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
1. Students:
   1. Rasika Phutane - rasika.suresh2022@vitstudent.ac.in
   2. Pojesh Kumar R - pojeshkumar.r2022@vitstudent.ac.in
   3. C R Lavanya - lavanya.2022a@vitstudent.ac.in
2. Department:
   SCOPE & SENSE

Date: 23 Sep 2024

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

Kick Off
< 1st Month >

- Problem Briefing
- Check Feasibility
- Technical survey

Milestone 1
< 2nd Month >

- Finalize the model/pipeline
- Architecture diagram

Milestone 2
< 6th Month>

- Final product

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

- Original image contents should remain constant.
- No visible drop in image quality.

Issues faced

- Artifact distortion and noise generation.

Next Steps

- Conversion landscape to portrait.
- Frontend application for upscaling and perform landscape to portrait conversion.

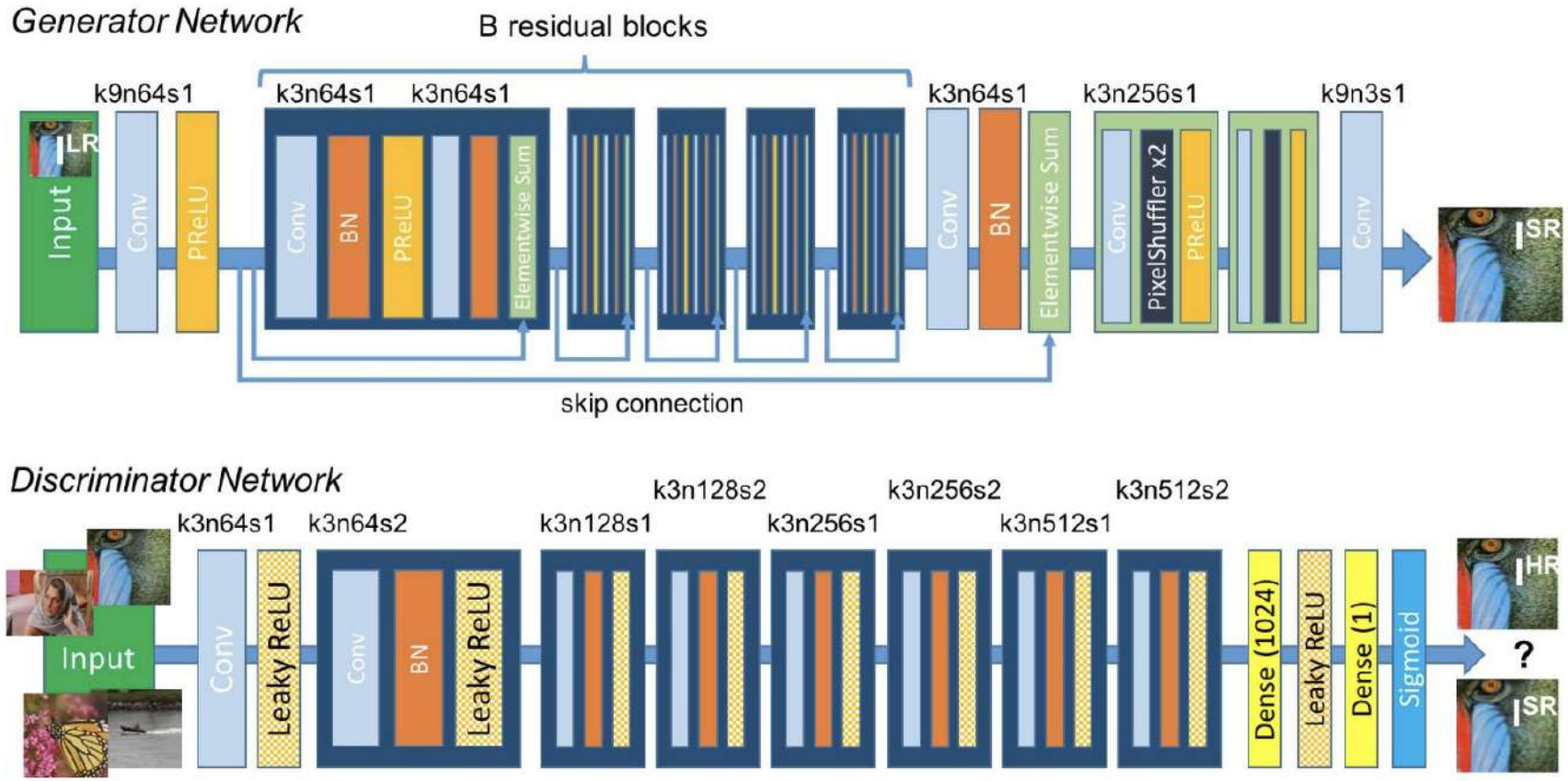
Key Achievements/ Outcome till now

- Identified best performing model.
- Successfully upscaled low-resolution images.

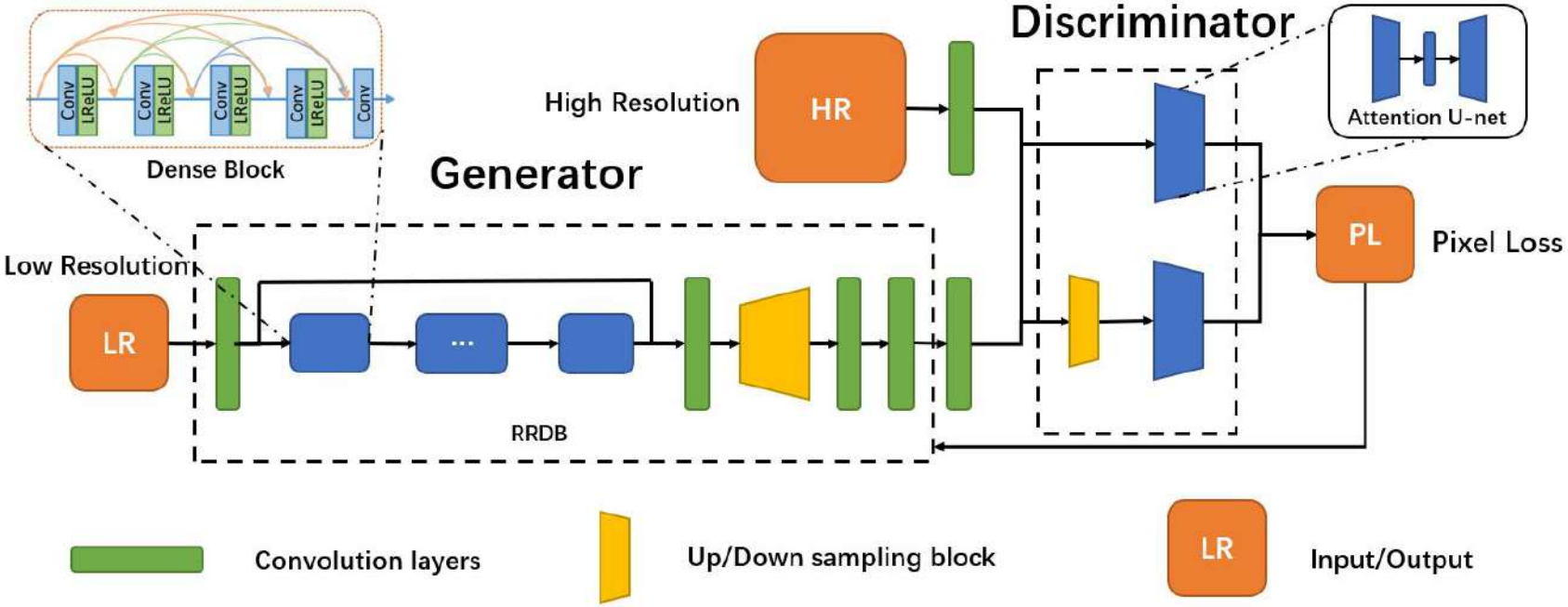
.

