Week 1 Plan: Foundations and Preparation

Objectives:

- 1. **Understand the Basics of Image Captioning**: Learn the principles behind generating captions for images.
- 2. **Familiarize Yourself with Relevant Technologies**: Gain knowledge of CNNs, transformers, and attention mechanisms.

Resources and Activities:

- 1. Readings and Research
 - Vision Transformers:
 - https://arxiv.org/abs/1706.03762 this paper introduces transformers
 - An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale - The paper that introduces Vision Transformers (ViTs) and discusses their applications in image processing.
 - Introduction to Image Captioning:
 - https://ieeexplore.ieee.org/document/9726074 A comprehensive survey of image captioning techniques
 - https://ieeexplore.ieee.org/document/10099161
 - https://arxiv.org/abs/1502.03044 this paper Combines visual features and attention mechanisms to generate captions

2. Basics on CNNS and LSTM

- https://youtube.com/playlist?list=PLuhqtP7jdD8CD6rOWy20INGM44kULvrHu&si=F B9SPZkMhhJwNsI0
- https://youtu.be/YCzL96nL7j0?si=SwQamx8LaQrAmoyg