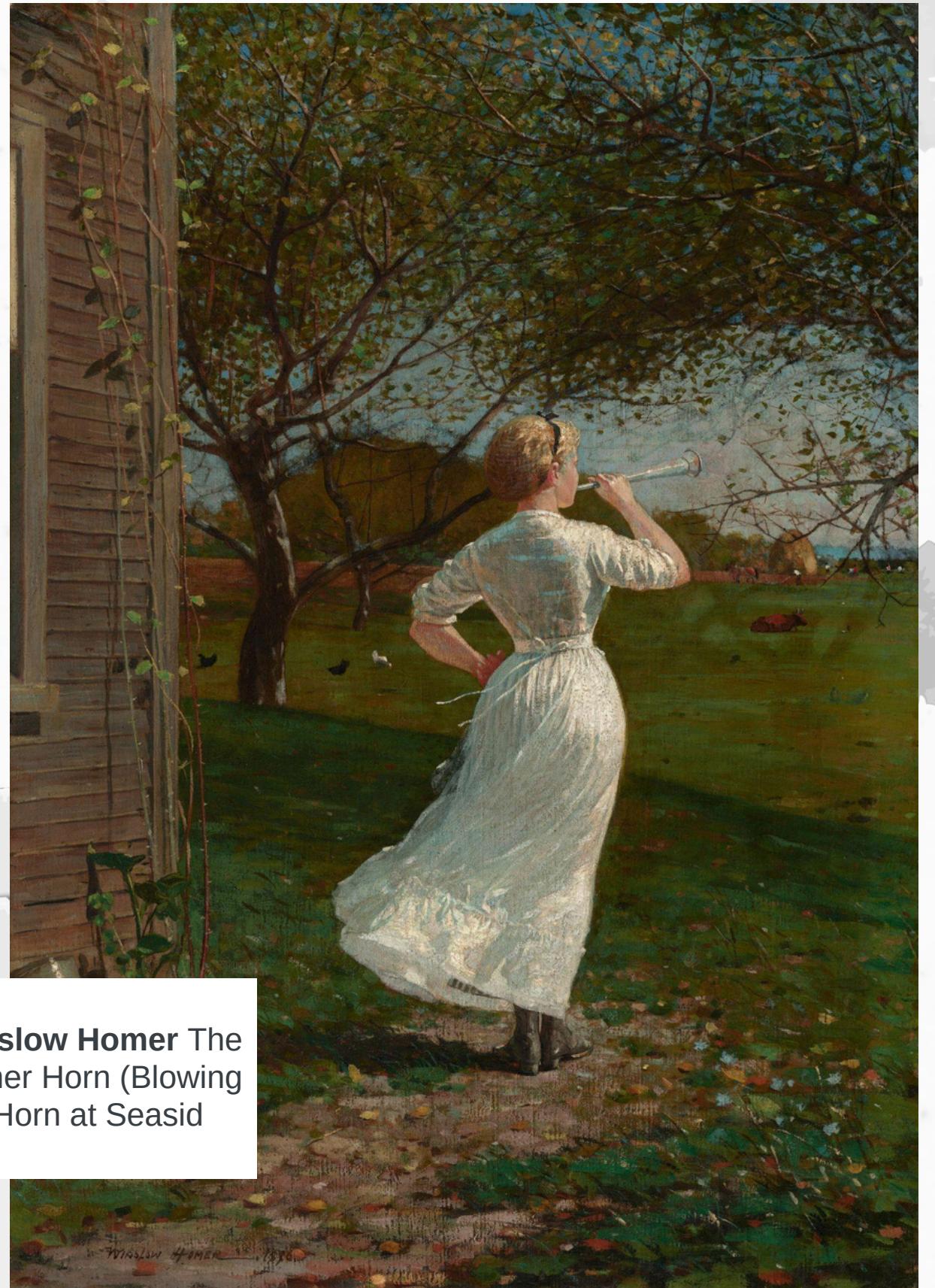


“Art-Vision”

**Art Critique &
Creation Tool**

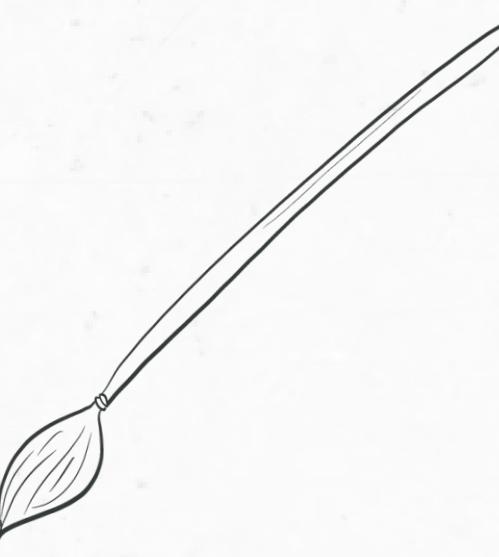
**Supervised By
Dr.Mohammed El-Said**



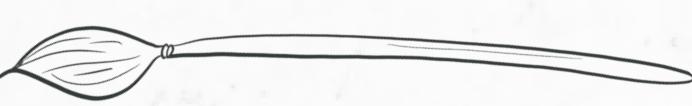
**Winslow Homer The
Dinner Horn (Blowing
the Horn at Seaside)**

