AI Anarchies

# Introduction to Adversarial Acoustics

# Introduction

**Murad Khan**
UAL: Creative Computing Institute

**Martin Disley**
University of Edinburgh: Institute for Design Informatics

# Workshop Structure

| | |
|---|---|
| **PART 1** | Lecture |
| **PART 2** | Practical: Colab Notebooks |
| **PART 3** | Group Discussion |

# Workshop Structure

https://tinyurl.com/aianarchies

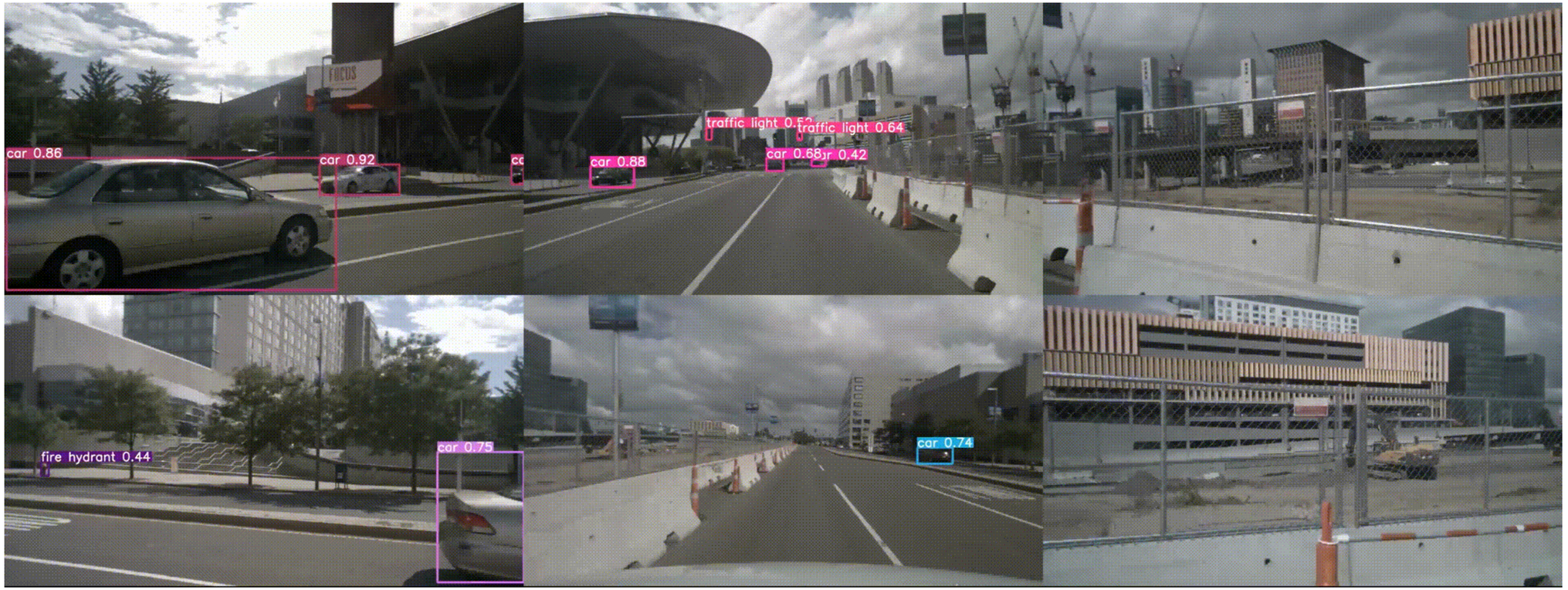
# Methods

The effort is not simply to import well-known methods - be they from humanities, social science or computing. Rather, the focus is on how methods may change, however slightly or wholesale, owing to the technical specificities of new media.

