Title: Manual validation of DeepFace demographic and emotion predictions appendix for "Toxic politics and TikTok engagement in the 2024 U.S. election"
Authors: Ahana Biswas (1), Alireza Javadian Sabet (1), Yu-Ru Lin (1)
Date: August 20th, 2025
Note: The material contained herein is supplementary to the article named in the title and published in the Harvard Kennedy School (HKS) Misinformation Review.

Appendix G: Manual validation of DeepFace demographic and emotion predictions

To evaluate the accuracy of DeepFace on our dataset, we conducted a manual validation of its predictions for gender, ethnicity, age group, and facial emotion. We randomly selected one keyframe from each of 50 randomly sampled TikTok videos, and a single human annotator labeled the demographic and emotional attributes visible in the selected frame. Given the visual clarity of the tasks and the binary or coarse-grained nature of the labels, a single coder was used for this evaluation. For gender (male/female) and ethnicity (white/non-white), DeepFace achieved an accuracy of 84.4% and 95.5% respectively. For age and dominant emotion, the proportion of correct predictions was 78% and 74%, respectively. When no face was detected, DeepFace returned null values. These results suggest that DeepFace performs reliably on the types of political TikTok videos in our dataset and is suitable for extracting high-level demographic and affective trends at scale.