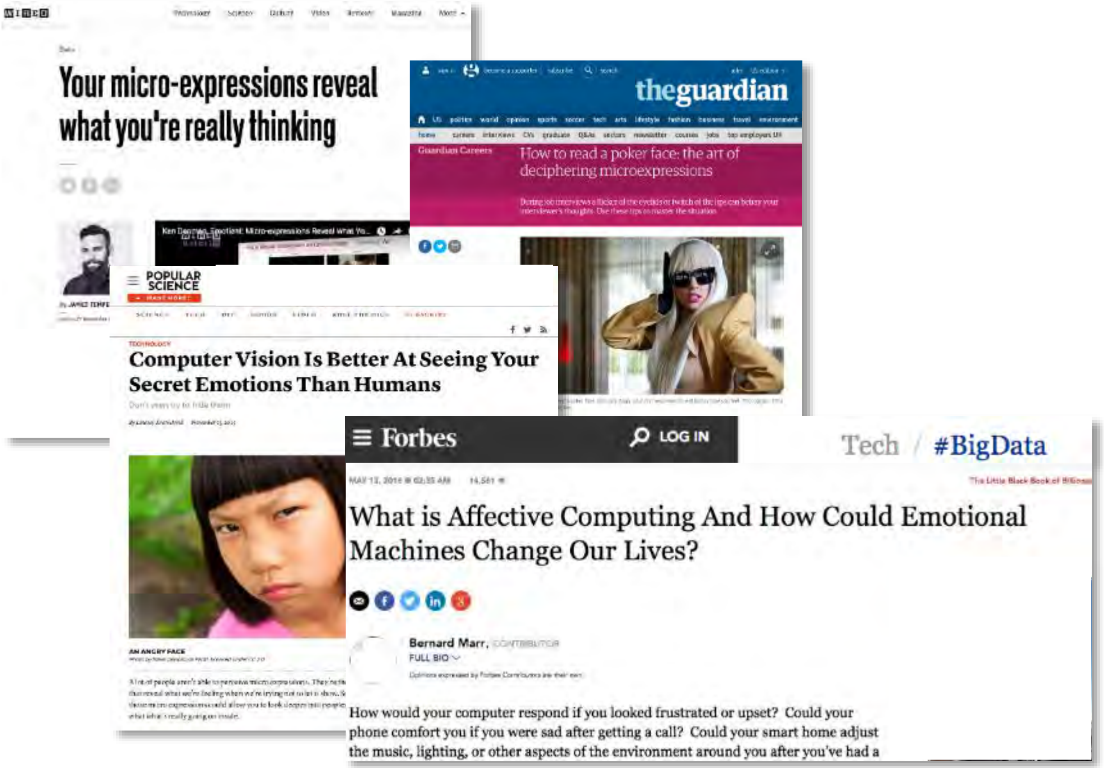
Research Review 2017

Micro-Expressions: More Than Meets the Eye

Oren Wright, Research Scientist

Satya Venneti, Senior Member of Technical Staff and Principal Investigator

Our Work in Micro-Expressions



- Develop algorithms to recognize facial micro-expressions from video.
- Improve upon the capability and accuracy of the state of the art in emotion recognition.
- Use deep neural networks, transfer learning, and signal processing techniques.

Revealing Leaked Emotions Using Software

CASME II database

Facial Micro-expressions

- Involuntary
- Fleeting
- Low intensity
- Universal across cultures
- Very difficult to suppress

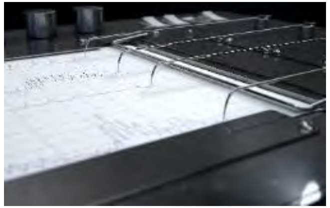
Revealing Leaked Emotions Using Software

CASME II database

Facial Micro-expressions

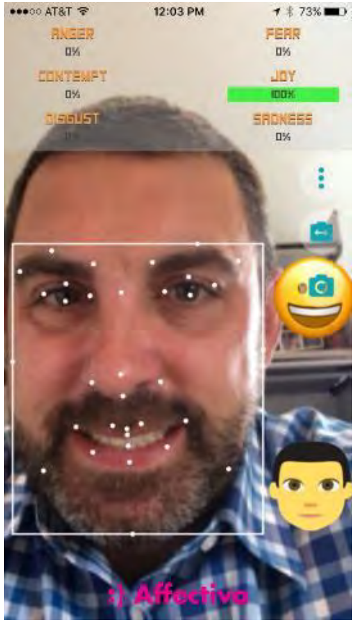
- Involuntary
- Fleeting
- Low intensity
- Universal across cultures
- Very difficult to suppress