—

Digital Methods Initiative Wiki

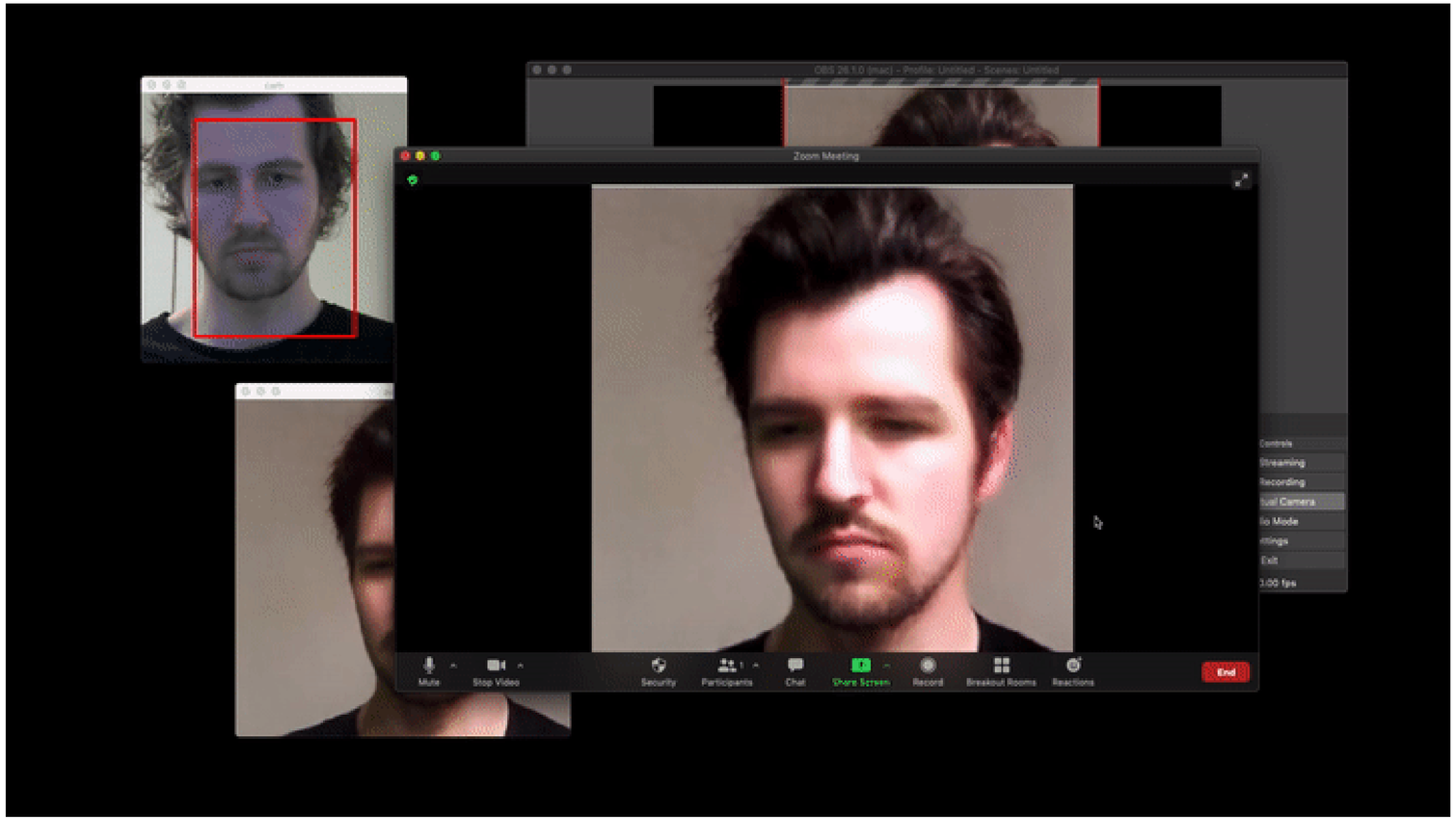
# Practice




—

Excerpt from Future False Positive (2019)
Martin Disley

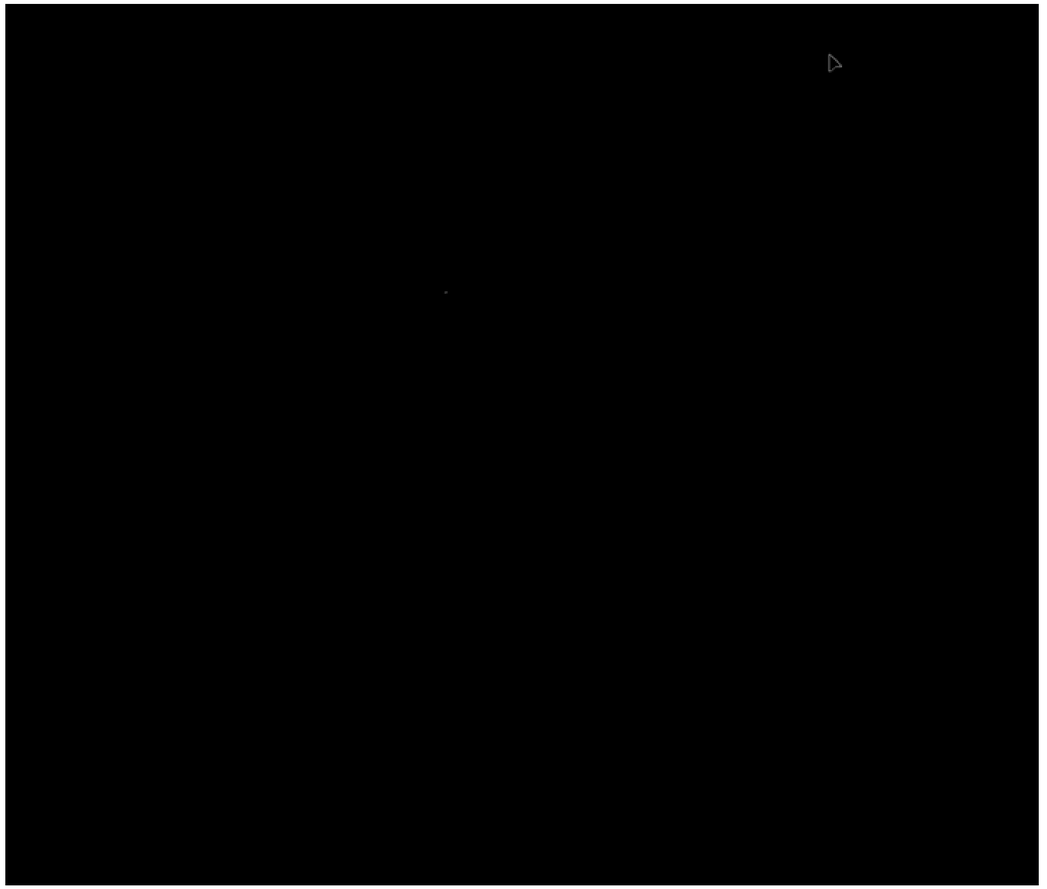
# Practice




—

Excerpt from How They Met Themselves (2021)
Martin Disley

# Practice



Excerpt from Layers of Abstraction: A Pixel at the Heart of Identity (2018)
Murad Khan

