Project Design

```
CS 7180 Advanced Perception
```

Title of the project:

> ## Watermark Aging of Images from Latent Diffusion Models

Group members:

Shreyas Prasad

Problem or hypothesis:

The internet age has enabled easy replication and distribution of multimedia content, making it challenging to verify the originality of any piece. Watermarking techniques, like the Stable Signature [1], aim to embed invisible markers in content to track and verify its origin. However, as images undergo transformations, filters, or edits, the watermark may degrade. The concept of "Watermark Aging" delves into this phenomenon of watermark decay over time and modifications.

Watermark Aging is Crucial is because of Content Life Cycle on the Internet: Multimedia content over the internet (Social Media), especially images, do not remain static. They undergo several edits, modifications, and repurposing. Over this life cycle, it's crucial to understand how a watermark behaves.

Although the paper evaluates robustness of the watermarking process on the transformations like cropping, rotation, coloring, and a combination of these. We are mainly concerned about these two types of **successive transformations.** Hence we aim to evaluating watermark resilience in:

1. **Compression Sensitivity Analysis**: Understanding how a watermark reacts to different compression levels and algorithms is critical. Both lossless (e.g., PNG) and lossy (e.g., JPEG) compressions should be tested.
2. **Format Transition Impact**: Study the watermark's integrity when images are converted across popular formats like JPEG, PNG, GIF, BMP, and WebP.

3. **Sequential Modifications**: Replicate the internet's real-world scenario by subjecting watermarked images to successive compressions and format conversions, mimicking an image's journey across various platforms.

Novel aspect of the work:

Prior work has focused on singular and combined transformation effects on watermarks, but the compounded effect of repeated transformations, which an image undergoes in its digital journey, has been unexplored. Furthermore, this work dives into the nuances of internet-specific alterations, like platform-driven compressions and user-initiated format conversions.

Data requirements and sources:

1. Pre-Trained Models [2]
2. COCO Dataset [3]

Computational requirements and resources:

- SSDs of at least 1TB
- Multi-GPU setups, preferably NVIDIA RTX 4090 or better
- PyTorch, OpenCV

Expected products or results:

Watermark Decay Index: A metric quantifying the level of watermark degradation relative to the number of successive and type of transformations.

References

1. Pierre Fernandez, Guillaume Couairon, Hervé Jégou, Matthijs Douze, Teddy Furon. *The Stable Signature: Rooting Watermarks in Latent Diffusion Models*. ICCV 2023. Mar, 2023.
2. https://github.com/facebookresearch/stable_signature
3. https://cocodataset.org/#home