Deception Detection: Predictive Computational Modeling

Dimitris N. Metaxas
CBIM, Rutgers University

Collaborators: MURI Team

Rutgers Students: Lezi Wang, Anastasis Stathopoulos, Ligong Han, Zachary Daniels (Graduated this May, now at SRI)

Thrust 2 Goal

We are focusing this year on analysis of the visual cues from faces (non-verbal) for Deception Detection

Research Activities over the past Year

Part1: Augmented facial behavior tracking systems are applied to analyze the entire games and corresponding videos, which include 285 players collected from 5 sites in 3 different countries.

Part2: Deception detection (mostly spies) is formulated as a classification problem.

- Temporal Convolutional Network (TCN) is used to model deceivers and truthtellers facial movement and expressions.
- We also added interpretability, by proposing the attention mechanism for TCN to discover and localize the dynamic cues of facial movements that the model attends to, which can detect deception behavior.

Part3: On-going work: Self-attention mechanism to offer improved localization based on new more detailed annotations, weakly/semi-supervised learning, multi-modality for human behavioral analysis, trust, dominance prediction.

Collaborative Research:

- Facial Data Collection and Annotation [**UA, UCSB**] (parts 1,2,3)
- The Communication & Deception theory is adopted as the guidance to link the facial behaviors for deceiver-vs-truthteller to the well-known facial action units, making our discovery interpretable [**UA, UCSB**] (part 2)
- New feature-based methods and ML to detect deception (**Dartmouth, UA and UCSB**)
- New machine learning methods for interpretability: Attention-based, weakly supervised, C3D, TCN
- Network centric methods (**Dartmouth and Stanford**) to be integrated with ML methods
- Co-authoring papers (BMVC 2019, FTC 2020) about attention mechanism for facial behavior analysis [**UA, UCSB, Dartmouth**] (part 2)

AUTOMATIC LONG-TERM DECEPTION DETECTION IN GROUP INTERACTION VIDEOS

Chongyang Bai
PhD student
Dartmouth College

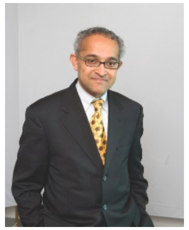
Maksim Bolonkin
PhD student
Dartmouth College

Judee Burgoon
Professor
Director of Research, Center for the Management of Information
University of Arizona

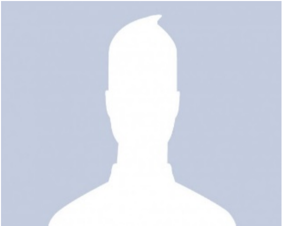
Chao Chen
Master student, Dartmouth
Software Engineer, Facebook