# Practice

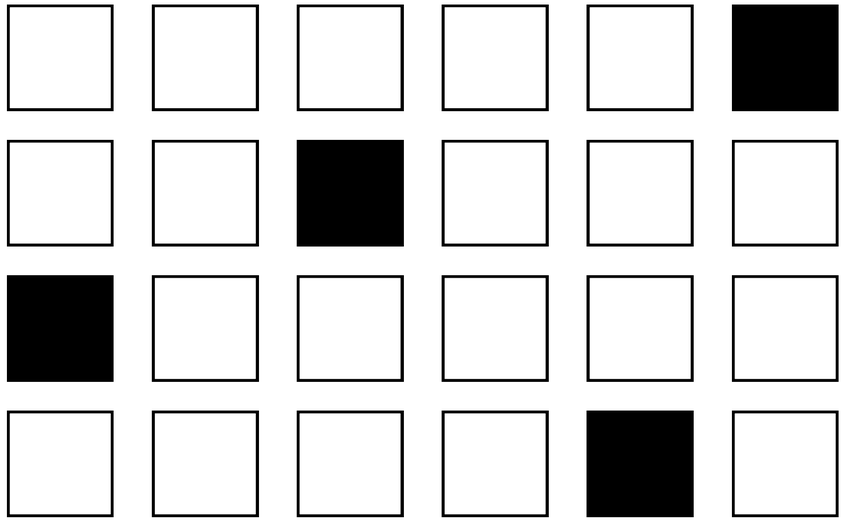
Excerpt from Speech2Face research (2022)
Unit Test

# Adversarial Machine Learning

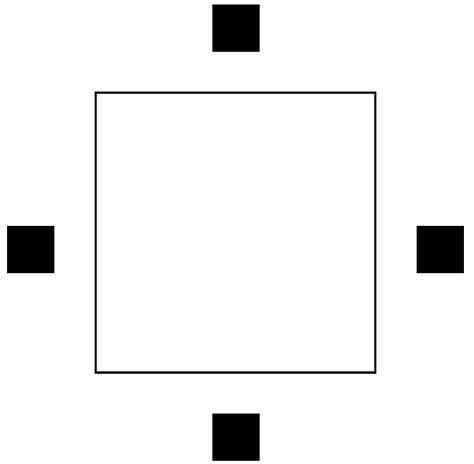
Adversarial machine learning is a field of research concerned with the study of vulnerabilities in machine learning systems, developing a set of techniques and toolkits to both test the robustness of a model and to improve accuracy.

# Types of Attack

01 **Causative** attacks occur at training time, focused on the data used to train a model.

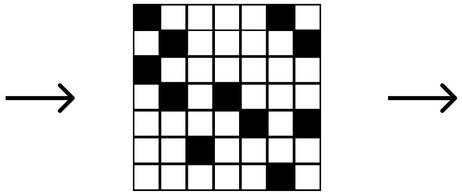


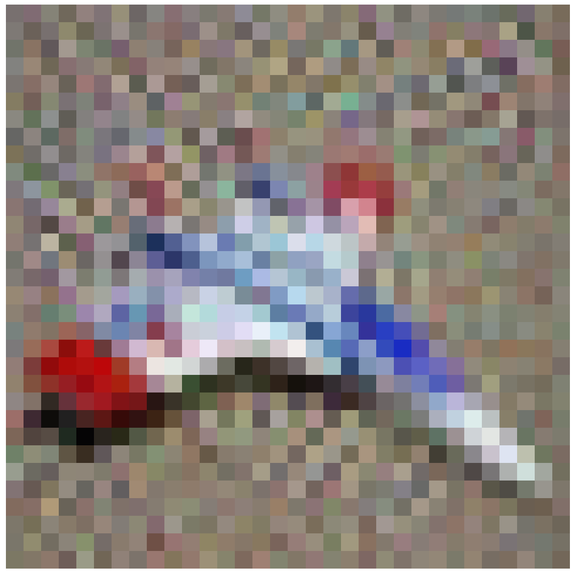
02 **Exploratory** attacks occur at testing time, probing the model via its inferential capacity.



