Pairwise Cosine Similarity Between ViT Models
(Same Architecture, Different Training)