Project Idea: Perceptual Enhancement for Social Media Videos

Project Title:

“One-Tap Perceptual Enhancement for User-Generated Mobile Videos”

Abstract:

The project aims to create a **“magic-enhance”** feature within a simple mobile video editor. This feature will use a single, efficient AI model to improve the perceptual quality of user-generated content (e.g., Reels, YouTube Shorts) by simultaneously performing denoising, deblurring, and super-resolution, optimized for visual appeal rather than pure technical accuracy.

Models & Approach:

- Use a pre-trained GAN-based model known for producing visually pleasing results, such as Real-ESRGAN or a similar lightweight alternative.
- The focus is not to re-implement the model from scratch, but to optimize its pre-trained weights for mobile deployment using TensorFlow Lite’s quantization and pruning tools.
- The system will act as a post-processing tool, similar to apps like Remini, but specifically designed for video enhancement.

Hardware & Tools:

- **Development:** Laptop
- **Deployment:** Android Phone
- **Software:** Python, TensorFlow Lite, Android Studio

Existing Solutions

Currently, there are several AI-based applications and research models that enhance images and videos such as **Remini**, **Topaz Video Enhance AI**, and **Real-ESRGAN**.
These tools mainly focus on improving image or video resolution (Super-Resolution) or removing noise and blur individually. However, most of these existing systems:

- Require **high-end GPUs** or **cloud servers** to run efficiently.

- Are **not optimized** for real-time enhancement on **mobile devices**.
- Focus on **technical metrics** like PSNR or SSIM rather than **human perceptual quality** (how natural and visually appealing the video looks).

Gap Identification

While current tools provide good technical improvements, they are **not directly usable for casual mobile users** who create short videos (like Reels or TikToks).
There is a **lack of a lightweight, mobile-optimized solution** that can:

- Perform **multiple enhancements (denoising, deblurring, super-resolution)** together in one step.
- Focus on **perceptual quality** making videos look clearer, smoother, and more visually pleasing to the human eye.
- Run **efficiently on standard smartphones** without the need for powerful hardware.

Our New Contribution

Our project aims to fill this gap by developing a **"One-Tap Perceptual Enhancement"** feature for user-generated mobile videos.
Key contributions include:

- **Mobile Optimization:** We take a state-of-the-art model (like Real-ESRGAN) and **optimize it using TensorFlow Lite** through pruning and quantization, so it can run efficiently on mobile phones.
- **All-in-One Enhancement:** Instead of handling denoising, deblurring, and super-resolution separately, our approach performs **all enhancements in a single step**.
- **Perceptual Focus:** We prioritize **visual appeal and user experience** over pure technical accuracy ensuring videos look naturally enhanced and engaging for social media.
- **Practical Deployment:** Unlike most research models limited to high-end hardware, our solution demonstrates **real-time enhancement directly on an Android phone**, making professional-level quality **accessible to everyday users**.

Project Development Approach

Phase 1: Standalone Implementation

- In the initial phase, we will build a **standalone desktop version** of the enhancement system.
- We'll use a **pre-trained GAN-based model** (like **Real-ESRGAN**) to enhance video quality by performing **denoising, deblurring, and super-resolution**.
- This phase helps us:
    - Understand how the model works and measure its enhancement performance.
    - Fine-tune or test perceptual quality improvements on sample videos.
    - Verify that the model produces visually appealing results before mobile optimization.

Phase 2: Model Optimization and Mobile Deployment

- After validating results on the standalone setup, we'll **optimize the model** using **TensorFlow Lite** (through pruning and quantization).
- Then, we'll **integrate the optimized model** into a simple **Android app** (or Flutter-based app) with a "**One-Tap Enhance**" feature for user videos.