# Adversarial Examples

Original Image
**Prediction: Plane**





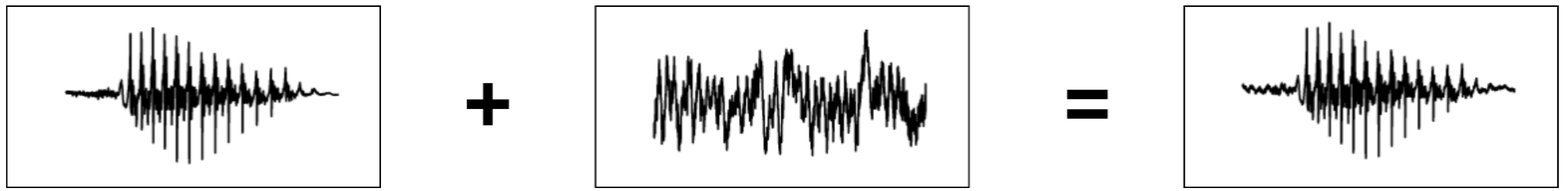
Adversarial Image
**Prediction: Bird**

—

Method: Carlini and Wagner algorithm

Model: CIFAR-10

# Adversarial Examples



"it was the best of times, it was the worst of times"

"it is a truth universally acknowledged that a single"

—

Method: Carlini and Wagner algorithm

Model: Deep Speech (Mozilla Implementation)

# Experiments

The development of techniques for facial recognition models have reinvigorated the prosopoetic relationship between interior and exterior common to the pseudo-sciences of physiognomy and phrenology.

The exterior body is made to speak for a necessary and hidden interior character - an essential and stable subject.

# Crip Technoscience

Disability technoscience reinforces the sense that disabled people are not already making, hacking, and tinkering with existing material arrangements. Disability is cast as an object of innovation discourse, rather than as a driver of technological change

—

Hamraie, A. and Fritsch, K., 2019. Crip technoscience manifesto. Catalyst: Feminism, Theory, Technoscience, 5(1), pp.10

# Questions

01 How can we think about adapting the tools of machine learning to develop a more experimental vocal practice?

02 How do we negotiate the boundaries that are enforced in the development and application of sociotechnical systems?

03 How do we point beyond their limits?

# Experiments

## Measurable

Acoustic properties of speech typically made amenable to scientific analysis.

## Symbolic

The dynamics of listening and interpretation.

# Source Filter

# Pink Trombone

https://dood.al/pinktrombone/

# Measuring Difference

The introduction of the spectrograph focused attention away from the voice as unique aspects of the individual, and towards a standardised framework for speech sounds.

