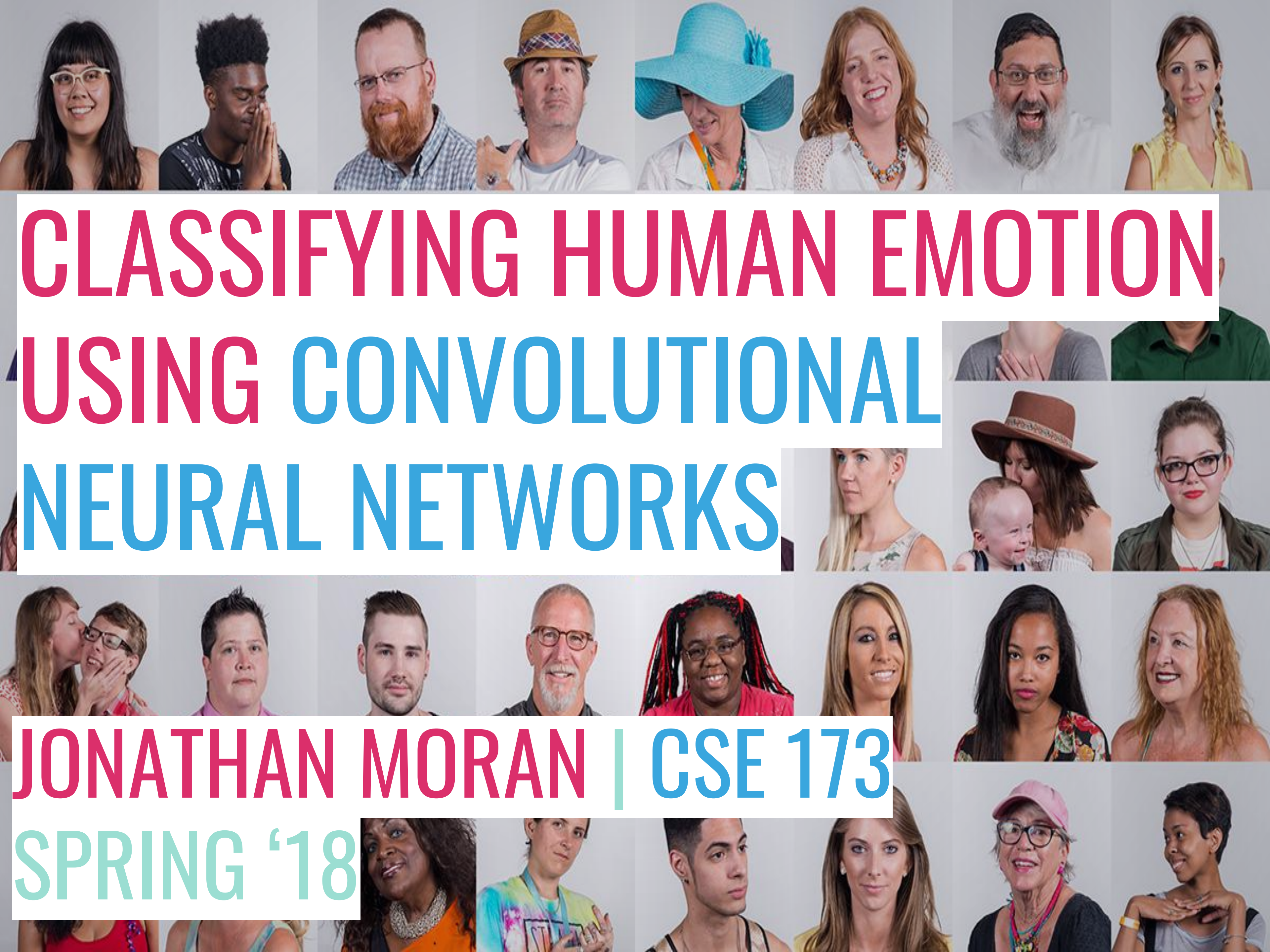


CLASSIFYING HUMAN EMOTION USING CONVOLUTIONAL NEURAL NETWORKS

JONATHAN MORAN | CSE 173
SPRING ‘18

1

EMOTION

AND ITS ROLE AS A UNIVERSAL LANGUAGE

facial expressions of emotion are universal, not learned differently in each culture

– Charles Darwin, *The Expression of Emotions in Man and Animals*

THEORETICAL RATIONALE

WHY EMOTION?