Art-Vision Team

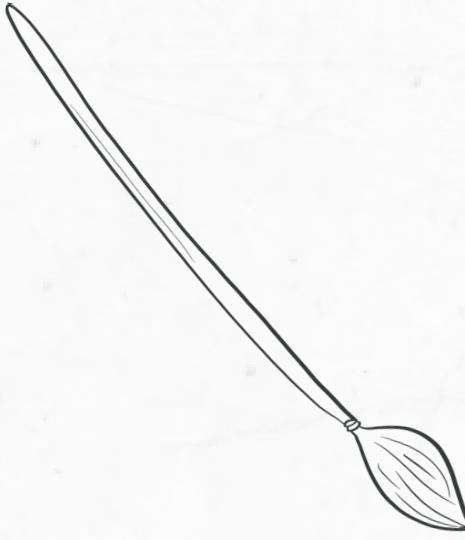
Sama Gomaa



Sama Mostafa



Sara Mahmoud



Sara Ashraf



Azza Mohamed



Israa Hossam Eldin



Agenda

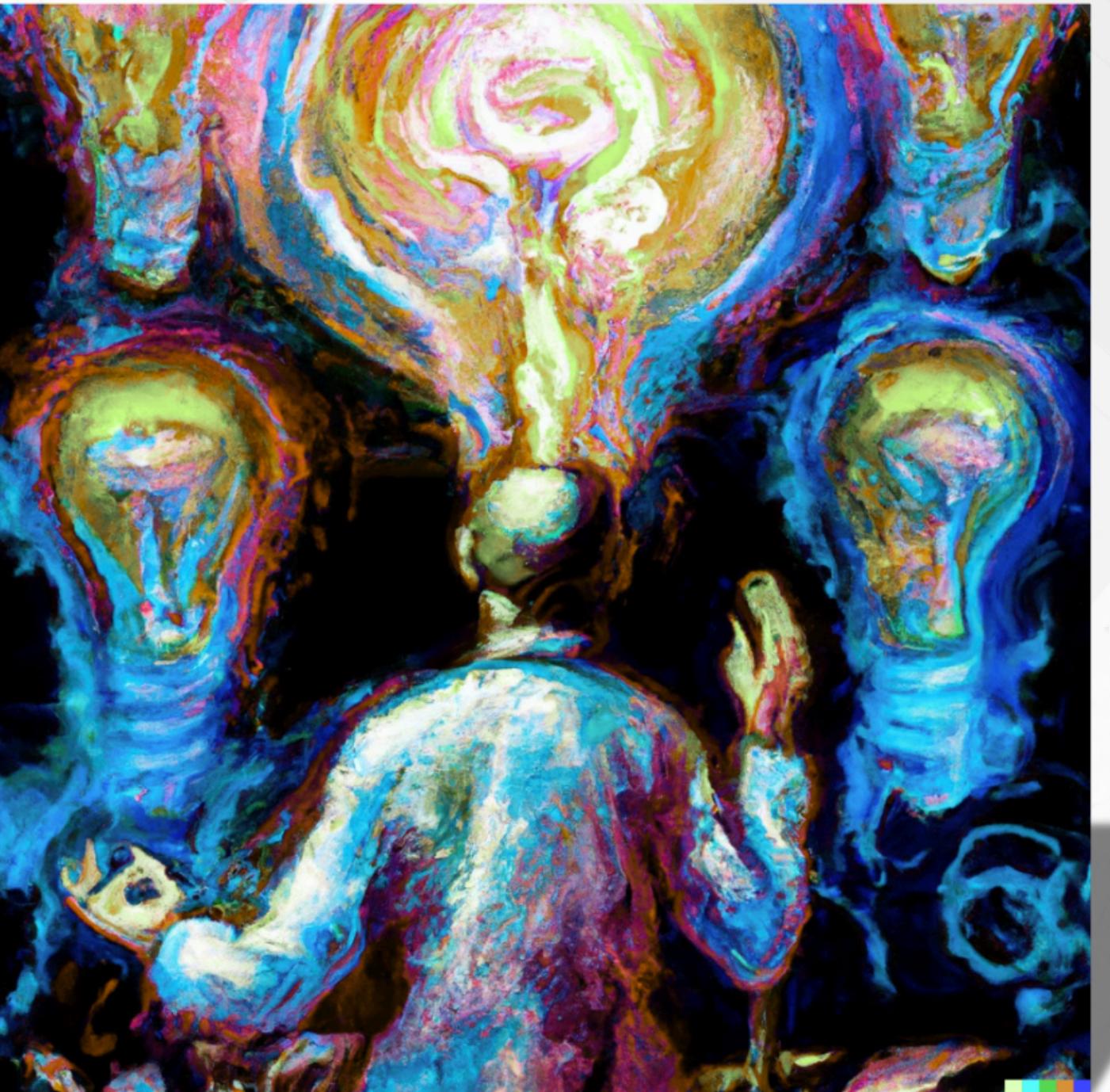
- 01 Problem Statement
- 02 Proposed Solution
- 03 Business Vision
- 04 Key Features
- 05 System Architecture
- 06 Models Architectures & Datasets

- 07 Testing & Evaluation
- 08 Limitation
- 09 Demo QR
- 10 Future work
- 11 Selected References

For more info:
[Documentation link](#)

01 Problem Statement

- Lack of constructive feedback for artists.
- Artists Need for inspiration and creative direction.



Organize Your Ideas



02 Proposed Solution

- AI-powered tool for generating creative artworks based on user descriptions.
- AI-driven critique tool analyzing artistic elements.

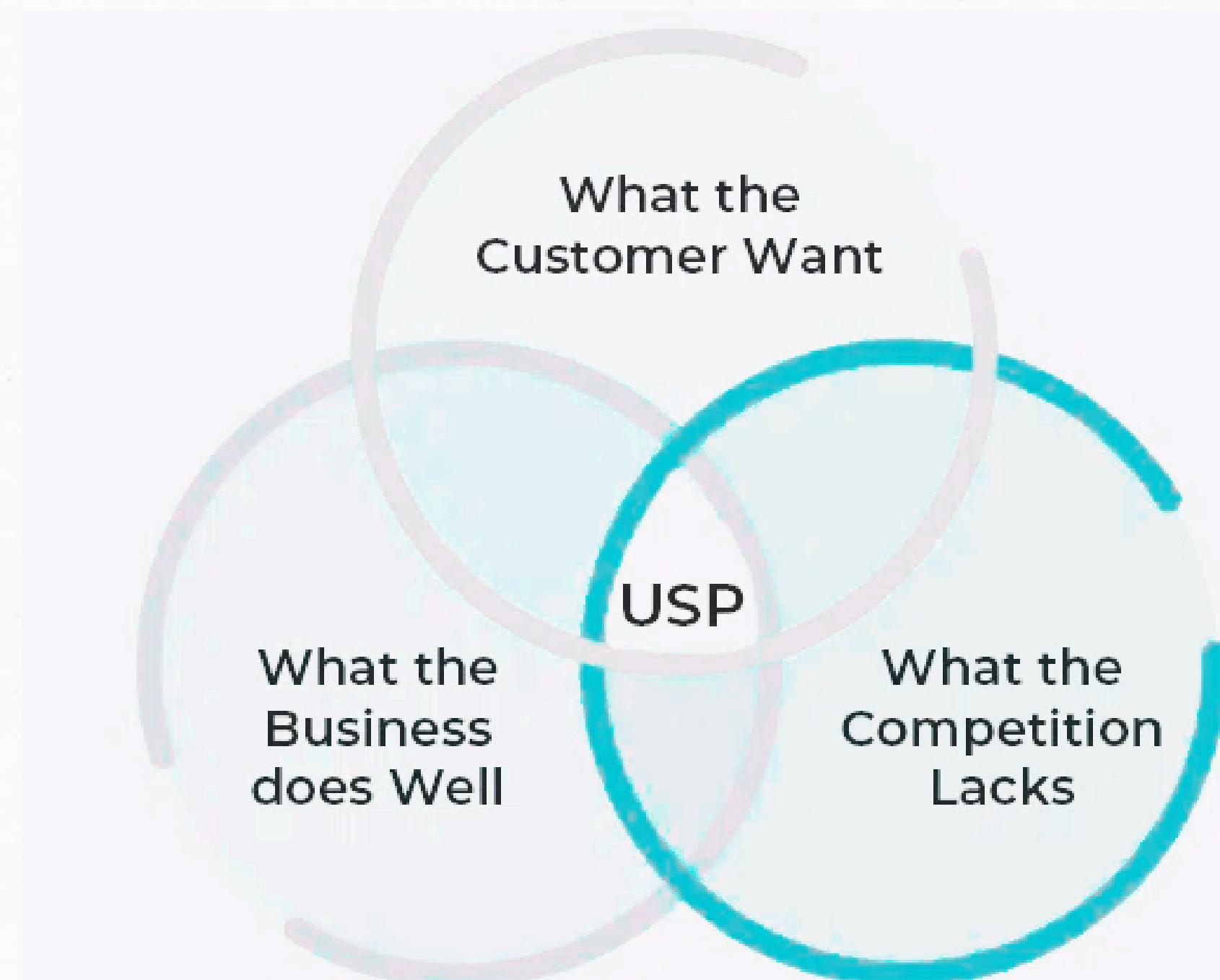
03

Business vision



- Unique Selling Proposition

- ArtVision is trained exclusively on art-specific dataset.
- AI-Powered Dual Functionality.



- Revenue Streams

- Subscription Plans
- Sponsored Content & Ads



For more info:
[Business Model link](#)