Relevant DoD Applications



- Security checkpoint encounters
- Interrogations
- Polygraph testing
- Media analysis
- Media exploitation
- Detection of stress, PTSD

Current State of the Art: Recognizing Emotion



Macro-expressions: Broad smiles, exaggerated frowns, easy to fake. Current approaches work well.

Micro-expressions: Low intensity, fleeting, hard to fake. Current approaches are inadequate.

Challenges

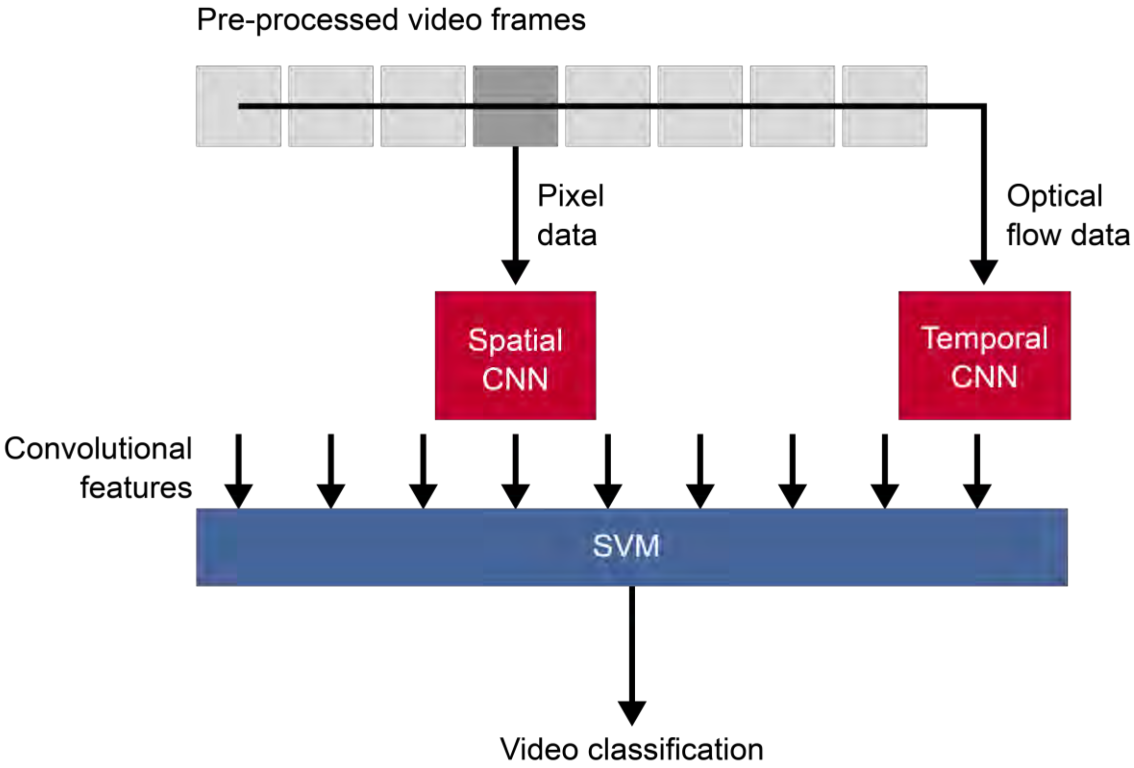
Inherent nature of micro-expressions

- Low intensity, short duration

Paucity of data

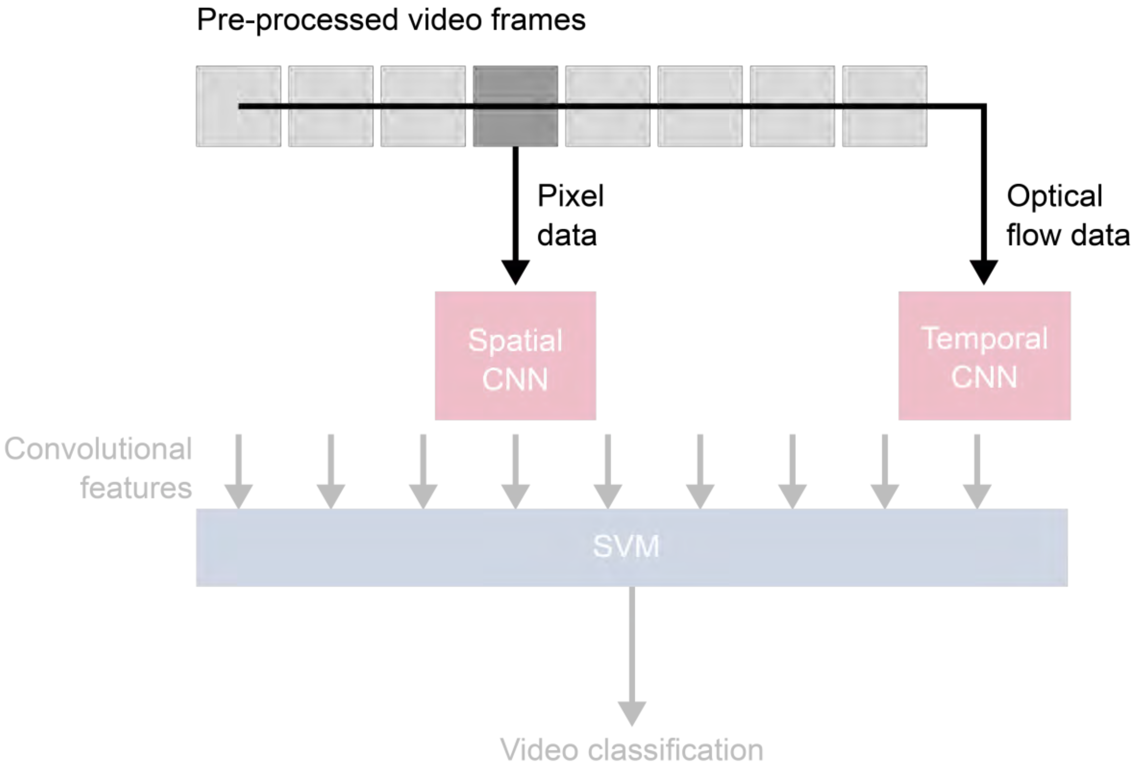
- Difficult to collect
- Limited size
- Contrived data, does not extend well to practical applications
- Imbalanced labels, some emotions are easier to simulate than others
- Varying conditions across different datasets

