PixelForge: Al-Driven Digital Marketing System



ABSTRACT-

An innovative web application that uses advanced NLP and ML to automatically create high-quality image posts, captions, and video ads based on user prompts and product images. Unlike traditional methods, our system generates product-based video ads frame-by-frame, no scraping or stitching. This scalable, Al-driven tool streamlines digital ad creation and sets the stage for fully Al-generated campaigns.

METHODOLOGY

Image Captioning

- Used BLIP (Salesforce) for initial captions results were often vague or inaccurate.
- Improved outcomes via fine tuning and post-processing using Gemini API to generate market-ready slogans and engaging captions.

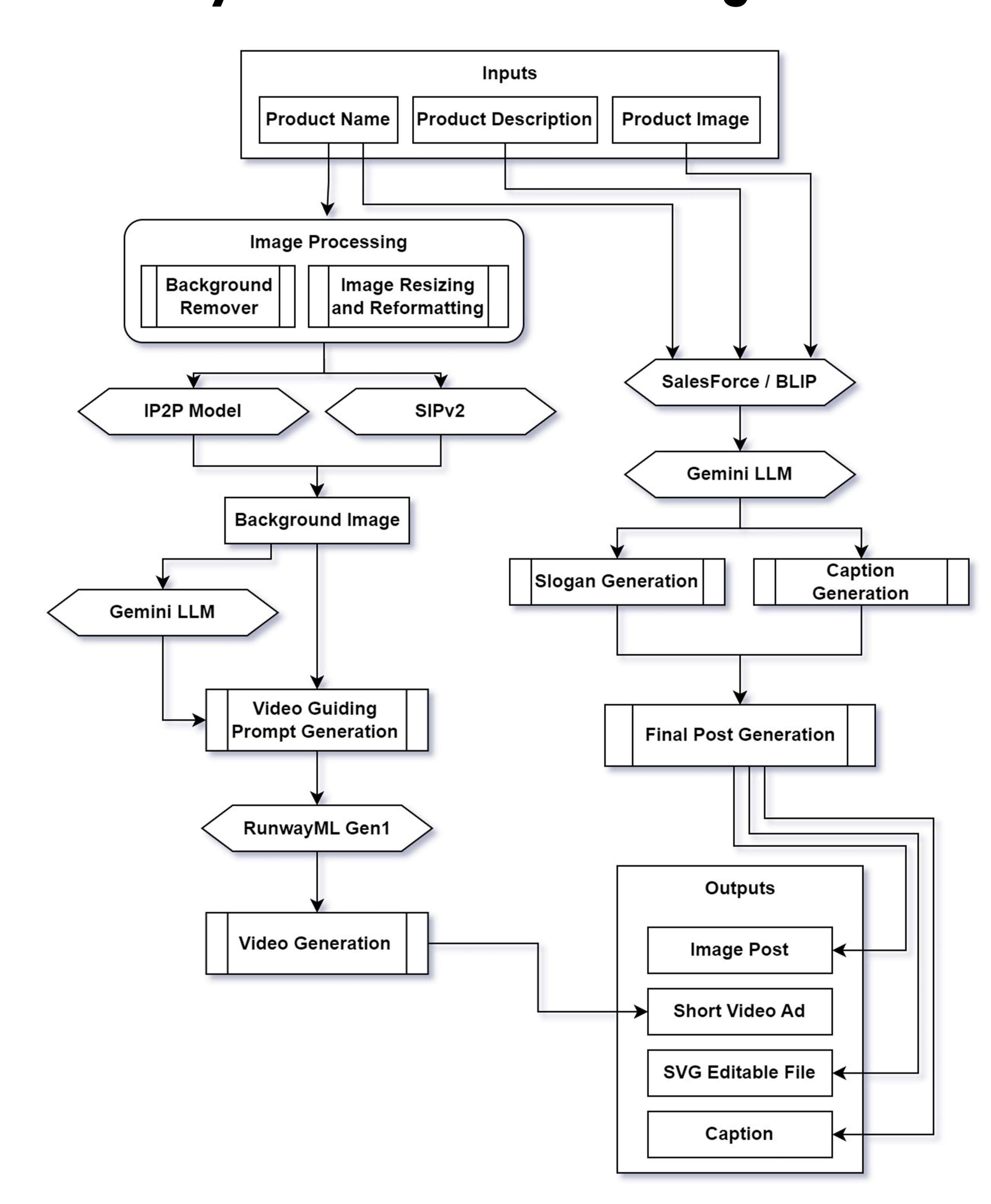
Image Generation

- Tested IP2P and SIPv2 for product to promotion image generation.
- Parameter tuning and transfer learning with a more specific dataset to overcome initial distortions.

