

Hannah Babe and
Alessandra Vennema
AIT-Budapest

Spring 2022, Deep Learning
Final Presentations

Audio Spoof Detection



Introduction

Background: With systems being increasingly vulnerable to scams., spoof detection applications are an important tool to protect consumer privacy and security

Goal: Implement a model that can identify when a piece of audio is fake and/or manipulated

Deepfake Audio Scores \$35M in Corporate Heist

A combination of business email compromise and deepfake audio led a branch manager to transfer millions to scammers, in a case that serves as a warning to organizations.

**Robert Lemos**

Contributing Writer

October 20, 2021



Security News

> Cyber Attacks

> Unusual CEO Fraud via Deepfake Audio Steals US\$243,000 From UK Company

Unusual CEO Fraud via Deepfake Audio Steals US\$243,000 From UK Company

September 05, 2019

An unusual case of CEO fraud used a deepfake audio, an artificial intelligence (AI)-generated audio, and was [reported](#) to have conned US\$243,000 from a U.K.-based energy company. According to a report from the [Wall Street Journal](#), in March, the fraudsters used a voice-generating AI software to mimic the voice of the chief executive of the company's Germany-based parent company to facilitate an illegal fund transfer.

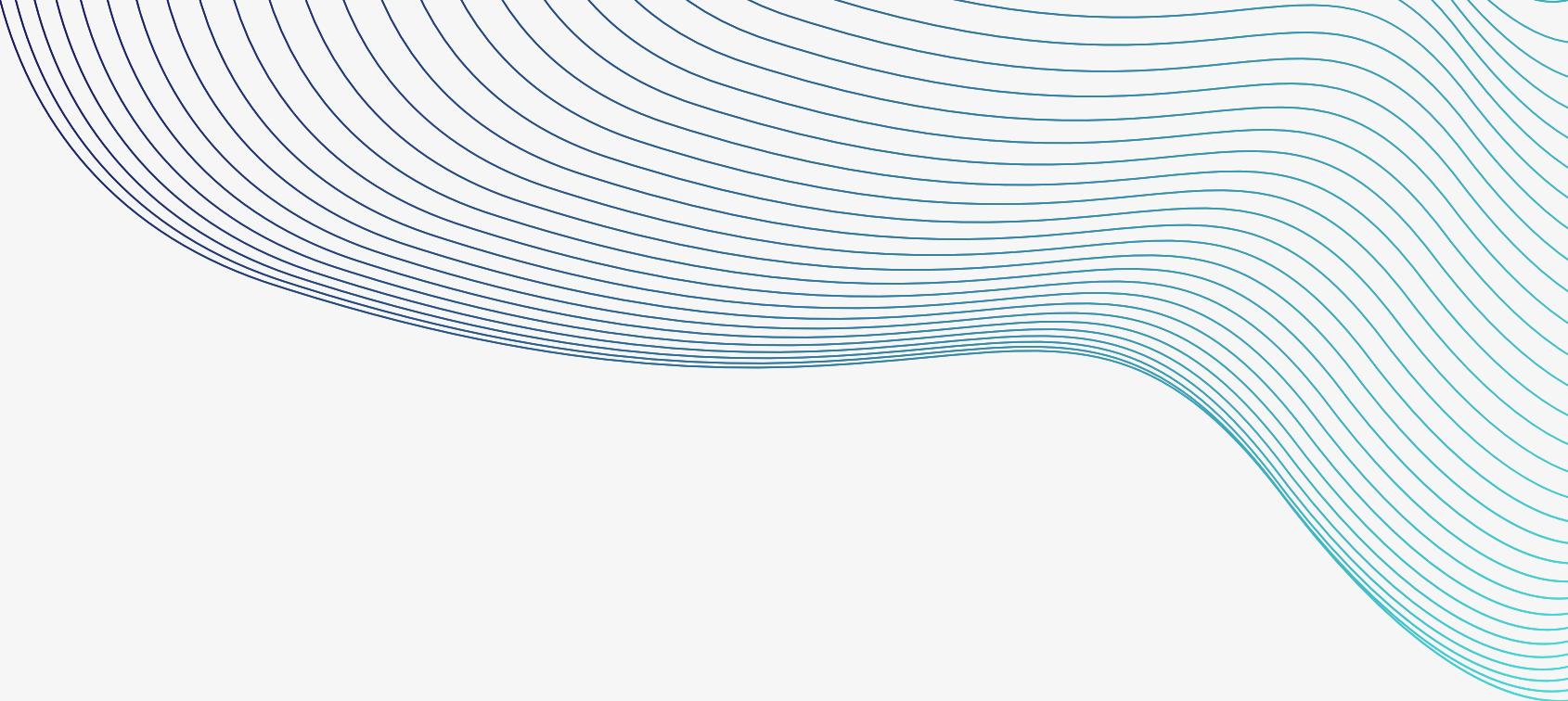
The cybercriminals called the U.K. company's CEO pretending to be the CEO of the parent company. The attackers demanded that an urgent wire transfer be made to a Hungary-based supplier and the U.K. company's CEO was assured of a reimbursement. After the money had been transferred, it was forwarded to an account in Mexico and then other locations, making the