Norah Dunbar
Professor & Chair
UC Santa Barbara

Bharat Singh
PhD Student, University of Maryland
Research Engineer, Amazon



V. S. Subrahmanian
Professor, Director of ISTS
Dartmouth College



Zhe Wu
PhD Student, University of Maryland
Senior Researcher, Comcast

Sales events

Job/Security interviews

Deception Detection

Negotiations

Witness testimonies

Challenges

Single person → Multiple people

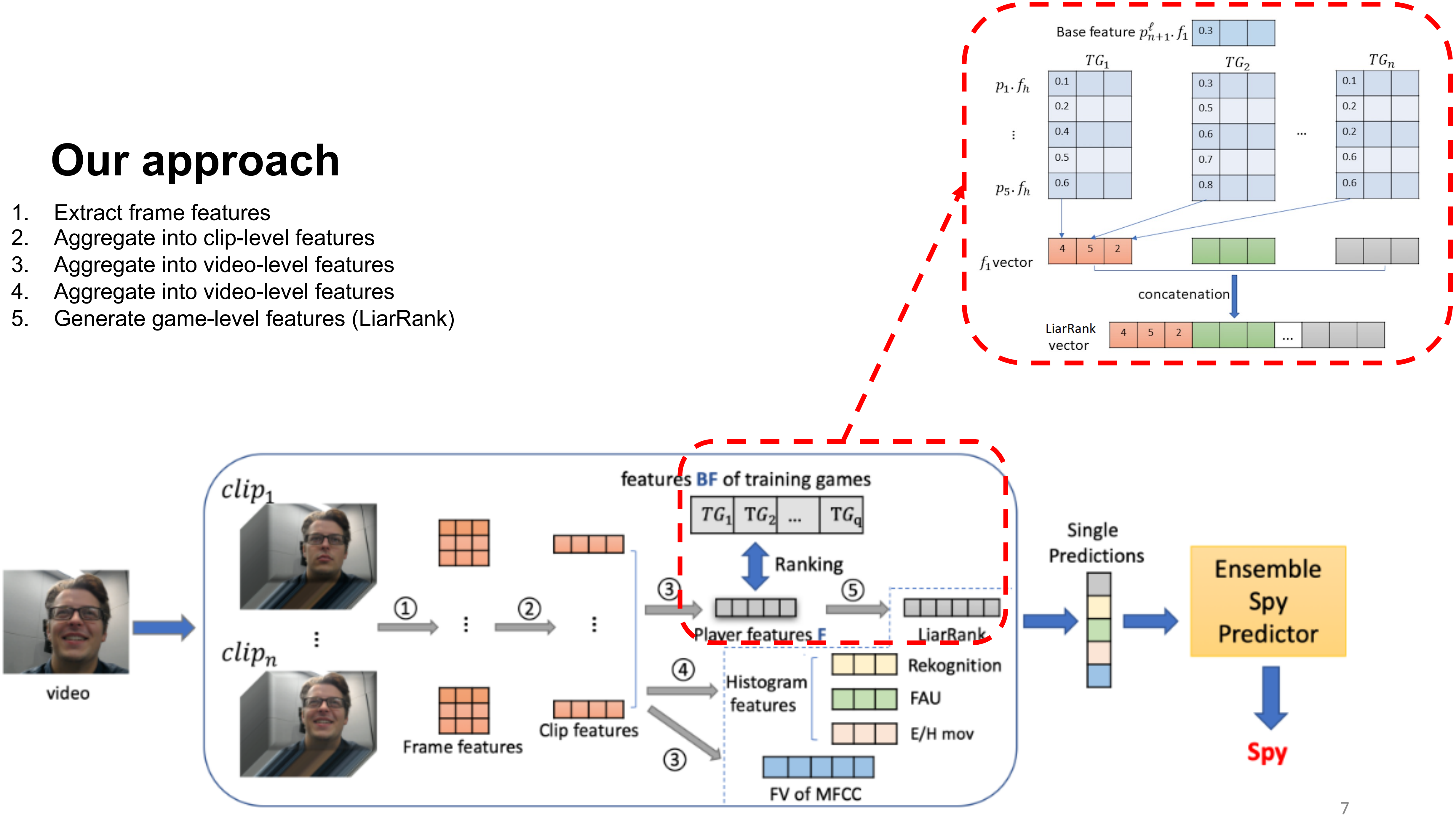
Interview → Group interaction

Short video (1-4 mins) → Long video (30-60 mins)

Strictly controlled setting → Less controlled setting

Our approach

1. Extract frame features
2. Aggregate into clip-level features
3. Aggregate into video-level features
4. Aggregate into video-level features
5. Generate game-level features (LiarRank)



Results

Audio/Visual features

Features	RF	L-SVM	NB	LR	KNN
Average VGG Face (baseline)	0.516	0.533	0.549	0.546	0.50
VGG Face clip-level voting	0.503	0.520	0.550	0.527	0.479
FV of VGG Face	0.468	0.573	0.502	0.584	0.502
FV of VGG Face + FS	0.506	0.470	0.491	0.467	0.522
LiarRank of FV of VGG Face + FS	0.639	0.647	**0.663**	0.652	0.603
FV of MFCC frame-level	0.606	0.395	0.56	0.608	0.579
FV of MFCC clip-level	0.586	0.441	0.533	0.579	0.595

