Gradescope Overview

- (2pts) Problem 1
 - Visualize the similarity between different pairs of positional embedding and briefly explain the result.
 - Additionally, attach the code that you used for visualization.
- (2pts) Problem 2
 - Clip gradient norm and visualize the changes of gradient norm in different steps. Circle two places with gradient explosion.