EGYPT
IoT & AI
CHALLENGE



2025
Egypt IoT & AI Challenge

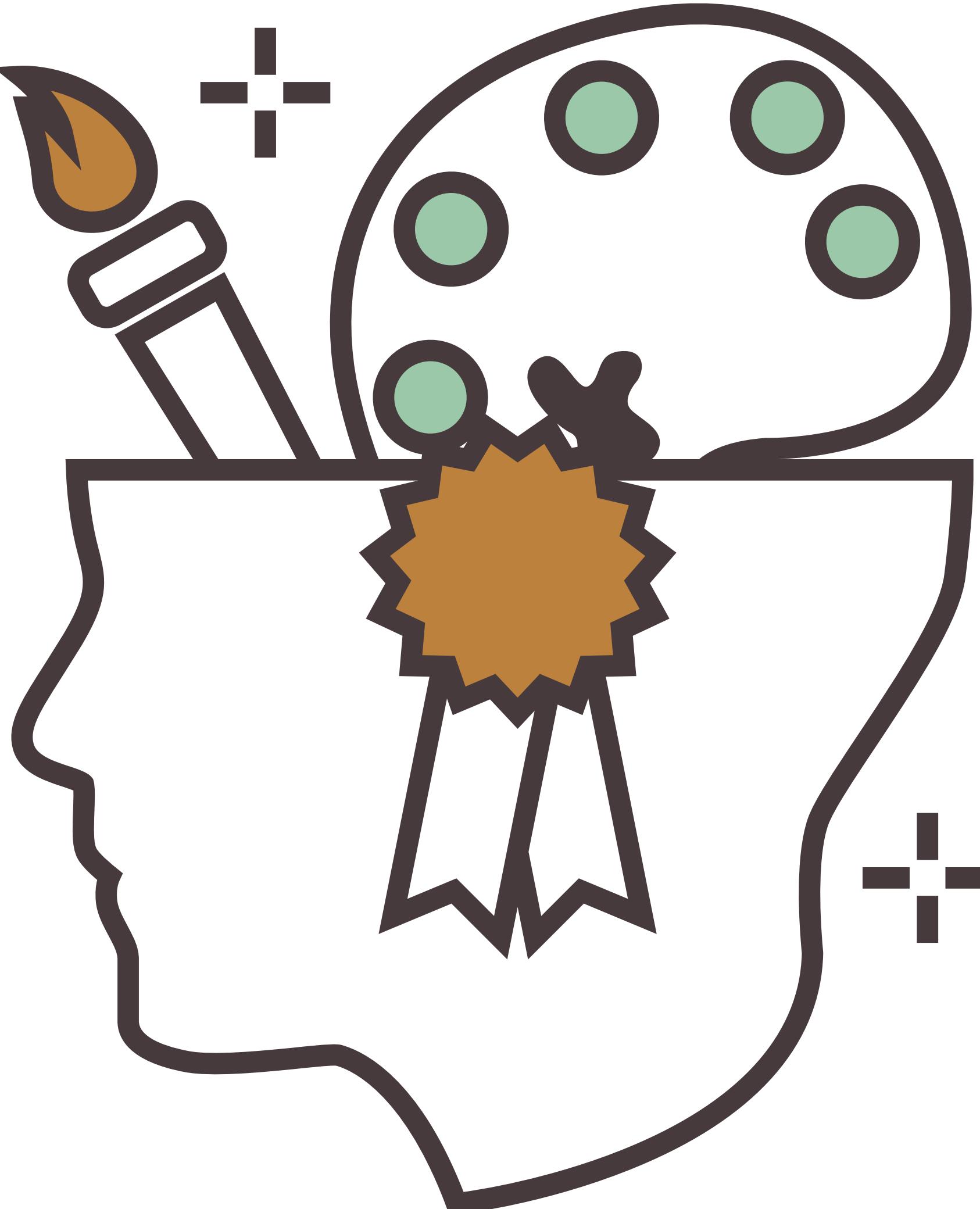
ORGANIZED BY

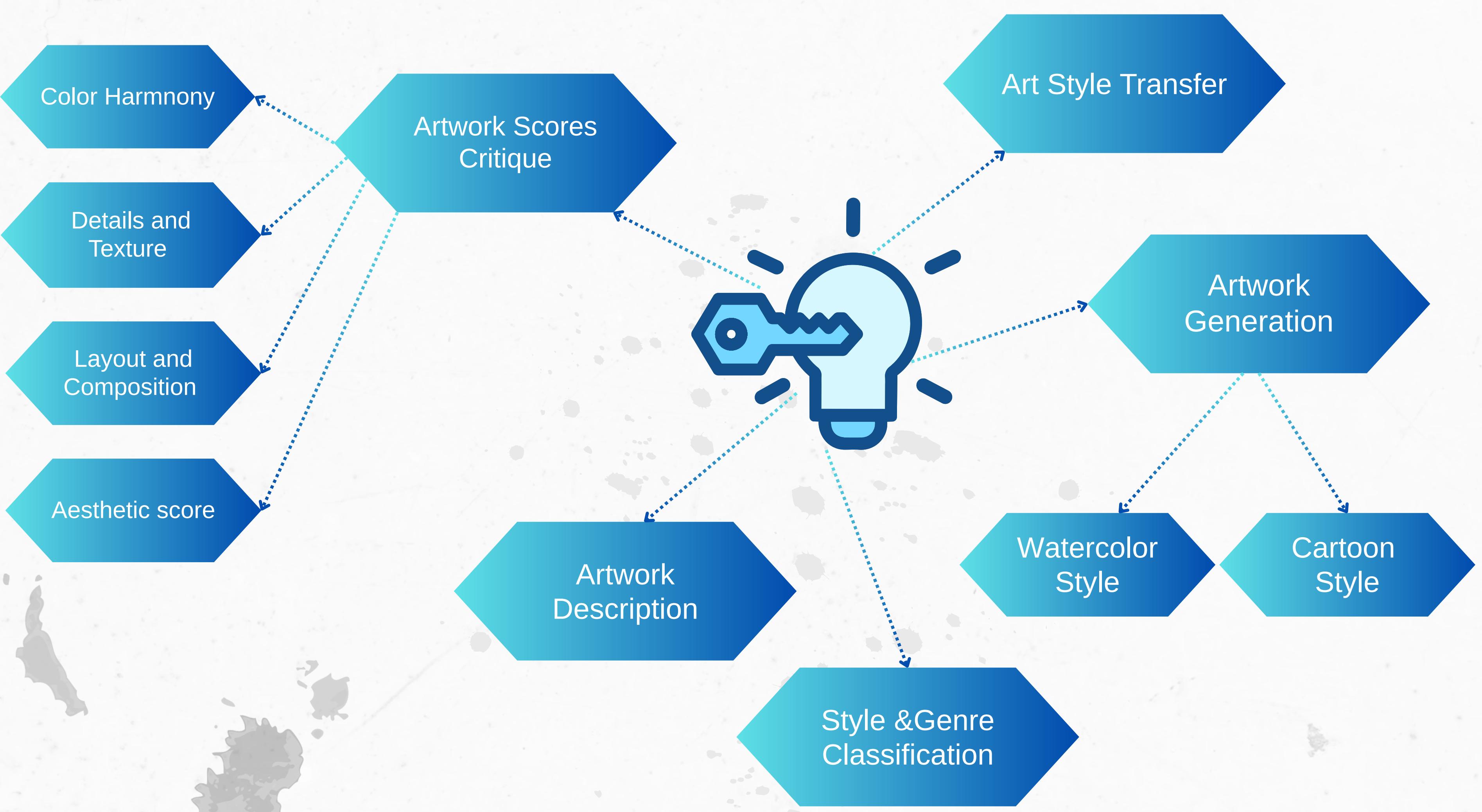


04 Key Features



Why ArtVision
Stands Out?