Related Posts

Recent Post

[Using Custom Co](#)[Serverless Enviro](#)[Security](#)[Abusing Argo CD](#)[Artifact Hub: An A](#)[Chain Attacks in A](#)[pplications](#)[Crafting an Azure](#)[Threat Model](#)



Can you guess the
real audio?

Import Text ▾



← Spoof Detection

Download Speech

Voice: Aria · Female · English (US) ▾

Voice style: Regular News (formal) News (casual) Customer Service Chat Cheerful Empathetic

Speed: Default ▾

12 words • 0:05 est.



Hi honey, can you send me a picture of the Chase card



Listen

Pronunciation

Tone

Pause ▾

Tone

Change the voice tone

Volume

x-soft

soft

medium

loud

x-loud

or



0

dB

Rate

x-slow

slow

medium

fast

x-fast

or



100

%

Pitch

x-low

low

medium

high

x-high

or



0

%



Preview

Clear

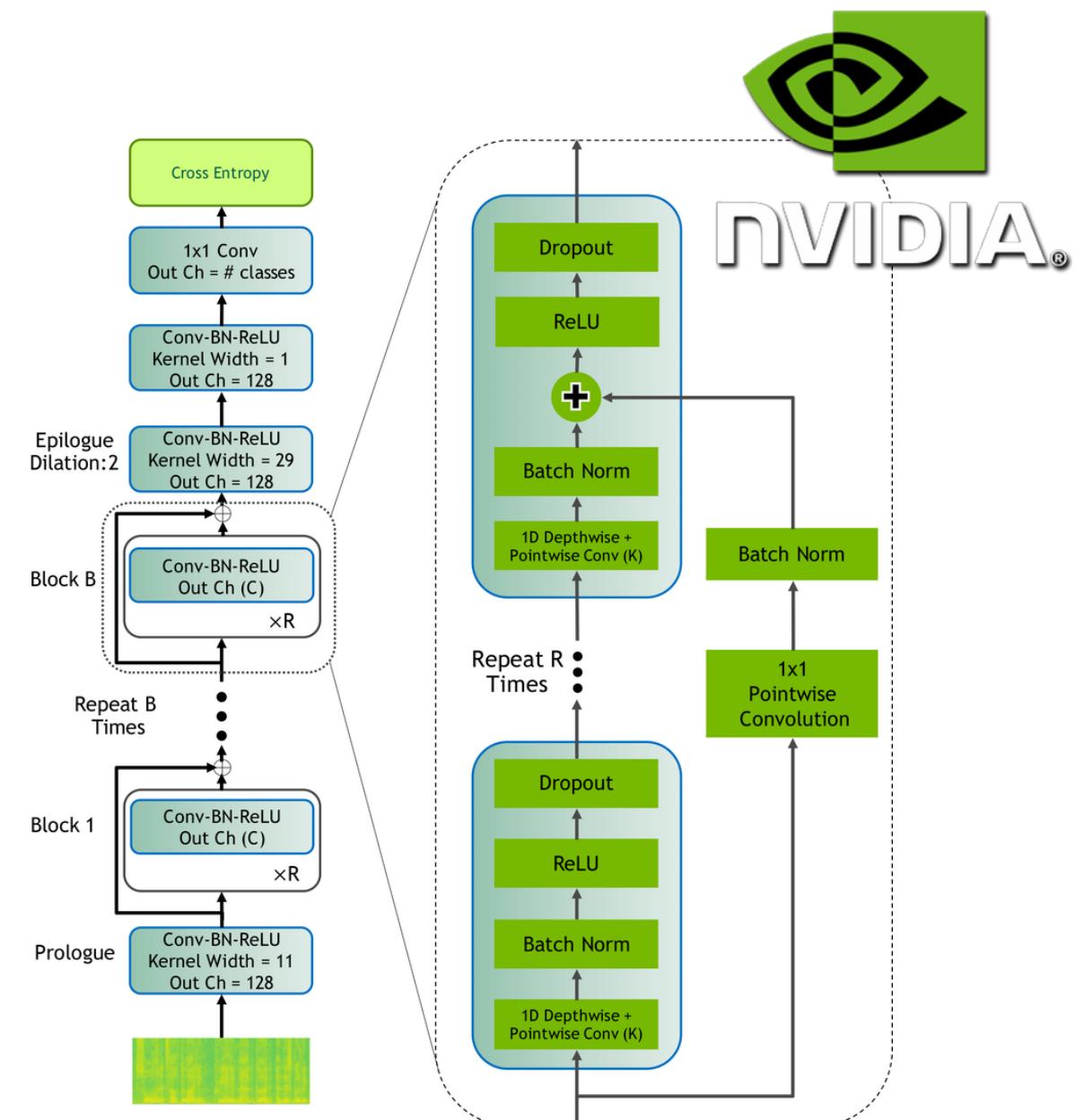
Apply



