GEN AI - IMAGE RESIZERS

Team

1. College Professor(s):
   Dr. Ranjeet Kumar - ranjeet.k@vit.ac.in
   Dr. Jayanthi R - jayanthi.r@vit.ac.in
2. Students:
   1. C R Lavanya - lavanya.2022a@vitstudent.ac.in
   2. Pojesh Kumar R - pojeshkumar.r2022@vitstudent.ac.in
3. Department:
   SCOPE & SENSE

GenAI | Image Resizers

SAMSUNG

Problem Statement

Context

GenAI image generation models are good at generating images in trained resolution (For example 1024*1024). However, this limitation is not desirable in real life situation where we need to have images in different resolutions.
Another aspect is generally images generated are in square format which is not suitable for usage on mobile or laptop having rectangle format predominant.

Statement

Resizers and auto-upscalers (2X, 4X) for the images. Change to landscape and portrait.

Worklet Details

6	4	Gokul Nair +91-9496059290 gokul.gn@samsung.com
Duration (Months)	**Members Count**	Mentors

Pre-Requisite

- https://openmodeldb.info/
- https://paperswithcode.com/task/image-super-resolution

Expectations

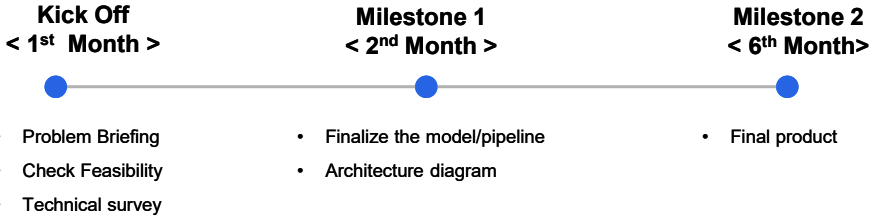
Undertaken Tasks

- Conduct Literature survey
- Identify the suitable framework
- Build a framework for image resizing and upscaling

KPI

- Web application with simple UI. ComfyUI is preferred.
- It should seamlessly integrate with the backend GenAI models. SDXL etc.
- Latency should be <10 seconds
- Original image contents should remain constant.
- No visible drop in image quality.

Timeline



Complexity

Work-let Name: **GenAI Image Resizers**

Worklet Details

1. Worklet ID: 24GAI15VITC
2. College Name: Vellore Institute of Technology Chennai

KPIs achieved till now

- Next.js Web application built
- Server with FastAPI endpoint
- Upscale latency less than than 10s
- No visual drop in image quality
- Original image contents remain constant

Issues faced

- Insufficient GPU Memory for Stable Diffusion XL

Next Steps

Key Achievements/ Outcome till now

- Identified best performing model.
- Built the frontend webapp
- Built server with models inference framework
- Seamless integration with FastAPI endpoint

.

DATE: 05-06-2025

Work-let Name: **GenAI Image Resizers**

Worklet Details

1. Worklet ID: 24GAI15VITC
2. College Name: Vellore Institute of Technology Chennai

Resource Requirement

- Nvidia GPU with atleast 12GB of VRAM
- 16 to 32GB of system memory

Github Repository and Prism Portal Status

- Samsung enterprise Github repository updated
- Prism portal updated
- Public repo link: Webapp and Server