Our Approach



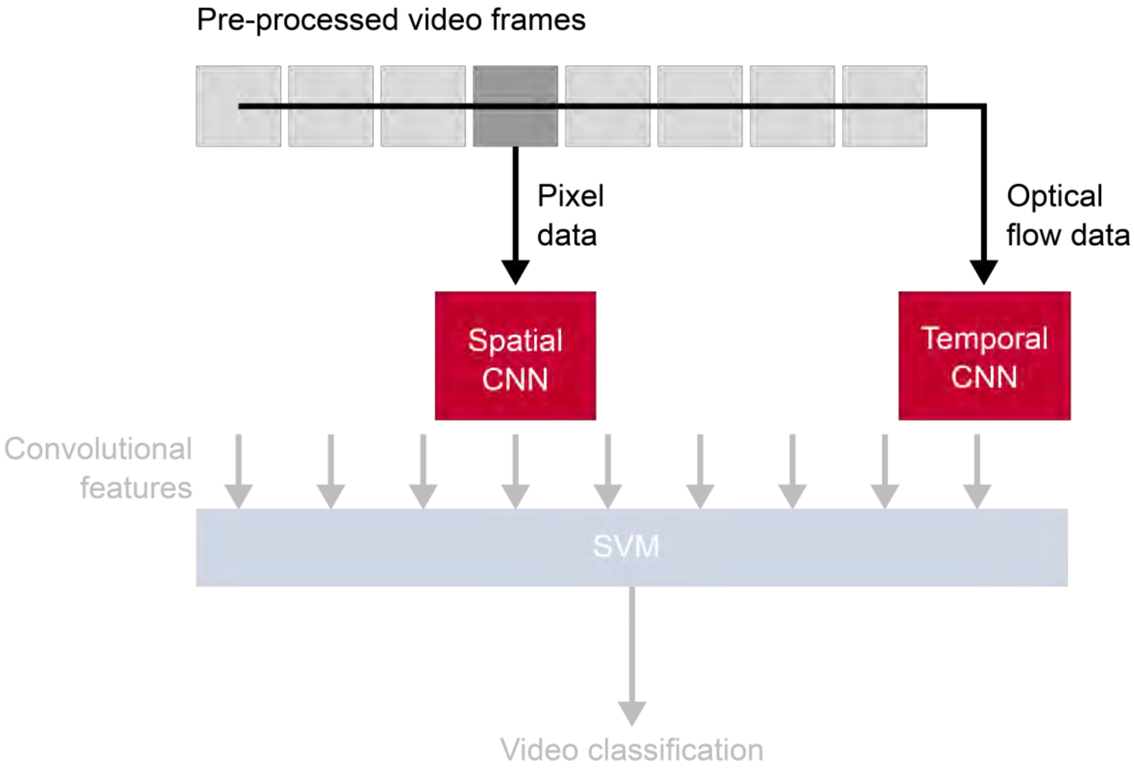
- Use both pixel data and optical flow data to capture spatial and temporal information
- Use convolutional neural networks to create machine-learned data features
- Integrate both data streams into a single classifier

Our Approach



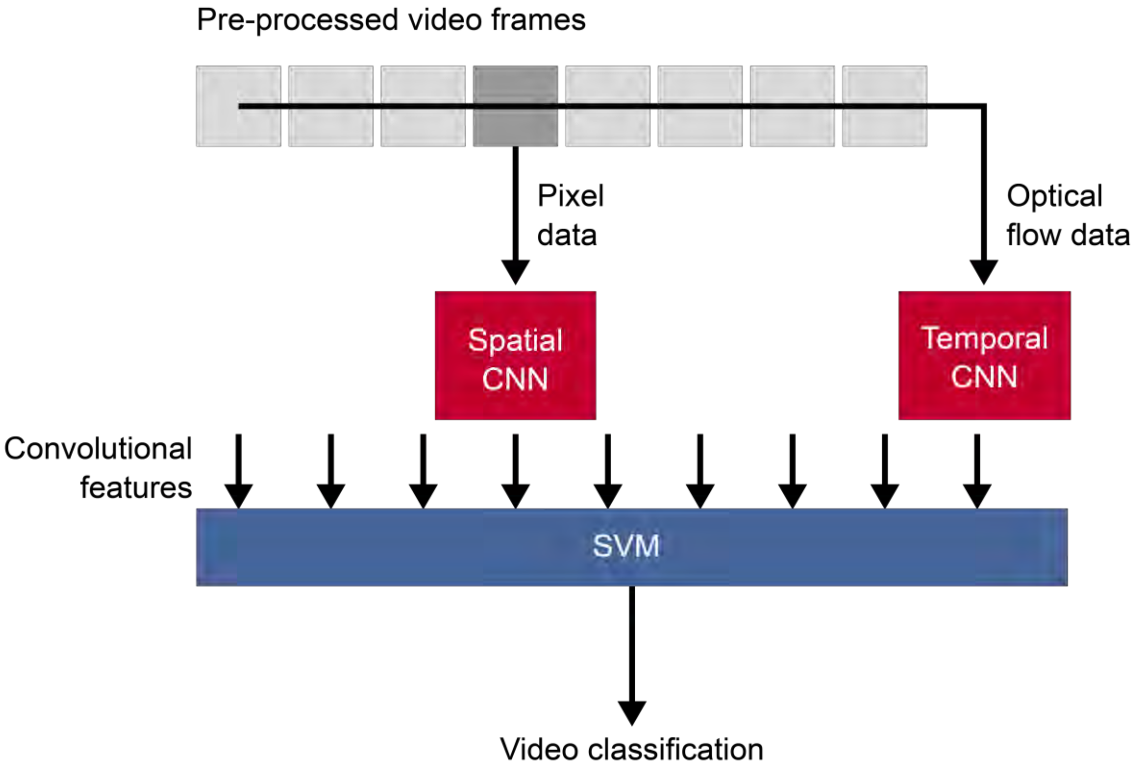
- Video frames are time-interpolated via graph embedding
- Pixel data are pre-processed via techniques such as histogram normalization and horizontal flipping
- Lucas-Kanade optical flow data are extracted to add temporal structure

