Embeddings of clean/poisoned images after denfense using average