# Measuring Difference

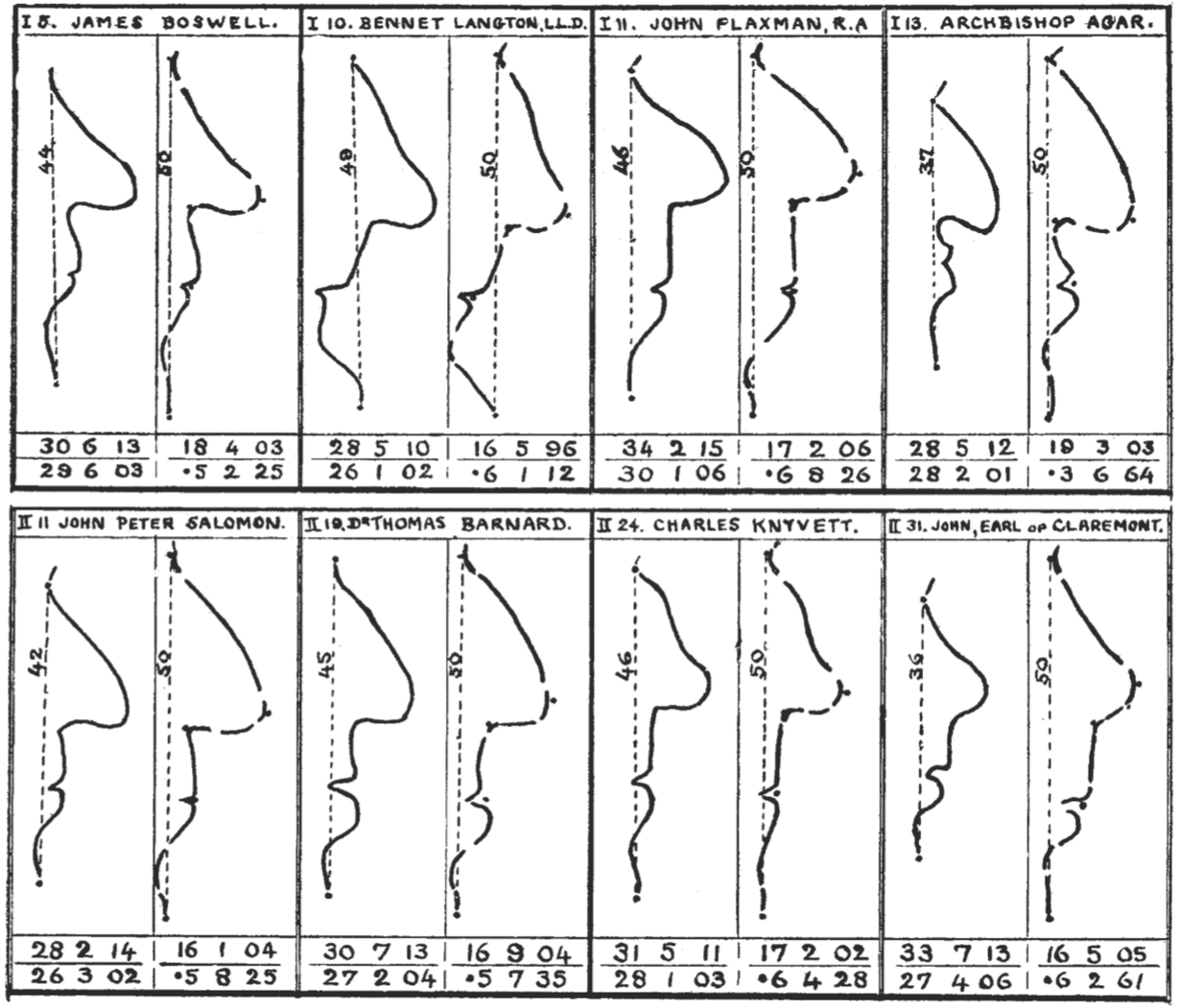



FIG. 5.

*Explanation* of the first formula, namely, that of James Boswell; the others are to be read on the same principle. $N_x$, 30; $N_y$, 13. $U_x$, 28; $U_y$, 03. $L_x$, 18; $L_y$, 03. The small letters are, *n*, 6; *u*, 6; *lp*, 4. *b*, 5; *g*, 2; *ll*, 2; *k*, 5.

# Measuring Difference

What we see emerge in an acoustic model of the voice is the extrapolation of acoustic features, mediated by digital representation (a ‘numeralised’ profile), which is then taken back out to the contingencies of the body in order to *reveal* something about the speaker, about the subject that stands static behind the signal.