Emotions are universal

Many scientists believe that there are **seven universal emotions** that humans are biologically-hardwired to express

Facial expressions account for nearly 70% of all non-verbal communication

WHY AI?

AI is everywhere

Artificially-intelligent systems can be trained to identify faces and recognize emotions using the same basic visual scanning techniques done by humans

These systems have many different applications in consumer technology

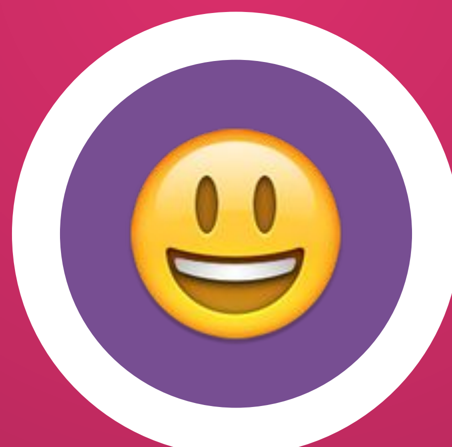
CLASSIFYING EMOTION

Anger

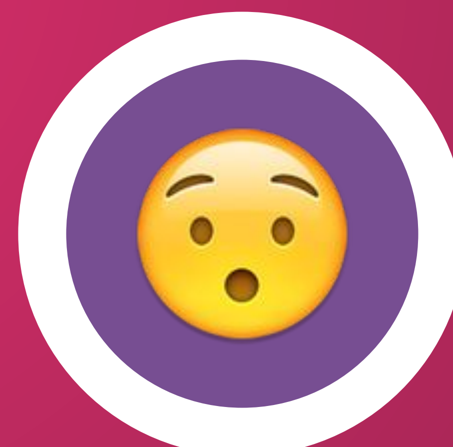
Disgust

Fear

Sad



Happy



Surprised

HOW EMOTION IS EXTRACTED

PRE-PROCESSING

Images from a dataset are normalized, so inconsistencies do not affect learning

Cropped and resized to a 48x48 input image

ANALYSIS

Viola-Jones algorithms from OpenCV and TFLearn detect features

Subsets are scanned for facial markers

CLASSIFICATION

Input is then mapped to the softmax output layer nodes

The unit with the highest activation gets selected

THE FER2013

A large, publicly-available dataset used to train and validate the CNN’s emotion-detecting capabilities