05 System Architecture

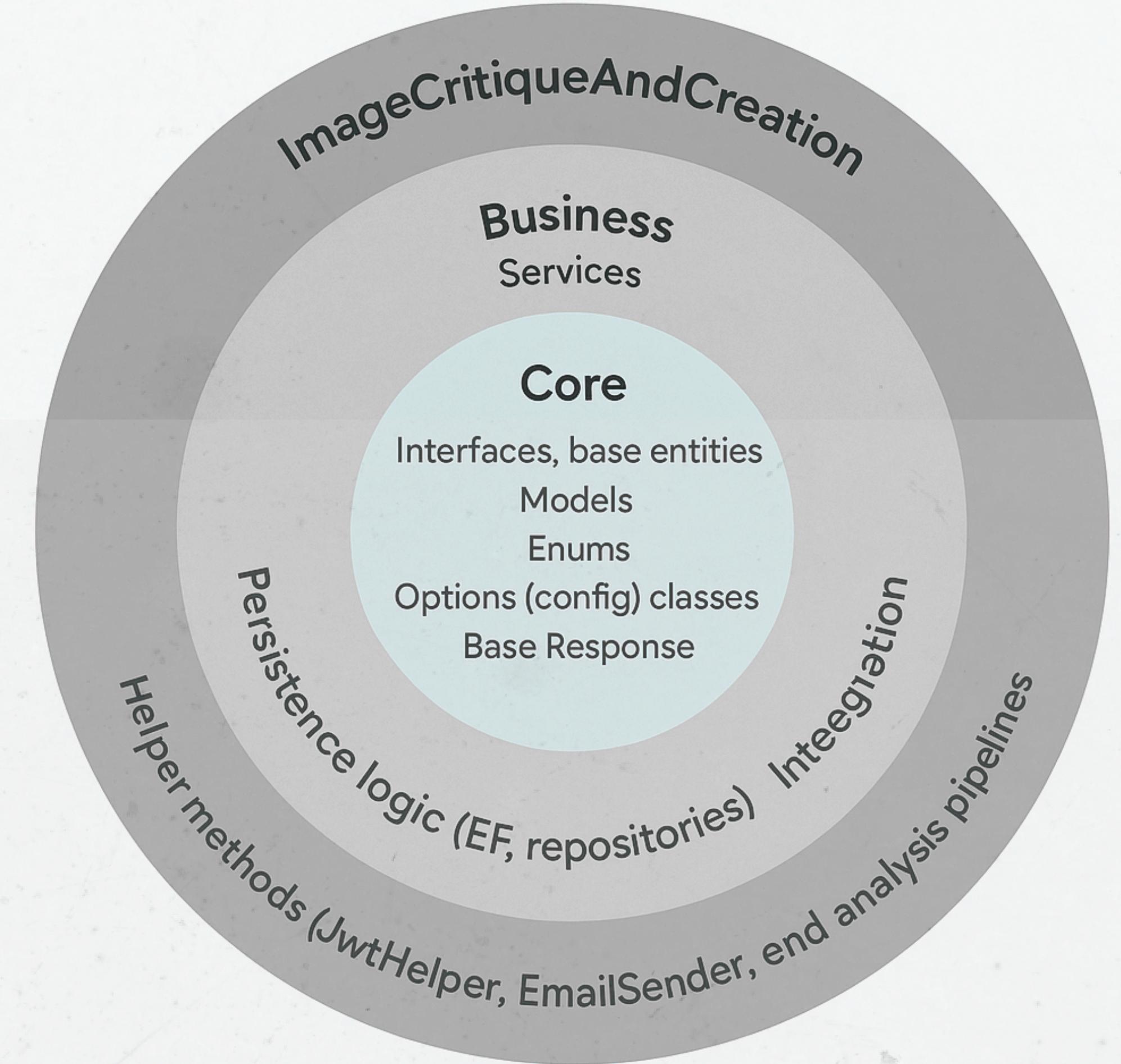
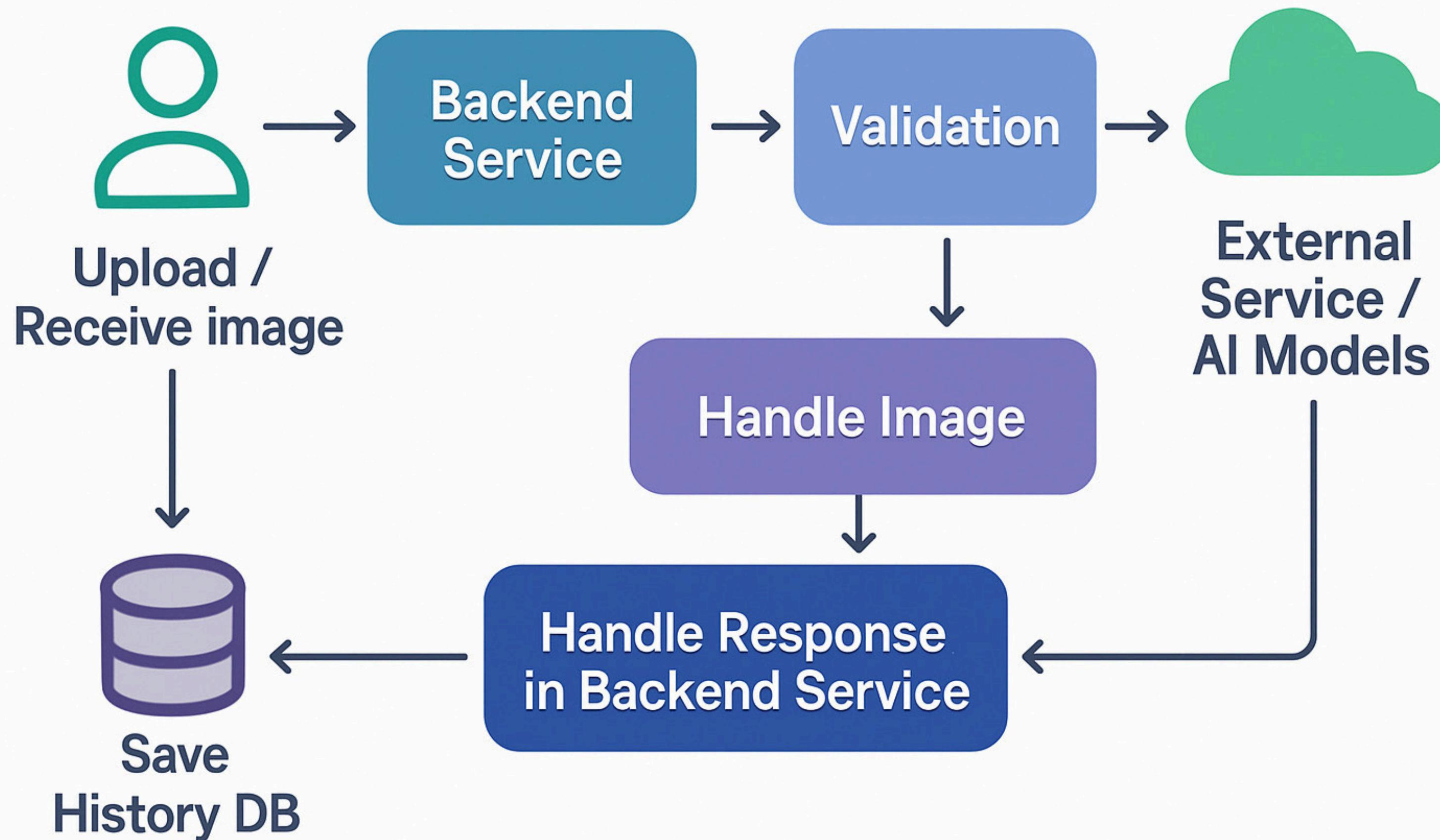
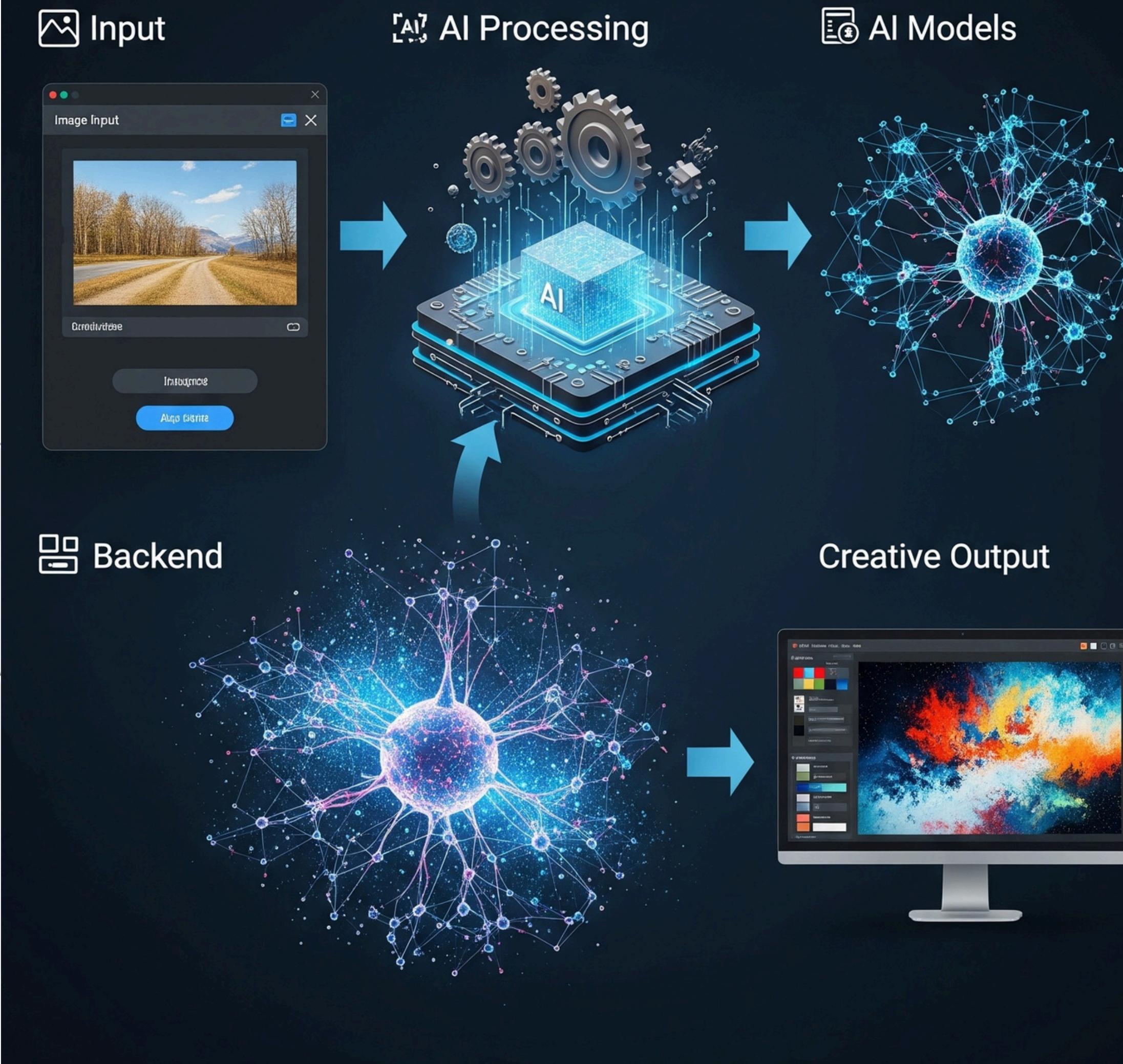