Source: <https://play.ht/text-to-speech-online/>

Resources

The screenshot shows a blog post on Towards Data Science. The title is "Audio Deep Learning Made Simple: Sound Classification, Step-by-Step" by Ketan Doshi. The post is published on Mar 18, 2021, and has a reading time of 12 minutes. It is part of the "HANDS-ON TUTORIALS, INTUITIVE AUDIO DEEP LEARNING SERIES". The post content includes a yellow background image of a woman dancing while listening to music. At the bottom, there are navigation icons for home, search, and profile.



SASV challenge 2022

Chosen Method

Conv2D

Recognize patterns

Batch Normalization

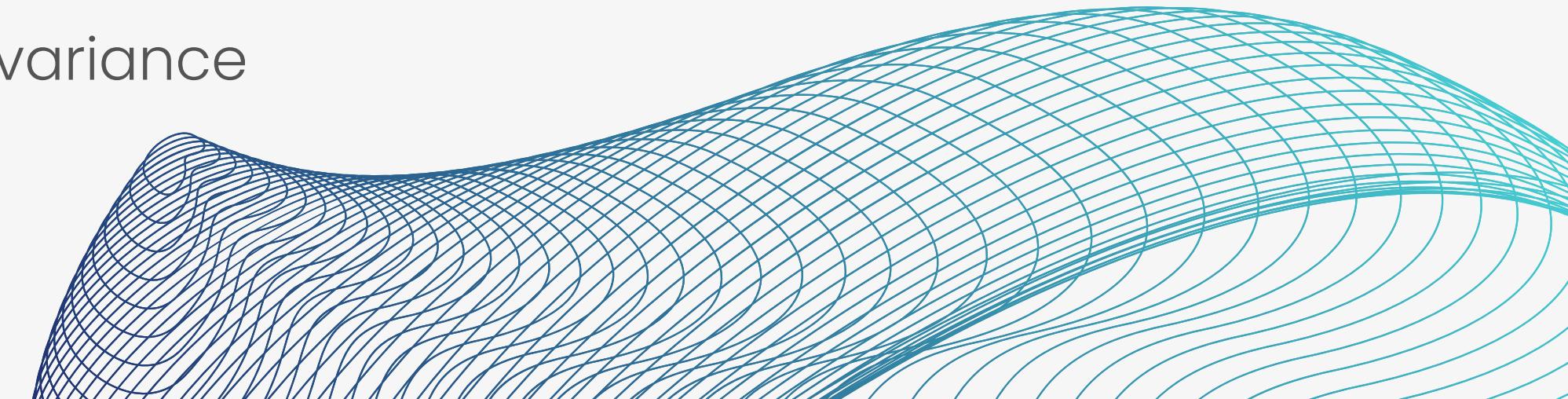
accelerates the training of deep neural nets

Dropout

prevent co-adaptation, where the neural network becomes too reliant on particular connection