28,709 faces

Used to train and validate the model

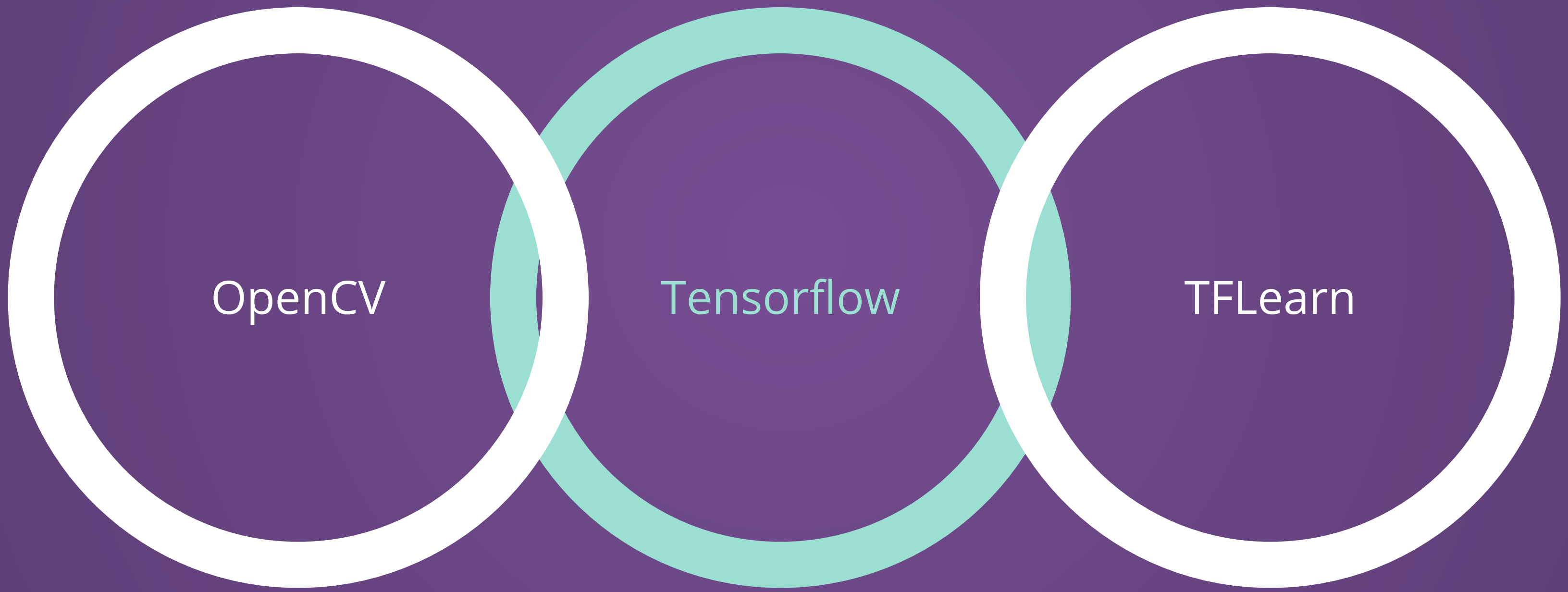
48x48 pixel

Tokenized grayscale images

100 epochs

To converge to 67% accuracy (approx. ~40 hrs)