# Fourier Transform

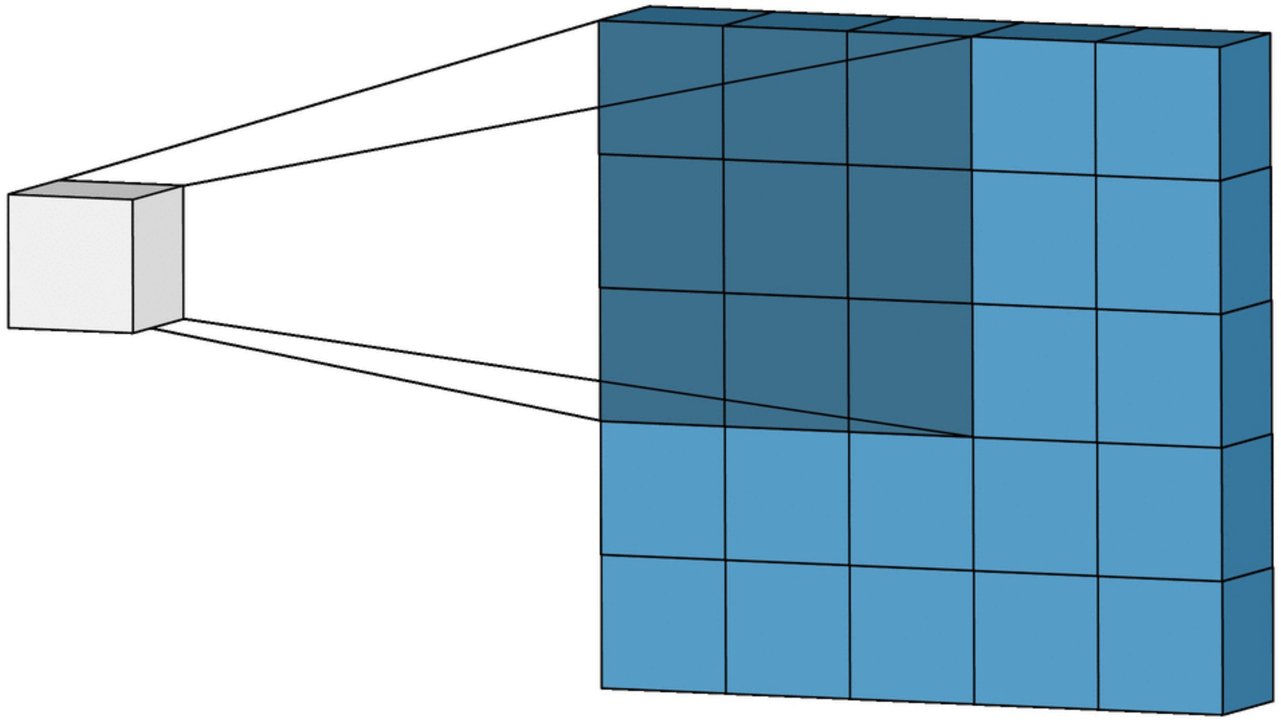
## Sonicity is where time and technology meet

—

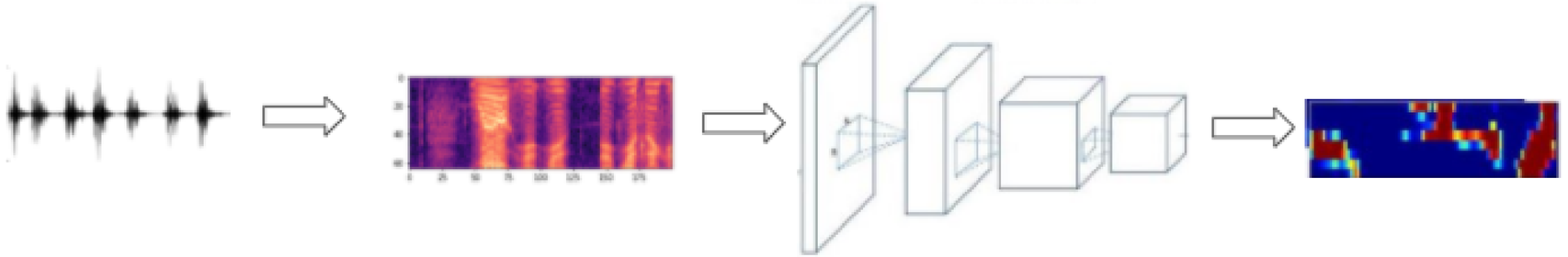
Ernst, W., 2016. Sonic time machines: Explicit sound, sirenic voices, and implicit sonicity. Amsterdam University Press, 21.

# Fourier Transform

# Convolutional Neural Networks



# Convolutional Neural Networks



Audio wave

Spectrogram

CNN Architecture

Feature Maps

# An Unstable Image

I call my vocal condition experimental: every day, every encounter is an experiment where my voice, once a constant in my self-conception, is now a variable.

—

Sterne, J., 2019. Ballad of the dork-o-phone: Towards a crip vocal technoscience. Journal of Interdisciplinary Voice Studies, 4(2), pp.179-189.