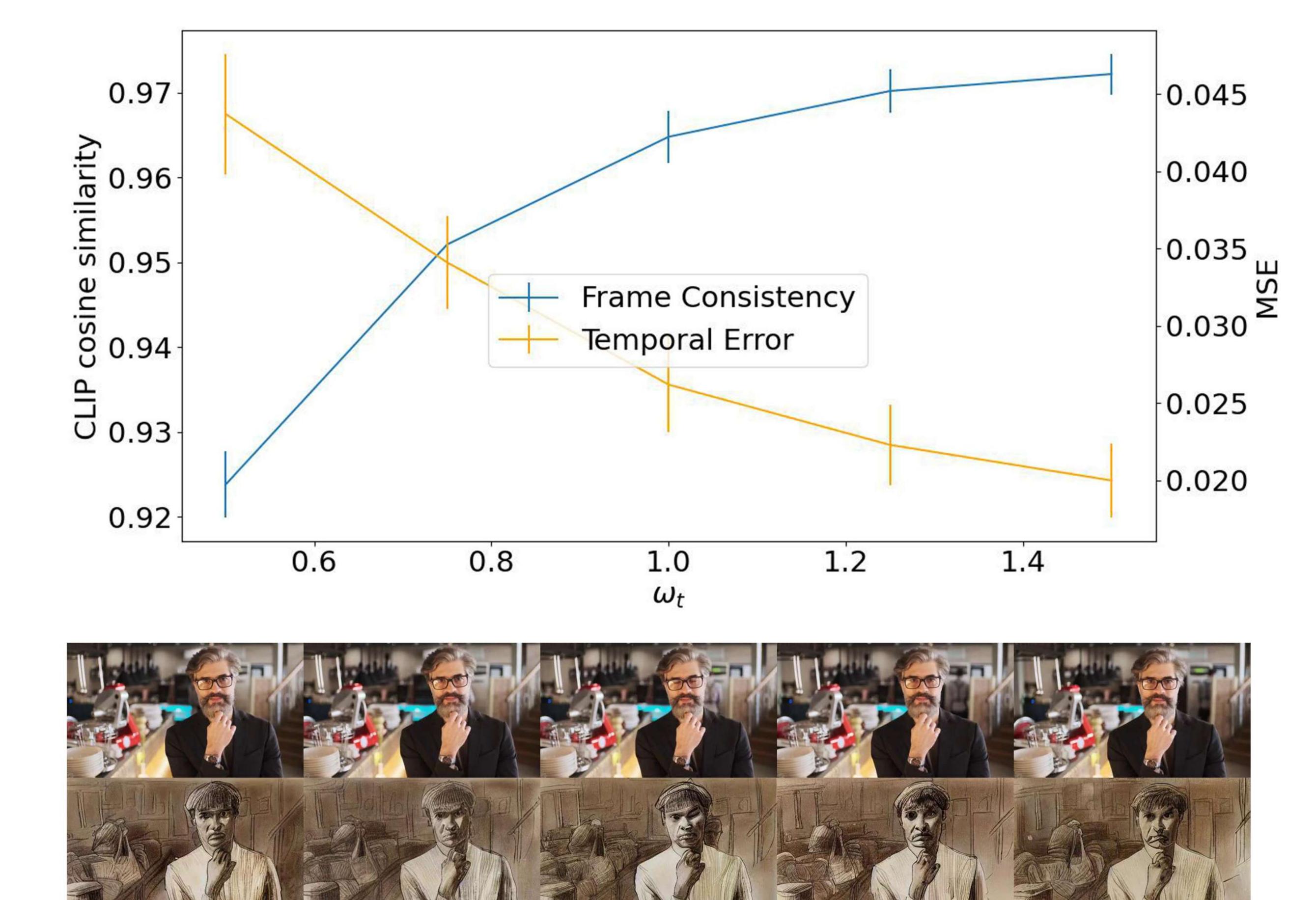
Video Generation

- Uses RunwayML Gen1 pretrained model.
- Our LLM genérates prompts from user uploaded images.
- Prompt guides model for proper video generation.
- Extreme Hardware Limitation (Multiple High End GPUs required) for fine tuning. Inference performed via Commmercial GPU.

System Architecture Diagram



RESULTS





	Model	Prompted Background		
		Empty	Likely	Unlikely
FID \	SI ₂	12.37	11.40	18.58
	Ours	6.44	7.12	10.50
LPIPS \	SI2	0.32	0.33	0.36
	Ours	0.32	0.30	0.33
CLIP Score ↑	SI2	-	25.89	29.01
	Ours		25.69	27.79
Obj. Sim. ↑	SI2	0.59	0.70	0.59
	Ours	0.65	0.72	0.62
Obj. Exp. ↓	SI2	0.117	0.104	0.102
	Ours	0.038	0.041	0.044

Comparison of SI2 and our model using prompts: (i) empty, (ii) likely setting, and (iii) unlikely setting. Metrics are averaged over all datasets.

Group Members

Khawaja Rabeet Tanveer (21K-3374)
Muhammad Shaheer Luqman (21K-4655)
Ismail Qayyum (21K-3188)

Supervisors

Mr. Basit Jasani Dr. Muhammad Rafi