Our Approach



- Convolutional neural networks serve as feature generators in place of traditional hand-crafted features
- The spatial CNN is built from VGG16 and is pre-trained on over 1.1 million images from ImageNet

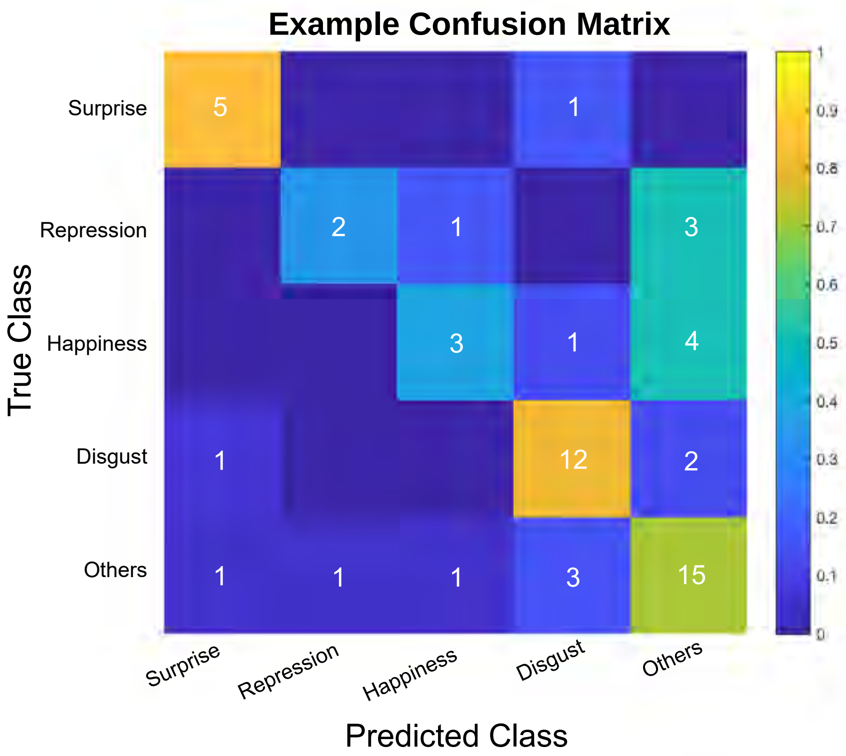
Our Approach



- Convolutional features for both spatial and temporal information are combined into a single SVM to score each video

Results

Architecture	Accuracy
CASME2 baseline	63.4%
Spatial CNN	67.7%
Spatial CNN + Temporal CNN	67.7%*



Micro-Expressions Example

Used with permission of the Poker Channel: https://www.youtube.com/user/sergeypoker/

Micro-Expressions Example

Used with permission of the Poker Channel: https://www.youtube.com/user/sergeypoker/

Future

Engineering Opportunities

- Improve the state of the art for detection
- Optimize feature generators
- Experiment with other pre-processing techniques

Research Questions

- How can we improve existing datasets?
- How can we improve detection accuracy?
- What more can we do to improve recognition accuracy?

Mission Application Challenges

- Accuracy vs. runtime
- Combine detection with recognition
- Deal with long-running videos
- Migrate to GPUs for faster performance
- Combine with other modes like audio for more accurate emotion recognition

Long Term: Advancing Human-Machine Teaming

Photo: Lt. Col. Deanna Bague, Fort Bliss Public Affairs Office

- Extracting heart rate from video (2016)
- Micro-expressions (2017)
- Emotion from voice (2018)

Contact

Oren Wright

Research Scientist

owright@sei.cmu.edu

Satya Venneti

Senior Member of Technical Staff and Principal Investigator

srvenneti@sei.cmu.edu