Histogram-based features

Amazon Rekognition					
Frame hist.		Clip hist.		Combined	
Disgusted, Surprised	0.630	Smile, Angry, Disgusted	0.634	Smile, Angry, Disgusted	**0.676**
Surprised	0.622	Smile , Angry	0.623	Smile, Disgusted	0.647
Calm	0.622	Smile, Disgusted, Calm	0.618	Angry	0.638
All features	0.557	All features	0.544	All features	0.563
Facial Action Units					
Frame hist.		Clip hist.		Combined	
AU07+AU10+AU12	**0.621**	AU06+AU14	0.609	AU07+AU09+AU10	**0.621**
AU12+AU23+AU25	0.614	AU07+AU09+AU10	0.606	AU07+AU10+AU23	0.617
AU09+AU10+AU12	0.612	AU07+AU14+AU45	0.603	AU12+AU25	0.611
All features	0.592	All features	0.577	All features	0.608
Eye/Head movement					
Frame hist.		Clip hist.		Combined	
3+8	0.632	1+6+8	**0.671**	1+3+4+5+6+8	0.643
3	0.624	1+6	0.642	1+3+5+8	0.627
3+7	0.615	1+3+6+8	0.636	1+3+5+6+8	0.625
All features	0.591	All features	0.560	All features	0.618

Ensemble results

Classifiers	AUC	F1	FNR	FPR	Precision	Recall
LR+RF+NB+L-SVM+NB	**0.705**	0.466	0.621	0.142	0.666	0.379
LR+L-SVM+NB+L-SVM+NB	**0.705**	0.466	0.610	0.169	0.660	0.390
KNN+RF+NB+RF+NB	0.704	0.403	0.673	0.173	0.622	0.327
NB+L-SVM+NB+L-SVM+NB	0.704	0.406	0.667	0.151	0.624	0.333
LR+KNN+NB+L-SVM+NB	0.704	0.468	0.620	0.143	0.684	0.380

Data for Analysis of players' faces and Geometric Setting

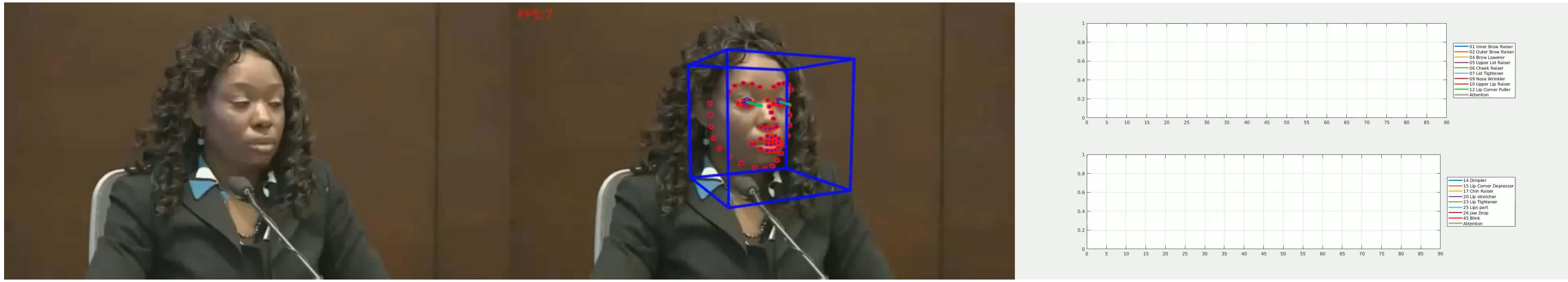
- From UA and UCSB: Both homogenous and heterogeneous games, 285 players/videos in total.

- Data: 285 players/videos collected from 5 sites in 3 different countries to account for possible heterogeneity in deceptive behavior among different cultures. The average video duration is 46 minutes.