Date: 17-12-2024

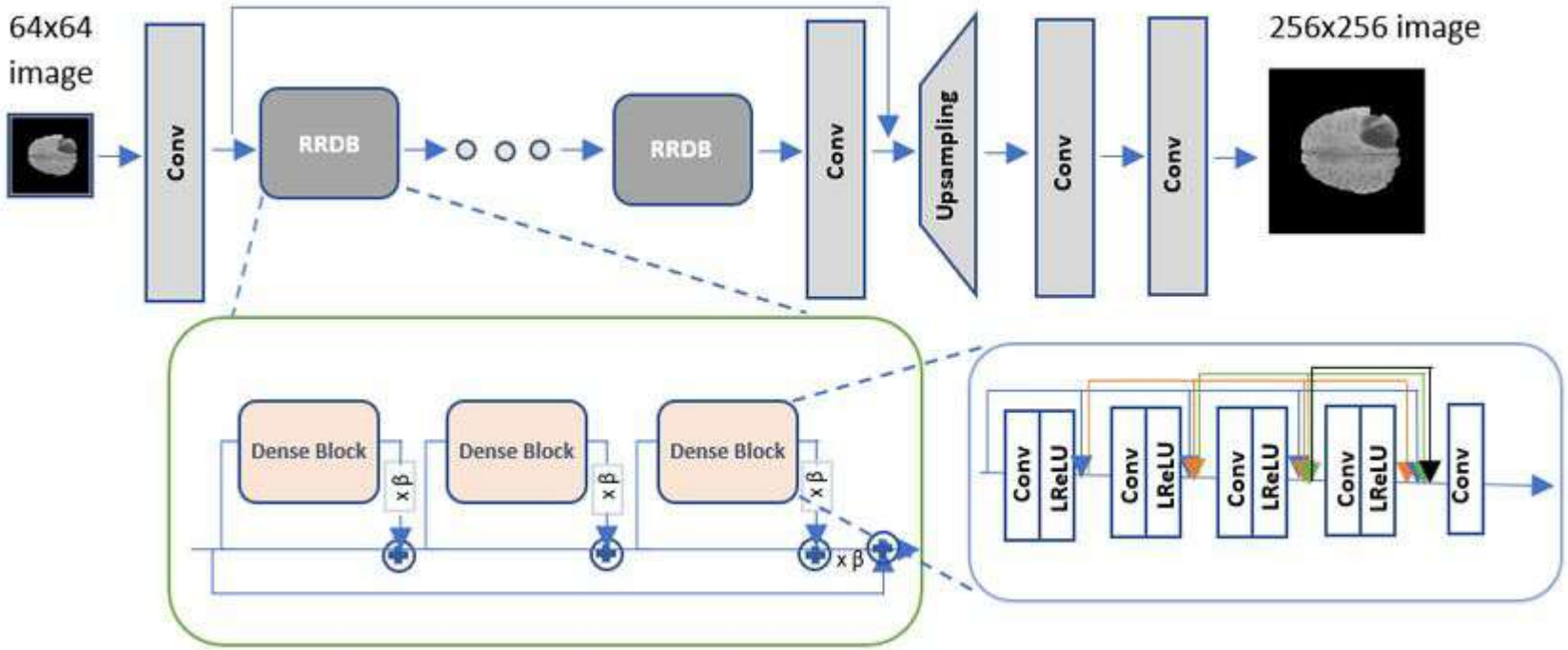
SRGAN - ARCHITECTURE



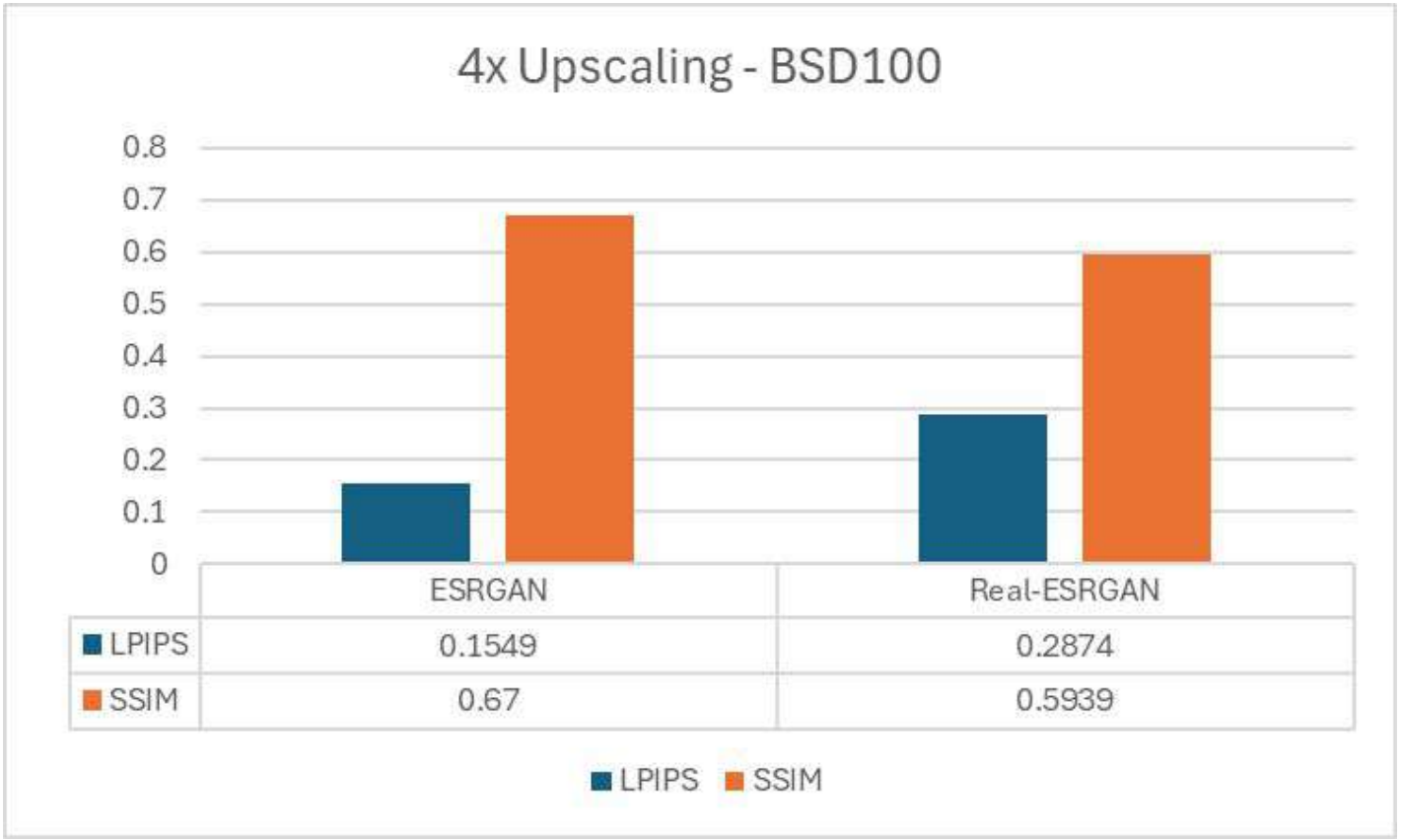
ESRGAN - ARCHITECTURE



REAL ESRGAN - ARCHITECTURE



Model Evaluation - BSD100 Dataset



Results of Image Upscaling

Original
481 x 321

Low resolution image
120 x 80

Real-ESRGAN
LPIPS : 0.3790
480 x 320

ESRGAN
LPIPS : 0.1960
480 x 320

Results of Image Upscaling

Original
1268 x 1132

Low resolution image
317 x 283

SRGAN
LPIPS : 0.21746
1268 x 1132

ESRGAN
LPIPS : 0.20872
1268 x 1132

Results of Image Upscaling

Original
1033 x 1026

Low resolution image
258 x 256

SRGAN
LPIPS : 0.30687
1033 x 1026

ESRGAN
LPIPS : 0.30019
1033 x 1026

Results of Image Upscaling

Original
1288x 1000

Low resolution image
322 x 250

SRGAN
LPIPS : 0.13159
1288 x 1000

