Abstract

Deepfake technologies are a form of artificial intelligence (AI) which are based on generative adversarial networks (GANs), a development which has emerged out of deep learning (DL) and machine learning (ML) models. Using a data range which spans the years 2018 – 2021, this research explores public perceptions of deepfake technologies at scale by closely examining commentary found on the social video-sharing platform, YouTube. This open source, ground-level data documents civilian responses to a selection of user-produced, labelled deepfake content. This research fills a gap regarding public perception of this emerging technology at scale. It gauges an underrepresented set of responses in discourse to find that users demonstrate a spectrum of responses which veer between irony and concern, with greater volumes of commentary skewed towards the former. This study of user commentary also finds that YouTube as a wild space ultimately affords reflexive and critical thinking around the subject of deepfake technologies and could prove to be effective as a form of inoculation against disinformation.