- A balanced subset of 230 videos is used in experiments. We use fivefold cross-validations on the combined video data.

Data for Analysis of players’ faces and Geometric Setting

- The data is analyzed with our in-house modified trackers (bounding boxes, landmarks, face tracking)
- This year we added more features:
  - eye gaze
  - facial action units (FAU)
- analyzed all videos



Previous Year

- Last year we developed attention methods that indicate features responsible for deception and truth-teller that were consistent to what we know based on communication theory.

- We apply a 3D convolutional neural (C3D) network to classify a player as the Deceiver or Truth-Teller

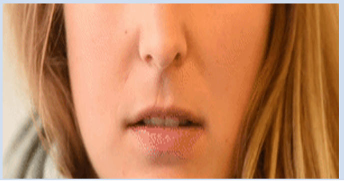
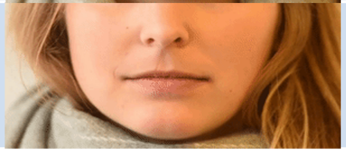
- We use attention mechanism to perform retrospective model analysis and discover AUs that relate to deception behavior.

Deception Behavior Analysis from Communication Theory

- In the latest deception theory, deception is represented by the combination of facial Action Units(AUs), including:
  - More blinks (AU45) with emotional responding and masking, fewer blinks with cognitively loaded responses and efforts at neutralization
  - Sneer (AU9 + AU10) while feigning sadness
  - Lip adaptors (AU18, AU19, AU23, AU24)
  - etc.
- Sources for the above come from various articles and include:
  - DePaulo (2003) (but this is seriously outdated)
  - Cohn, Zlochower, Lien & Kanade (1999)
  - Porter & ten Brinke (2008)
  - Waller, Cray, & Burrows (2008)
  - Kessous Castellano & Caridakis (2009)
  - Matsumoto, Willingham & Olide (2009)
  - Hurley & Frrank (2011)
  - ten Brinke & Porter (2012)
  - ten Brinke, Porter & Baker (2012)
  - Matsumoto & Hwang (2017)

Deception Behavior Analysis from Communication Theory

- Samples of Action Units are considered as deception:

Action Unit	Description	Facial Muscle	Example (Hover to Play)
AU45	Blink	Relaxation of *Levator Palpebrae and Contraction of Orbicularis Oculi, Pars Palpebralis.*	
Sneer **AU9** + AU10	Nose Wrinkler	*Levator labii superioris alaquae nasi*	
Sneer AU9 + **AU10**	Upper Lip Raiser	*Levator Labii Superioris, Caput infraorbitalis*	
Lip adaptors (**AU24**)	*Lip Pressor*	*Orbicularis oris*	
Faked happiness (**AU12**, but missing AU6)	Lip Corner Puller	*Zygomatic Major*	

Previous Year: Training the C3D with Attention-guided Frame Sampling

- 6 homogeneous games, 43 players in total
- Random Sampling (RS): We random sample 16 frames as the input of C3D and the temporal order is kept
- Attention-guided Sampling (AS): We sample the frames according to the attention importance and more frames with high attention probabilities are taken as input of C3D model
- The results show that the model with AS outperforms the RS in the notable margins, ~2% - ~4%
- The accuracy boosting validates the effectiveness of attention mechanism to identify the potential frames where Deceivers and Truth-Tellers show discriminative visual cues so that it is easier to train a model with better accuracy.