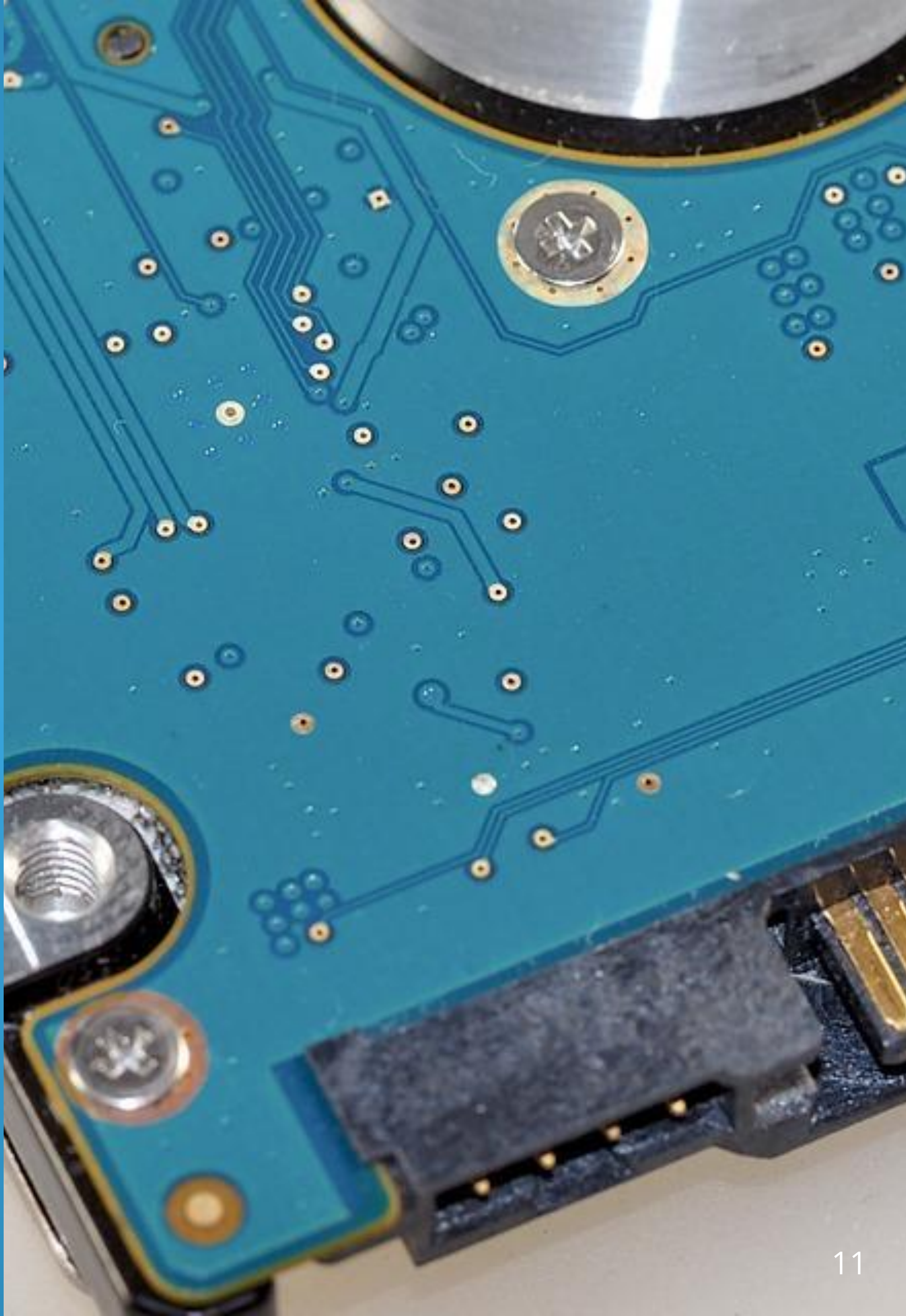
CONVENTIONAL FACIAL RECOGNITION
USING IMAGE CLASSIFIERS AND FEATURE DETECTORS



A. Gudi. Recognizing semantic features in faces using deep learning. *arXiv preprint arXiv:1512.00743, 2015*

NETWORK BASICS:
IMAGE CLASSIFIERS AND FEATURE DETECTORS

- **Tensorflow**
  - machine learning framework
- **TFLearn**
  - high-level API for deep learning
- **OpenCV**
  - open-source computer vision APIs
- **Docker**
  - python/tensorflow runtime environment



EMOTION RECOGNITION NETWORK STRUCTURE

Input layer	Conv. layer	Max-pooling layers	Conv. layer	Fully-connected layer	Softmax output layer
FER2013 grayscale images [48x48]	Detect input “features” [5x5] x 64	Apply local kernels [3 x 3]	Append + update weights [5x5] x 64	Map emotions to faces [14524x1]	Map faces to emotion [1x7]

SUMMARY OF NETWORK OPERATIONS

INPUT

FER2013 images come pre-cropped to 48x48 px, and have been cleaned up, rescaled and converted into a numpy-readable array

EDIT

OpenCV Viola-Jones algorithm normalizes factors like face location, low image quality, in-plane tilt and rotation to filter out poor data

ITERATE

Kernels (“filters”) operate on subsets of the input matrix to quickly discard unnecessary artifacts and retain the facial expression features

COMPARE

The AlexNet CNN model uses TFLearn and OpenCV libraries to calculate Haar-features and reduce negative windows (cascade filters)

UPDATE

Error backpropagation functions from TFLearn library are used to update the activation weights in the convolutional layers