Image Handling Workflow



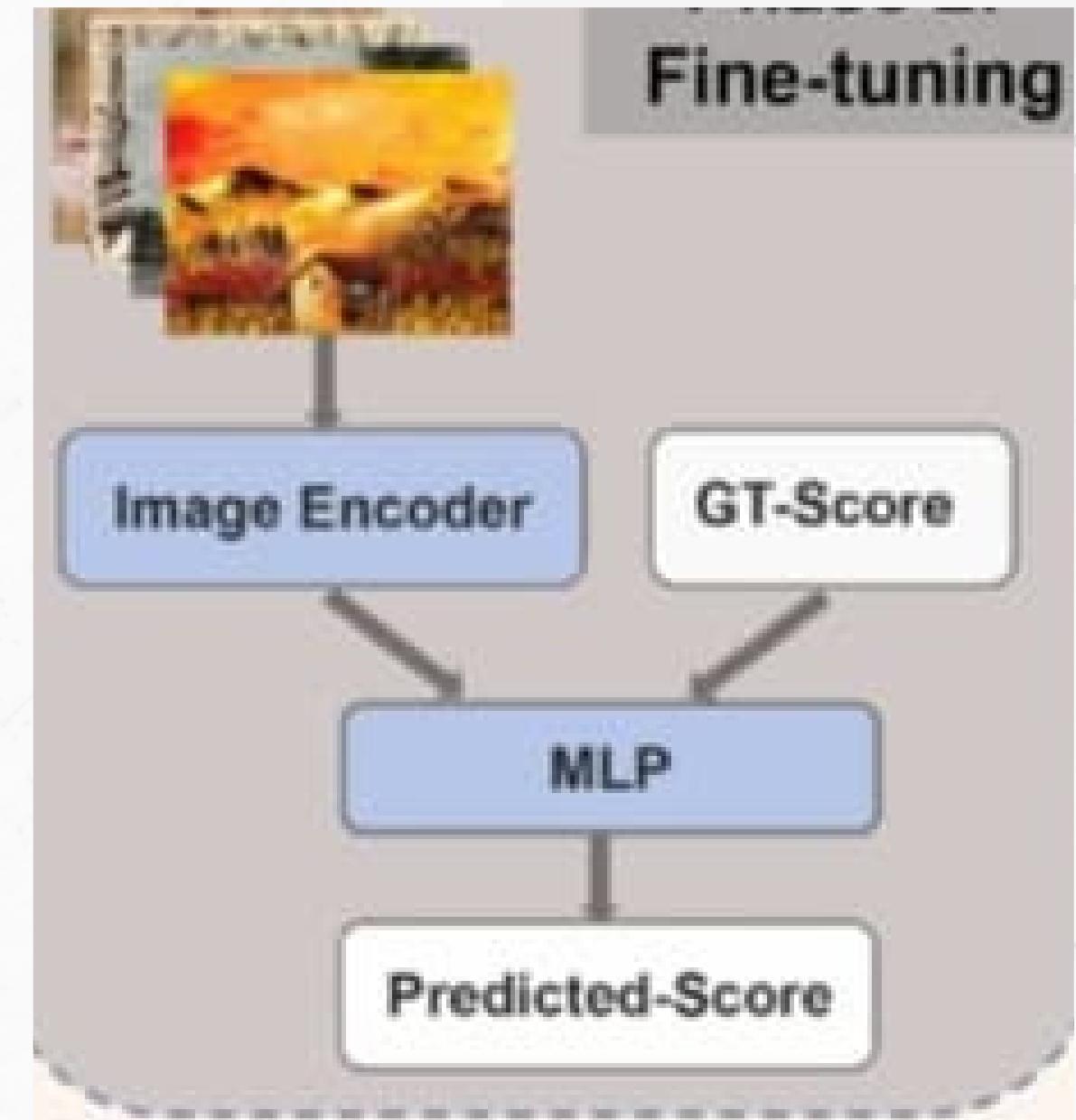
06

Models Architecture & Datasets



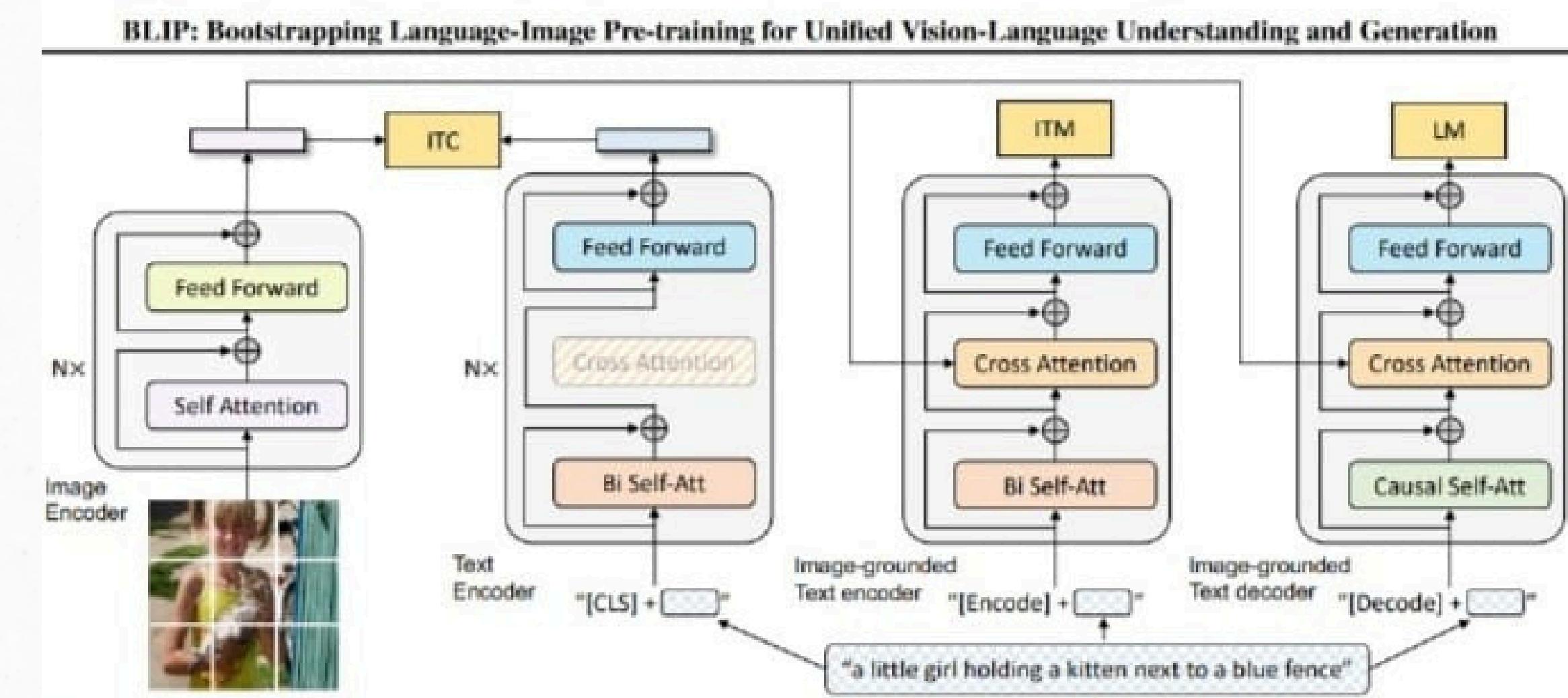
Aesthetic Score Prediction using AesCLIP