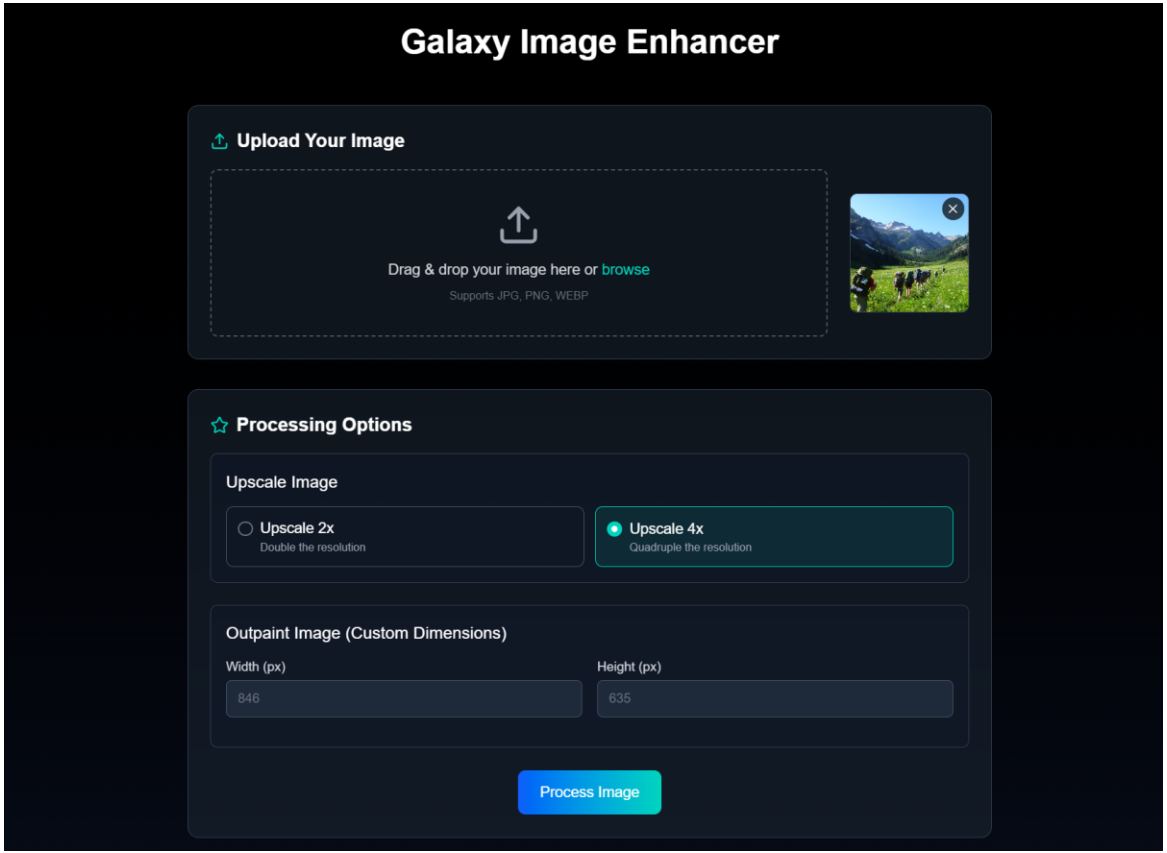
Anticipated Break

- None

Data collection source

- Not applicable
.

Frontend Webapp



- Next.js framework
- POST method to send Image and process option data to backend
- Form Data contains Image and Scale factor for upscaling or Target dimensions for outpainting