CLASSIFY

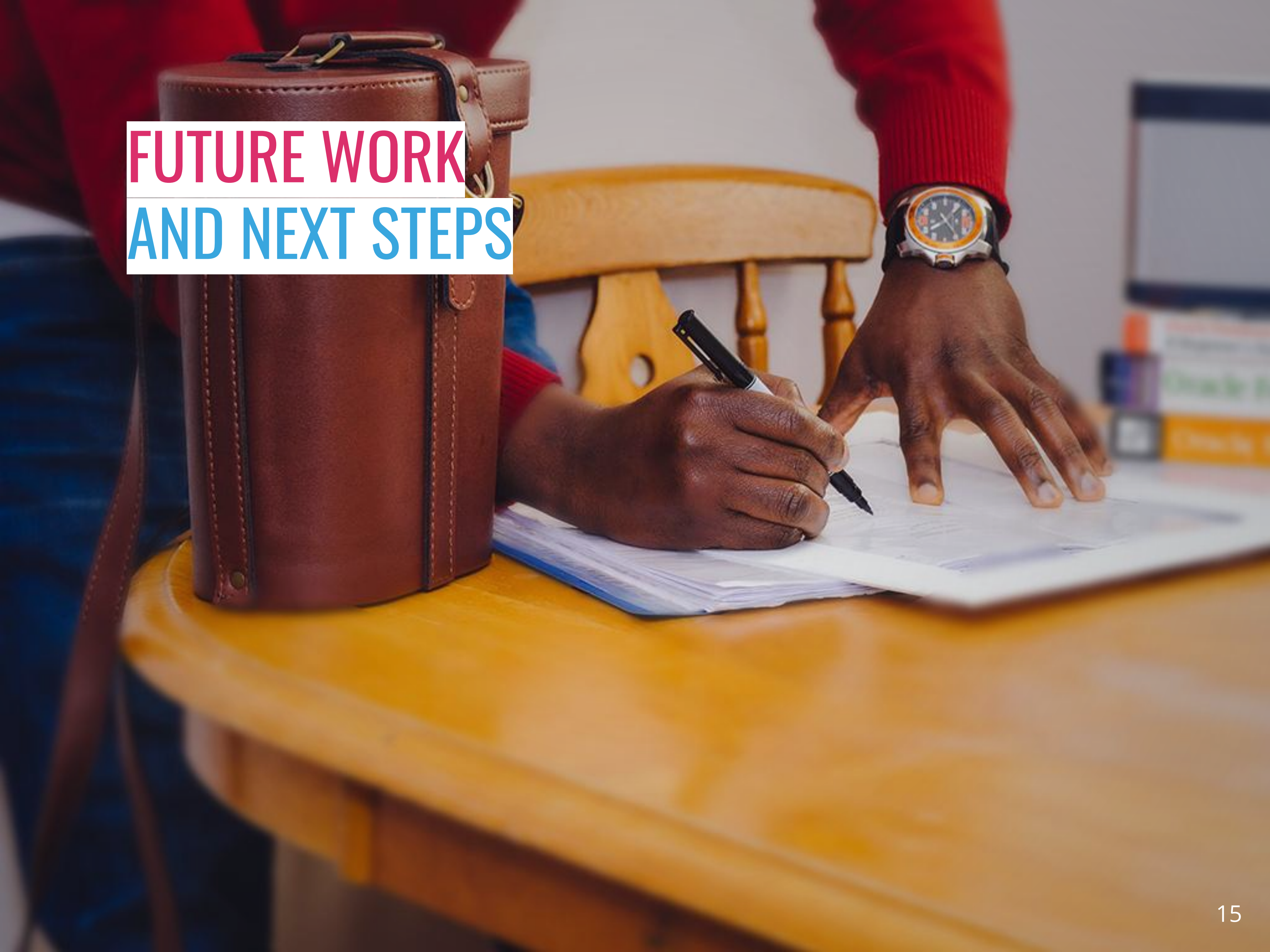
The fully-trained deep net selects the softmax output neuron that has the highest activation given by the ReLU Rectifier: `max(x, 0)`

PERFORMANCE MATRIX USING INITIAL DATASET

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	0.5						
Disgust		0.62					
Fear			0.37				
Happy				0.90			
Neutral					0.80		
Sad						0.28	
Surprise							0.77

* Data provided by TU Delft and @isseu on Github, ran using the same neural net and training set

FUTURE WORK AND NEXT STEPS



FUTURE WORK + CONSIDERATIONS

Train the network using various datasets

- Reduced FER2013
- Japanese Female Facial Expressions (JAFFE)
- CK+ dataset

Analyze for other facial feature characteristics

- Gender
- Age
- Race

Implement emotion-detection in consumer tech

- Automatic playlist generation
- Mood prediction in behavioral health apps

REAL-TIME CLASSIFICATION

Trained classifier operates on individual frames from a live video stream

CREDITS

A. Gudi. Recognizing semantic features in faces using deep learning. *arXiv preprint arXIV:1512.00743, 2015.*

C.R. Darwin. *The expression of the emotions in man and animals.* John Murray, London, 1872

OpenSourceComputerVision Face detection using haar cascades. URL `https://docs.opencv.org/master/d7/d8b/tutorial_py_face_detection.html`

TFLearn. TFLearn: Deep learning library featuring a higher-level API for Tensorflow. URL `http://tflearn.org`

Kaggle. Challenges in representation learning: Facial expression recognition challenge, 2013.

THANKS

ANY QUESTIONS?

Github: @jonathanloganmoran
jmoran23@ucmerced.edu