- Uses CLIP + MLP to predict aesthetic dimensions.
- Trained on APDDv2 dataset that has created with participation of over 40 experts from the painting domain, comprising 10,023 images, 85,191 scoring annotations, and 6,249 language comments.
- Outputs scores only for: Total aesthetic, color, texture, composition.



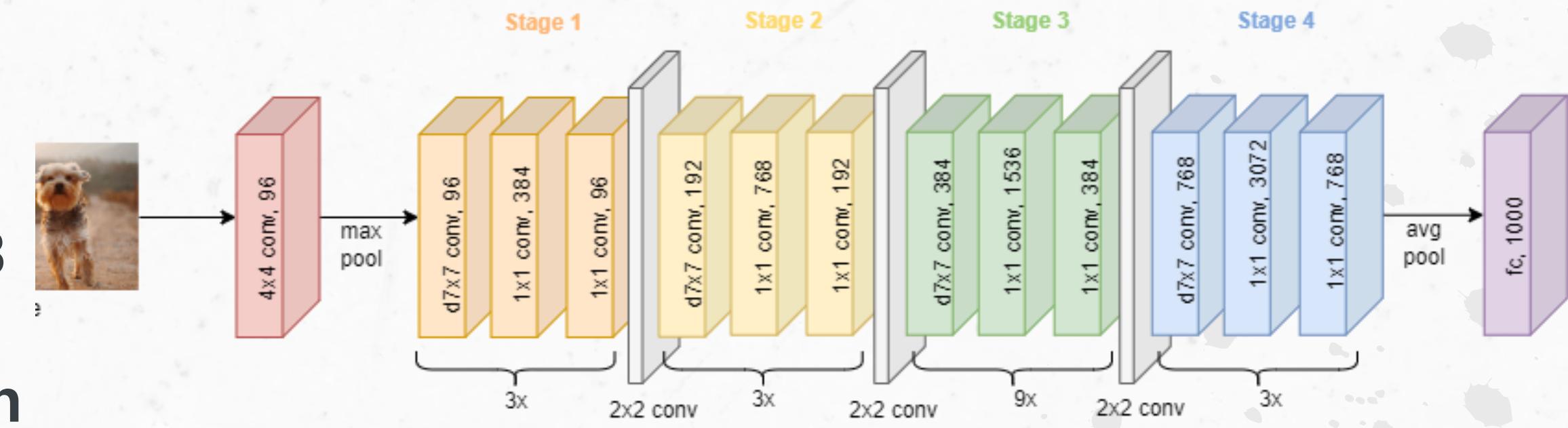
Art Captioning with BLIP

- This model uses the pretrained BLIP-Base architecture with a frozen vision encoder, fine-tuned on 20k WikiArt image-caption pairs to generate rich, descriptive captions of artworks



MultiTask Genre & Style Classification Model

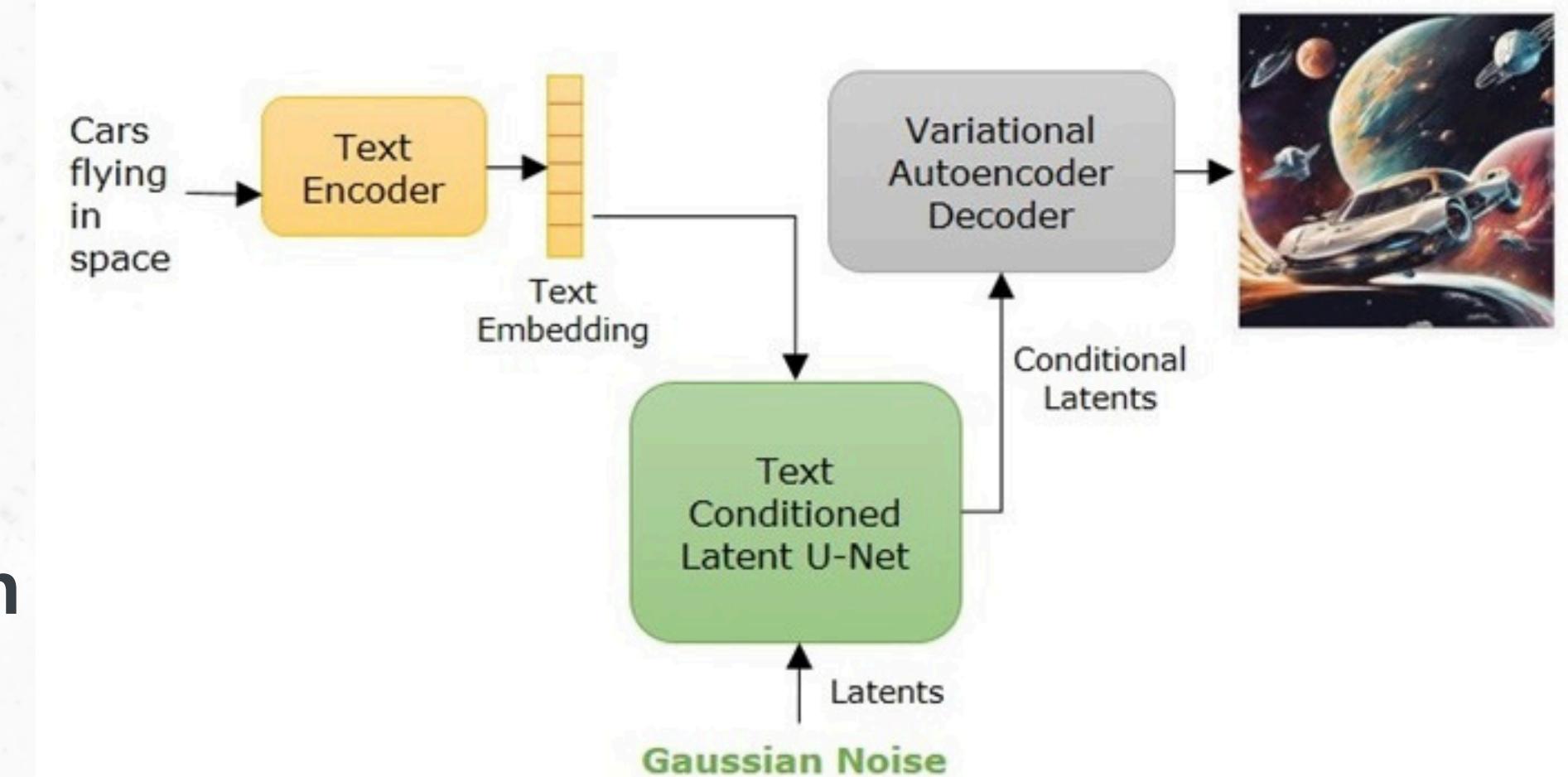
- Uses ConvNeXt-Base as backbone for deep feature extraction.
- Built on WikiArt dataset after cleaning, with: 44,868 training images, 19,149 validation images and with 10 genres and 27 styles.
- Trained to predict both genre and style simultaneously from artworks.



Watercolor Artwork Generation

- Fine-tuned Stable Diffusion XL with LoRA to generate images in a watercolor painting style based on text prompts.
- Custom dataset of only watercolor paintings from WikiArt, each image paired with a caption to guide the model during training.