Image Upscaling Pipeline - RealESRGAN

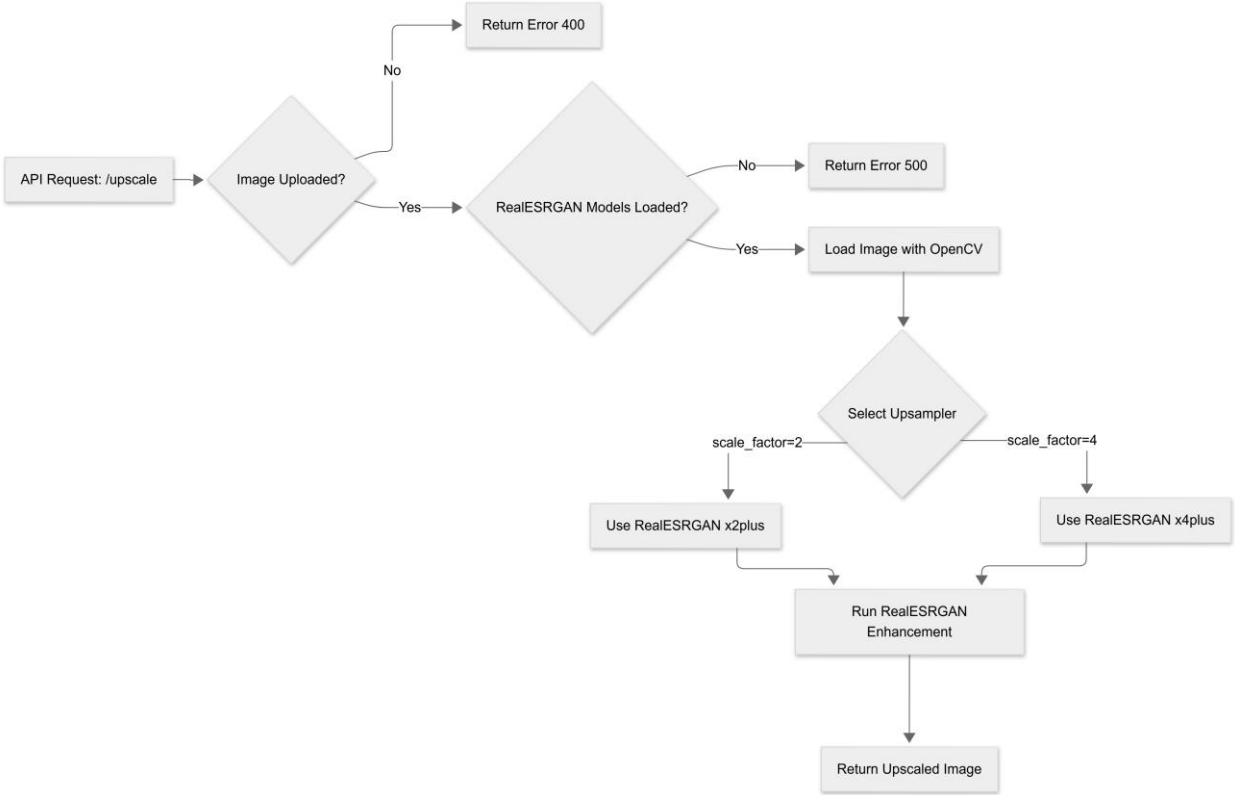
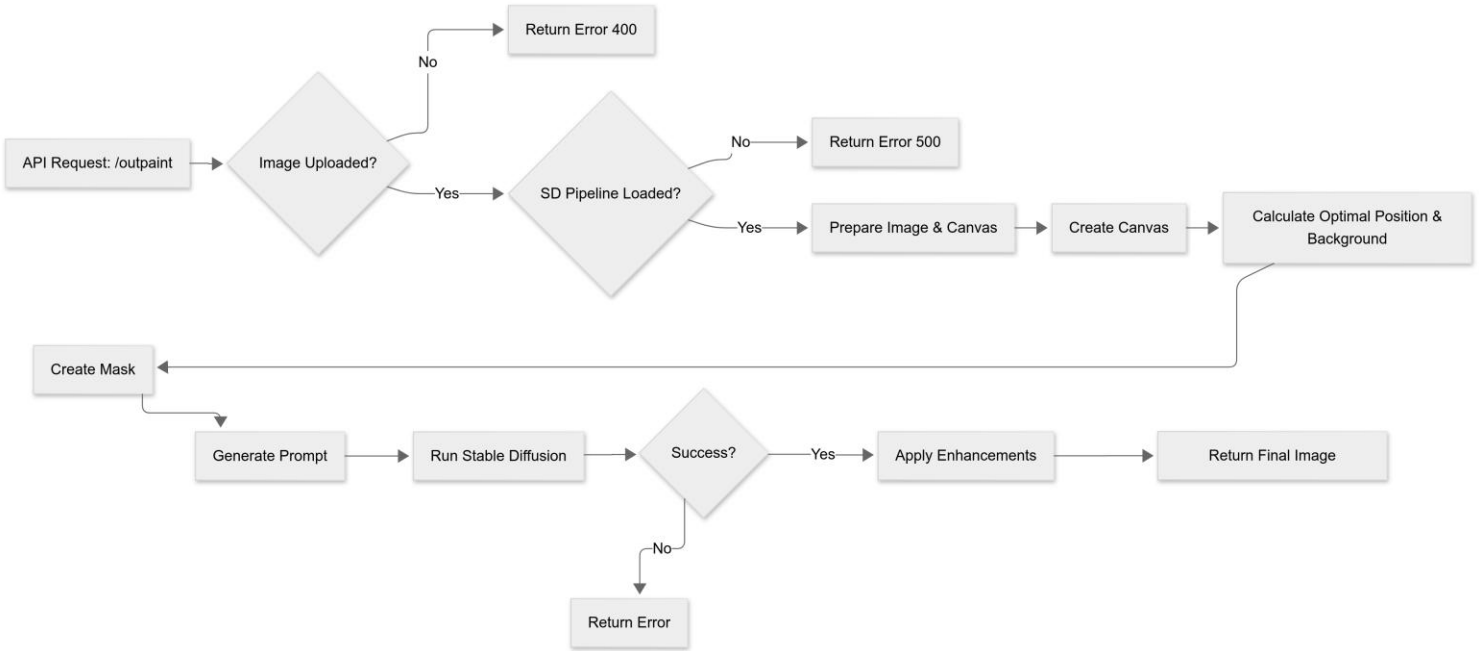


Image Outpainting Pipeline - StableDiffusion 2.0 Inpaint



Results

Process Option	Inference Time (RTX 4060)	VRAM Usage (Initialization)	VRAM Usage (Inference)
Upscale-2x	1-2 seconds	0.2GB	0.7GB
Upscale-4x	4-8 seconds		2.8GB
Outpainting	25-60 seconds	2.5GB	3.9GB + 0.6GB

- Inference time and VRAM usage for upscaling varies based on image resolution
- Inference time and VRAM usage for outpainting varies based on input image dimensions and target dimensions
- Cuda cache is cleared once inference is finished

Results

Original(1024x683)

Original - Cropped

2x Upscaled(2048x1366) - Cropped

4x Upscaled(4096x2732) - Cropped

Results

1024x683

1920x1080

846x635

1920x1200