Max Pooling

adds a small amount of translation invariance



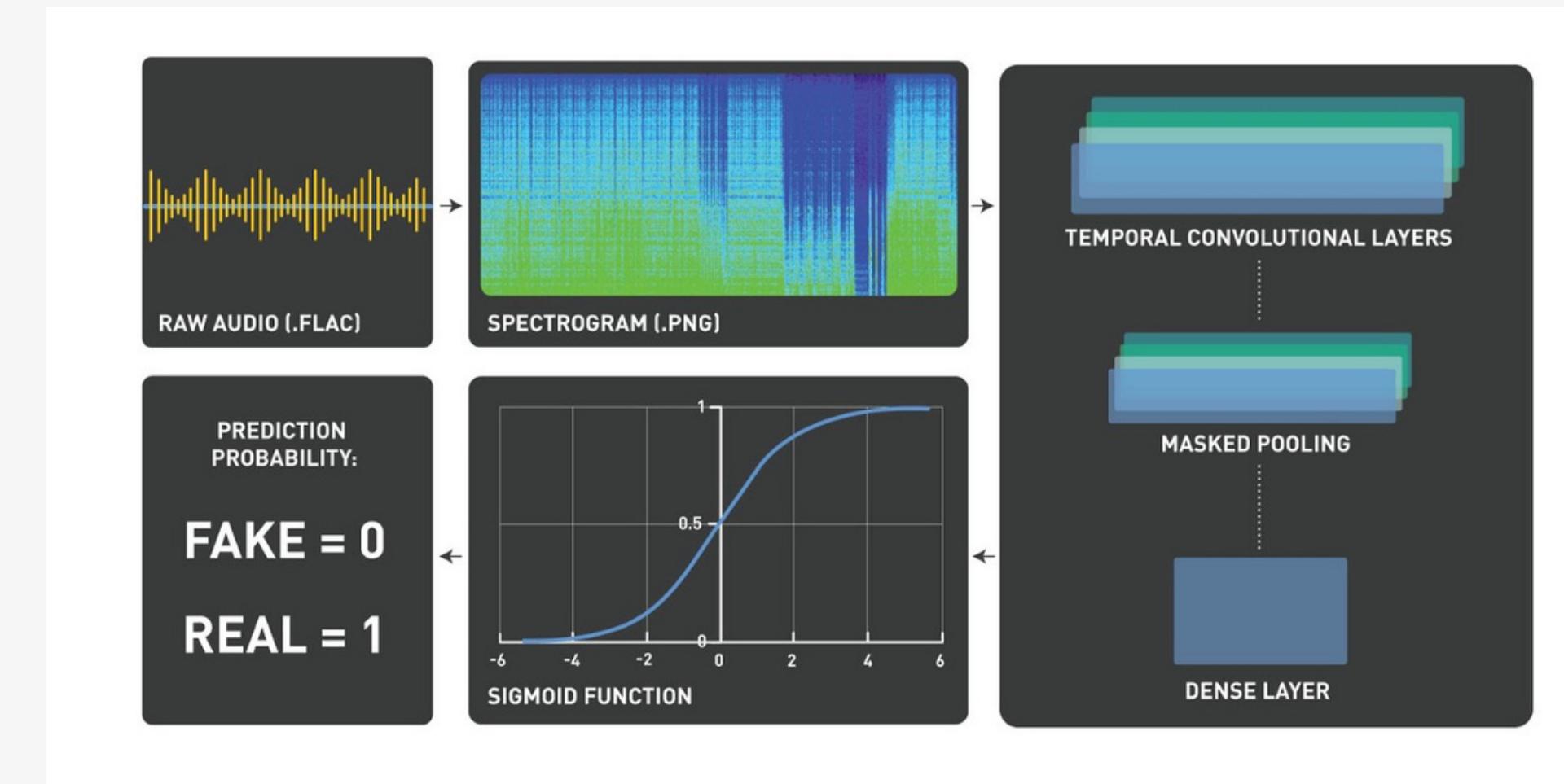
Results

- Initial accuracy of 99% and loss of 0.0348.
- On evaluation dataset 82% accuracy and loss of 4.88.



Discussion

- Increase model accuracy, potentially by using a pre-trained model.
- Usability of solution for real-time results.





Thank you for Listening
Questions?