ESRGAN
LPIPS : 0.12616
1288 x 1000

Results of Image Upscaling

Original
1276 x 1056

Low resolution image
319 x 264

SRGAN
LPIPS : 0.20463
1276 x 1056

ESRGAN
LPIPS : 0.18667
1276 x 1056

Results of Image Upscaling

Original
1240 x 772

Low resolution image
310 x 193

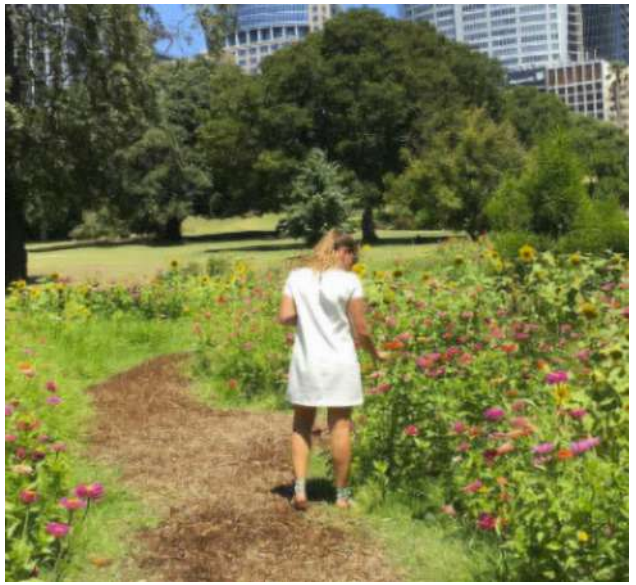
SRGAN
LPIPS : 0.26620
1240 x 772

ESRGAN
LPIPS : 0.28144
1240 x 772

Results of Image Upscaling

Original
1216 x 1124

Low resolution image
304 x 281



SRGAN
LPIPS : 0.20969
1216 x 1124

ESRGAN
LPIPS : 0.20233
1216 x 1124

Results of Image Upscaling

Original
1256 x 1032

Low resolution image
314 x 258

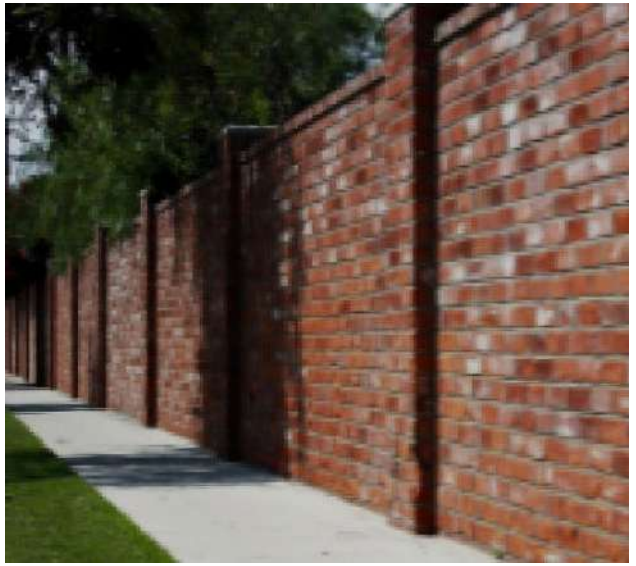
SRGAN
LPIPS : 0.19244
1256 x 1032

ESRGAN
LPIPS : 0.18088
1256 x 1032

Results of Image Upscaling

Original
1236 x 1108