#Testing/Training Games	Classification Accuracy
1/5	65.43 (± 0.27)
2/4	62.28 (± 0.30)

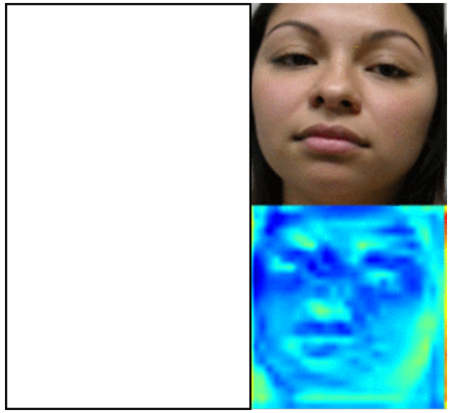
Table 1: Random Sampling

#Testing/Training Games	Classification Accuracy
1/5	67.85 (± 0.25)
2/4	67.03 (± 0.28)

Table 2: Attention-guided Sampling

Previous Year: Use Attention Results to Discover AUs

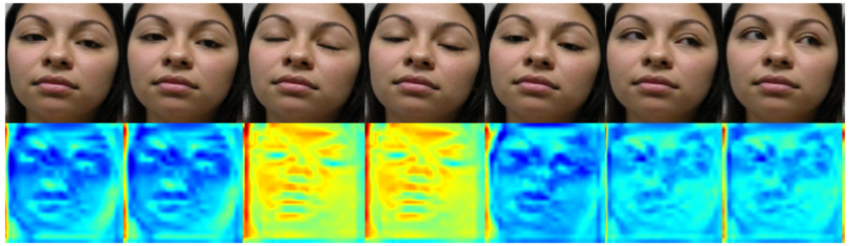
- A video is broken into different short clips
- The clips with high Deceiver/Truth-Teller probabilities from the C3D model are used for deeper investigation
- We compare what the attention model predicts to known deception cues provided by our group experts
    - The facial cues are converted to facial action units (AU)
- Based on the game **spies** are more often **deceptive** than **villagers** who are more often **Truth-Tellers**. But as we observe from the analysis roles can reverse.
- We do not know the locations and duration of when deception occurs.



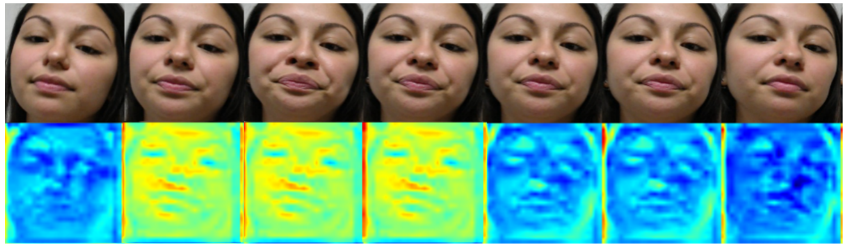
Previous Year: Deception Cues vs. Model attention

- Looks like the network is finding what seems to be known about deception: Here are some AUs addressed by the model attention
- But also their dynamics

