[Abstract]

According to one aspect of the present invention, there is provided a system for detecting a deep voice generated by generative artificial intelligence, the system comprising: a memory configured to store instructions; and one or more processors configured to execute the instructions, wherein the one or more processors are configured to: extract, from input voice data, a feature vector including a frequency characteristic and a temporal change pattern; input the extracted feature vector into an ensemble artificial intelligence model including a first AI model that is an acoustic model, a second AI model that is a language model, and a third AI model that is a pattern recognition model, to calculate an individual deep voice probability for each model; measure a signal quality of the input voice data to calculate a signal-to-noise ratio (SNR); calculate a final deep voice probability through an adaptive weight allocation algorithm that varies, in real time, a weight to be applied to each AI model based on the calculated SNR value; and generate a deep-voice-suspicion alert when the final deep voice probability exceeds a threshold value.