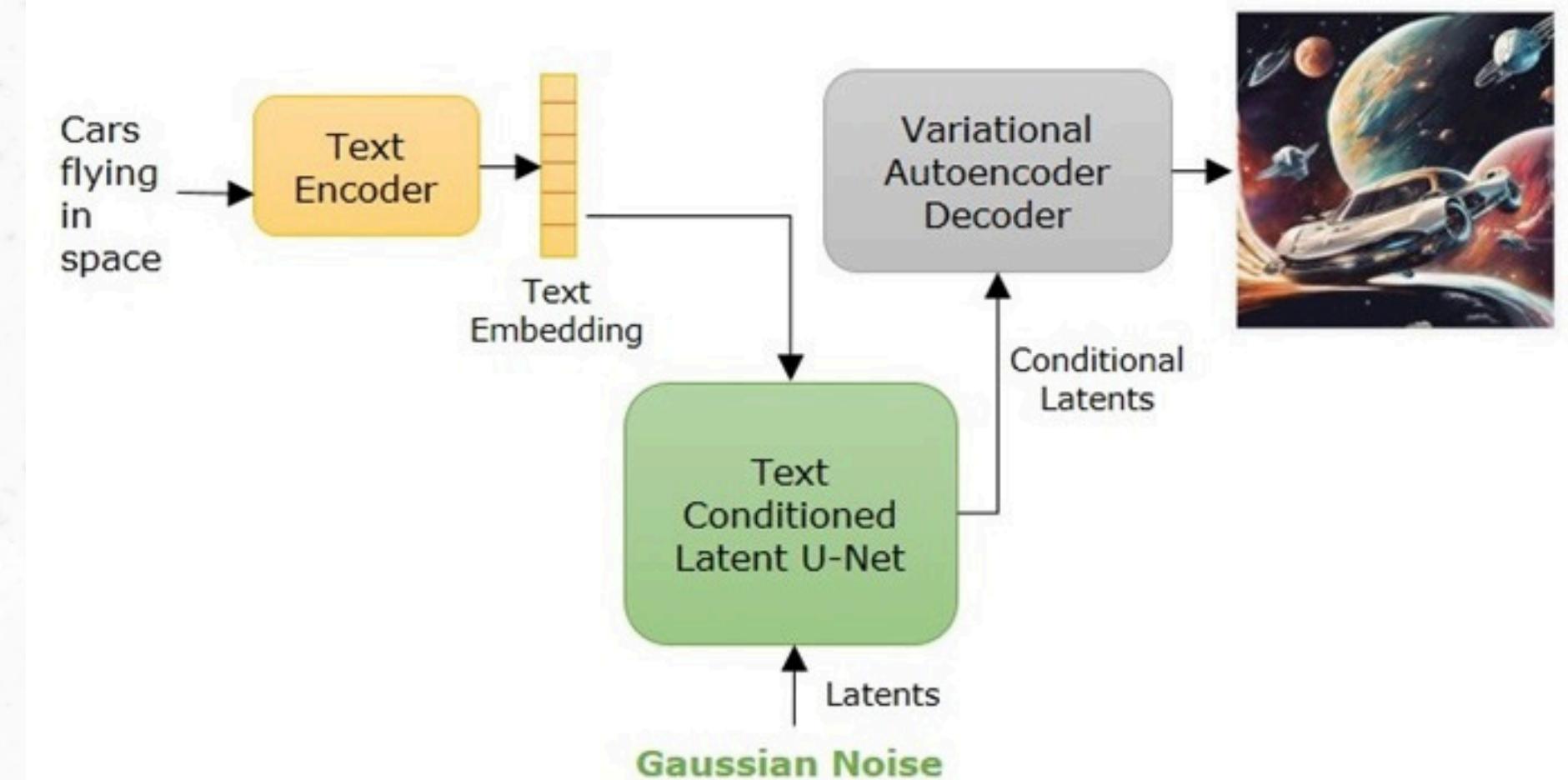
LATENT DIFFUSION ARCHITECTURE



Cartoon Artwork Generation

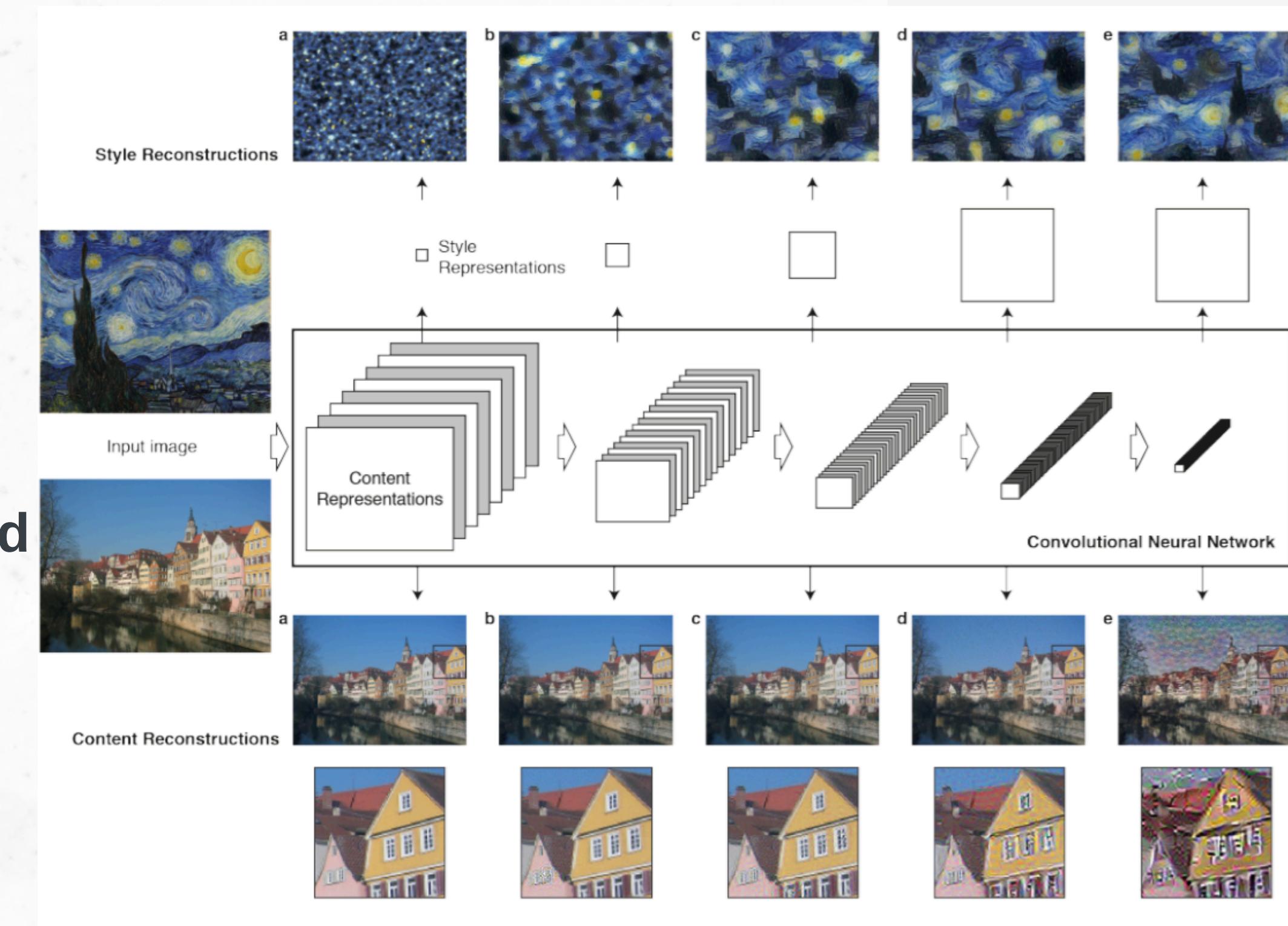
- Utilizing a pretrained Stable Diffusion model fine-tuned in Disney-Pixar cartoon style

LATENT DIFFUSION ARCHITECTURE



Art Style Transfer Model

- This Feature uses a pretrained model (`magenta/arbitrary-image-stylization-v1-256`) to apply artistic style transfer in real-time. Both content and style images are processed and blended to generate stylized outputs.



