

HOME > [TECH & INNOVATION](#)

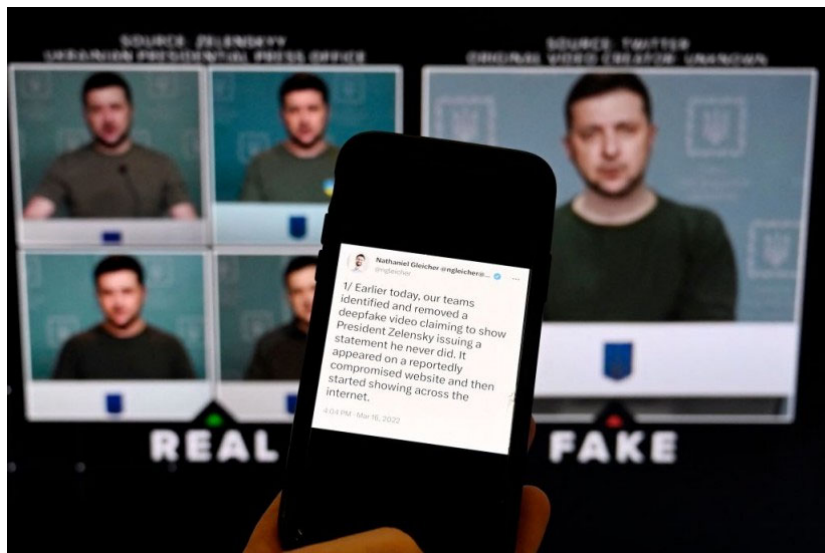
AI Program Generates Realistic Facial Animations from Audio and Face Photos

Margaret Davis Nov 17, 2023 09:07 AM EST



Researchers from Nanyang Technological University (NTU) in Singapore unveiled an advanced computer program called [Diverse yet Realistic Facial Animations \(DIRFA\)](#), crafting lifelike videos with facial expressions and head movements.

To ensure realistic and synchronized facial animations with spoken audio, the innovative artificial intelligence (AI) tool seamlessly generates 3D videos by utilizing just an audio clip and a face photo.



(Photo : OLIVIER DOULIERY/AFP via Getty Images)

This illustration photo taken on January 30, 2023 shows a phone screen displaying a statement from the head of security policy at META with a fake video (R) of Ukrainian President Volodymyr Zelensky calling on his soldiers to lay down their weapons shown in the background, in Washington, DC.

Training DIRFA To Make Realistic Talking Faces

Generating realistic facial expressions from audio signals presents a complex challenge due to the multitude of possibilities for a given signal, especially when dealing with temporal sequences. The team's objective was to develop talking faces that exhibit precise lip synchronization, rich facial expressions, and natural head movements corresponding to the provided audio.

While audio has strong associations with lip movements, its connections with facial expressions and head positions are weaker, prompting the creation of DIRFA—an AI model designed to capture intricate relationships between audio signals and facial animations.

MOST POPULAR

1 [12-Second Loud Boom Heard Over New England; Harvard Astrophysicist Believes It Could Come From UFO](#)



2 [Beachgoers Stunned as the World's Most Dangerous Bird Makes a Splash on Australia's East Coast on Halloween](#)



3 [Iron Age Ivory Elephant Tusk Found in Turkey Features Depiction of Sphinx](#)



4 [Novel Robotic System Can Assess Post-Stroke Mobility For Those Recovering From the Condition, Offer Objective Data For Rehabilitation](#)



5 [Low Testosterone Levels Linked to Increased Arthritis Risk, Study Reveals](#)



LATEST STORIES

[Robotics Bring Back Extinct Species Not With DNA But Engineering. \[Study\]](#)



[Autoimmune Disorder Hashimoto's Disease: Patients Face Challenges in Care, Bills, Diagnosis](#)



[Some Sub-Neptune Planets Shrink Down to Super-Earth \[Study\]](#)



[Hippo Descendants From Drug Kingpin Pablo Escobar's Pets Are Now Undergoing Sterilization in Colombia Amidst the Growth of Invasive Species](#)



[AI-Powered Robot Chemist From China Could Generate Oxygen in Mars](#)



Subscribe to The Science Times!

Sign up for our free newsletter for the Latest coverage!

Email Address

Sign Up

than 6,000 individuals, sourced from a publicly available database. The model was engineered to predict the likelihood of specific facial animations, like raised eyebrows or wrinkled noses, based on input audio.

This approach allows DIRFA to transform audio input into diverse, highly realistic sequences of facial animations, guiding the generation of talking faces with accurate lip movements, expressive facial features, and natural head poses.

According to Assoc. Prof. Lu, DIRFA's modeling capabilities enable the program to generate talking faces that authentically synchronize with audio, offering accurate lip movements and natural expressions. The team is actively refining DIRFA's interface to provide users with more control over specific outputs.

Notably, users will soon be able to adjust certain expressions, such as transforming a frown into a smile. Additionally, the researchers are committed to further enhancing DIRFA's capabilities by incorporating a wider range of datasets, ensuring it accommodates diverse facial expressions and voice audio clips for more comprehensive and versatile applications.

READ ALSO: [Humanoid Robot Ameca Is 'Sad' After Realizing It Will Never Experience True Love \[WATCH\]](#)

DIRFA Program Enhances Facial Animations with AI and Machine Learning

The developed DIRFA program from NTU addresses challenges in existing methods related to pose variations and emotional control. Researchers envision DIRFA having diverse applications, including healthcare, where it could enhance virtual assistants and chatbots, providing more sophisticated and realistic user experiences.

Additionally, DIRFA could be a valuable tool for individuals with speech or facial disabilities, enabling them to communicate thoughts and emotions through expressive avatars or digital representations.

Corresponding author Associate Professor Lu Shijian, the study's lead from NTU's School of Computer Science and Engineering, emphasized the profound impact of their research on multimedia communication.

DIRFA, incorporating AI and machine learning techniques, enables the creation of highly realistic videos featuring accurate lip movements, vivid facial expressions, and natural head poses, using only audio recordings and static images.

First author Dr. Wu Rongliang, a Ph.D. graduate from NTU's SCSE, highlighted the program's pioneering efforts in enhancing AI and machine learning performance, particularly in audio representation learning. The [findings](#) were published in the scientific journal *Pattern Recognition* in August.