Low resolution image
319 x 264



SRGAN
LPIPS : 0.25817
1236 x 1108

ESRGAN
LPIPS : 0.26188
1236 x 1108

Results of Image Upscaling

Original
848 x 1076

Low resolution image
212 x 269

SRGAN
LPIPS : 0.11214
848 x 1076

ESRGAN
LPIPS : 0.11125
848 x 1076

Results of Image Upscaling

Original
1344 x 892

Low resolution image
336 x 223

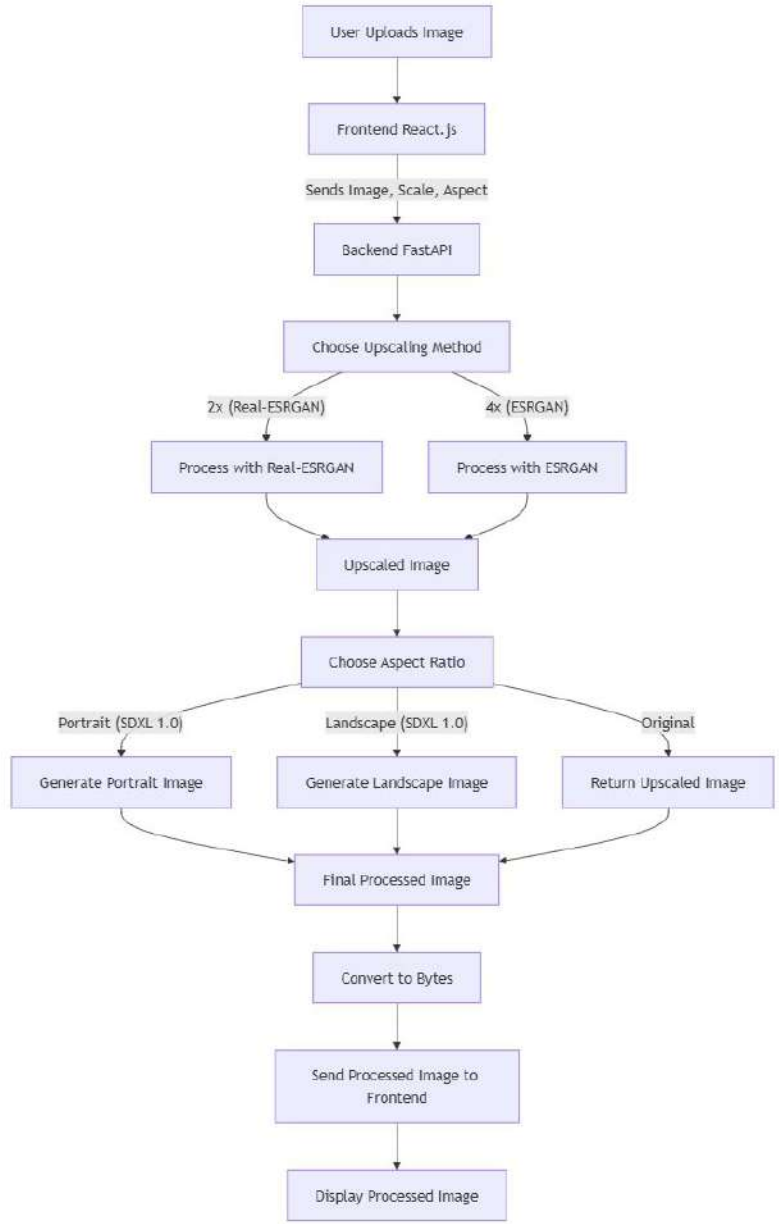
SRGAN
LPIPS : 0.23030
1344 x 892

ESRGAN
LPIPS :0.23656
1344 x 892

Results of Image Upscaling

Average LPIPS value of **SRGAN** - 0.212949

Average LPIPS value of **ESRGAN** - 0.209608


User Uploads Image
Frontend React.js
Sends Image, Scale, Aspect
Backend FastAPI
Choose Upscaling Method
2x (Real-ESRGAN)
4x (ESRGAN)
Process with Real-ESRGAN
Process with ESRGAN
Upscaled Image
Choose Aspect Ratio
Portrait (SDXL 1.0)
Landscape (SDXL 1.0)
Original
Generate Portrait Image
Generate Landscape Image
Return Upscaled Image
Final Processed Image
Convert to Bytes
Send Processed Image to Frontend
Display Processed Image

Academic Calendar Breaks

- **Mention any planned academic activity (Internals/Exams/Holidays/Events/Other Internships etc) in next 6 months which will impact project timeline.**
- **This helps in setting up the expectation of delivery timeline.**

1. 12/10/24 to 20/10/24 - Continuous Assessment 2
2. 18/11/24 to 22/11/24 - Final Assessment Test for LAB courses
3. 25/11/24 to 10/12/24 - Final Assessment Test for Theory courses
4. 22/12/24 to 5/01/25 - Winter Vacation
5. 25/01/25 to 1/02/25 - Continuous Assessment 1
6. 26/02/25 to 01/03/25 - Vibrance (College Event)