07 Testing & Evaluation

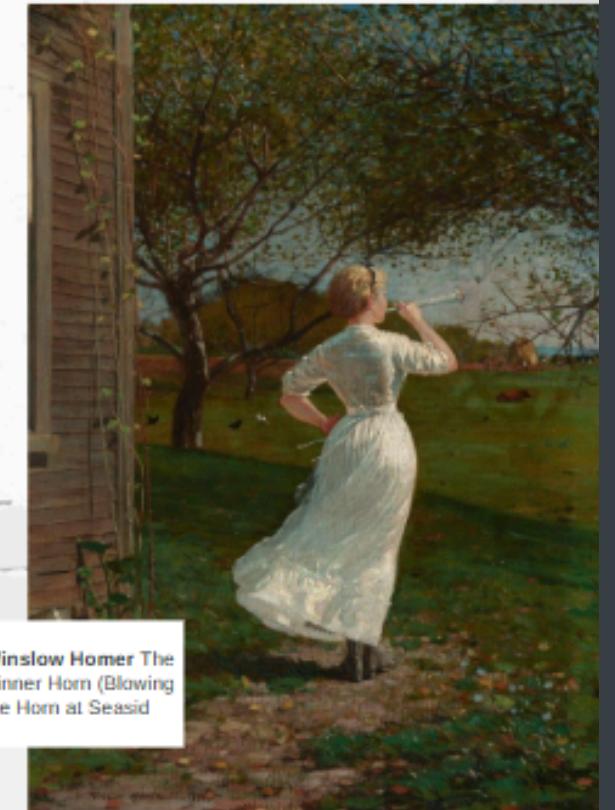
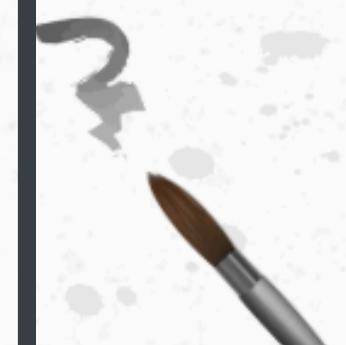
Feature	Task	Metrics / Comments
Artwork Captioning	Image-to-Text Generation	BLEU-1: 0.7744 , BLEU-2: 0.7336, BLEU-3: 0.6991 , BLEU-4: 0.6617 ROUGE-L: 0.8030 , SPICE: 0.6858
Genre & Style Prediction	Classification	Genre: Loss: 0.7268, Accuracy: 83.11% Style: Loss: 1.3600, Accuracy: 70.49%
Watercolor Artwork Generation	Text-to-Image Generation	Visually assessed on prompt-response quality and style match.
Aesthetic Scores Prediction	Regression	Color Harmony: Loss: 0.0749 , MAE: 0.2912 , MSE: 0.1537 Texture: Loss: 0.1893 , MAE: 0.4962 , MSE: 0.3988 Composition: Loss: 0.2034 , MAE: 0.5168 , MSE: 0.4294 Aesthetic score: Loss: 0.1831 , MAE: 0.4834 , MSE: 0.3905

08 Limitations

“Art-Vision”

**Art Critique &
Creation Tool**

**Supervised By
Dr.Mohammed El-Said**



Winslow Homer The
Dinner Horn (Blowing
the Horn at Seaside)

Limitation

Limited Style Diversity :

Only supports watercolor and cartoon generation.

High Hardware Requirements :

Some models require GPU to run efficiently.

Subjectivity in Critique :

AI scores may not fully match human artistic perception.

09 Demo QR



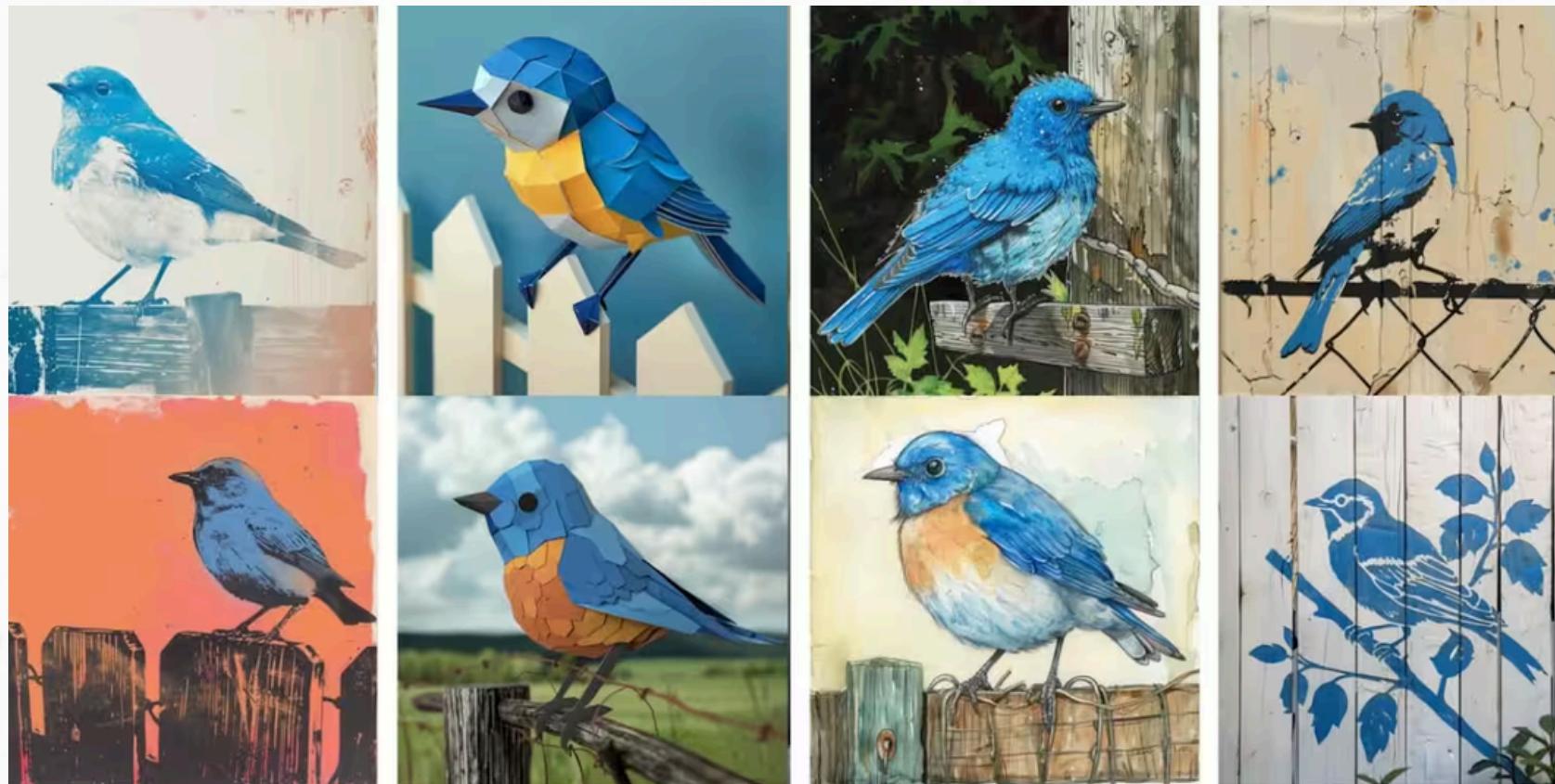
Scan me

A dark blue curved arrow originates from the text "Scan me" and points directly at the center of the QR code.

10 Future Work



Future work



3D

References

<https://github.com/openai/CLIP>

<https://huggingface.co/stabilityai/stable-diffusion-xl-base-1.0>

<https://www.kaggle.com/datasets/steubk/wikiart>

<https://github.com/BestiVictory/APDDv2>

<https://huggingface.co/stablediffusionapi/disney-pixar-cartoon>

https://www.tensorflow.org/hub/tutorials/tf2_arbitrary_image_stylization



Thanks!