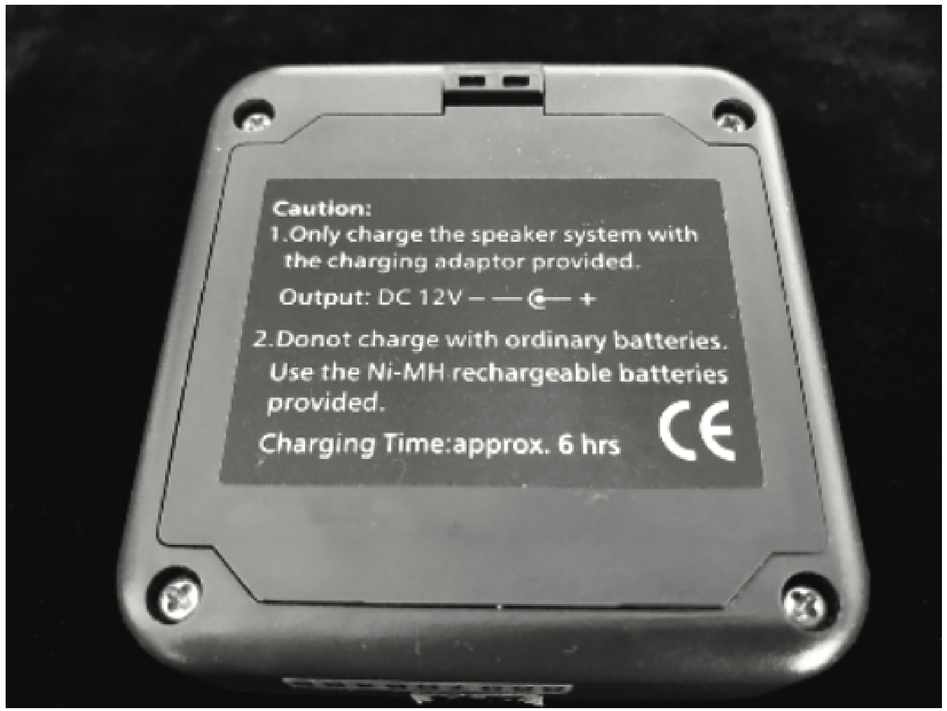
# An Unstable Image

...one of the core questions for a political phenomenology of vocal impairment: what happens to a subject when something that was stable becomes a variable, and that something is one of the mechanisms through which others infer the subject's subjectivity?

—

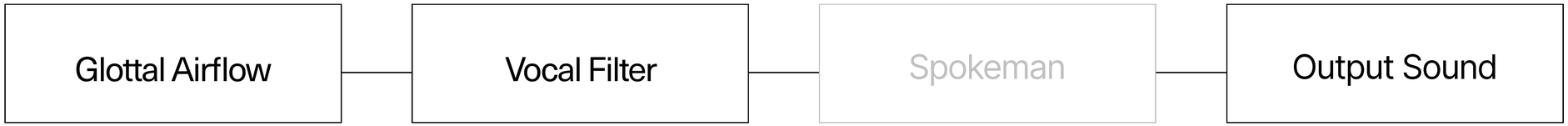
Sterne, J., 2019. Ballad of the dork-o-phone: Towards a crip vocal technoscience. Journal of Interdisciplinary Voice Studies, 4(2), 181

# An Unstable Image


Jonathan Sterne's Spokeman Personal Voice Amplifier (AKA the 'dork-o-phone)

# An Unstable Image



# Measuring Difference

Whilst vocal forensics often accounts for epigenetic reconfiguration of the voice, it fails to resolve this ontogenetic condition of Sterne’s voice - now developing in unison with the technical environment in noisy and non-linear fashion.

# Break

# Practical

## 01 Notebook 1: Encoding

Walk through the process of encoding and analysing your voice.

## 02 Notebook 2: Embedding

Create embeddings of your voice in a latent space and explore how difference is understood by the model.

## 03 Notebook 3: Decoding

Intervene in the encoded representation and decode this back to the audio domain.