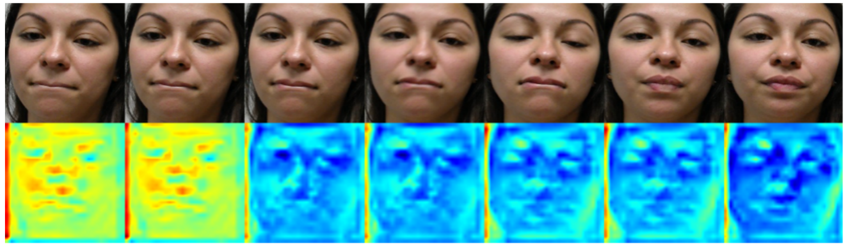
AU45: blinks



AU20: Lip stretcher
AU13: Cheek Puffer

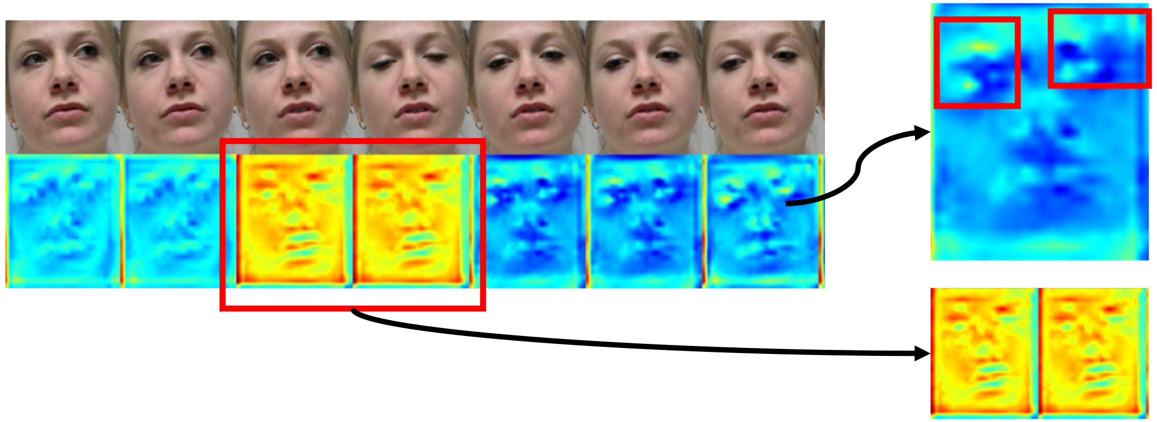


AU24: Lip Pressor



Previous Year: What does our model learn?

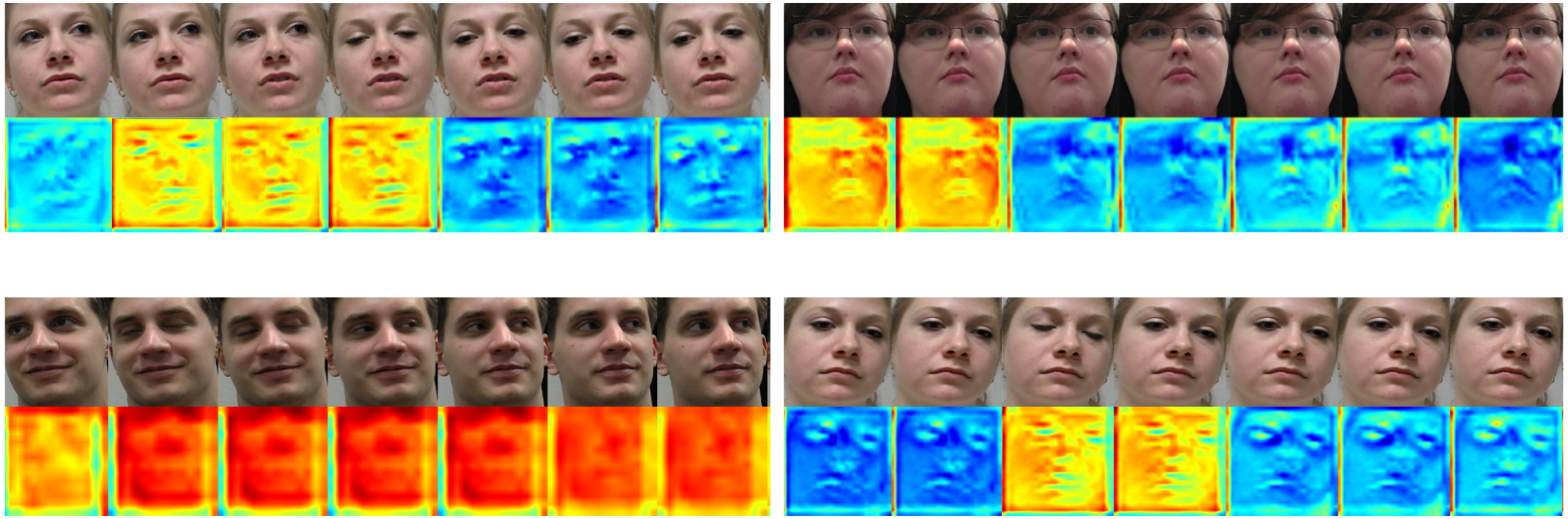
- We observe that C3D
  a) tarts by focusing on most contributing facial feature locations in the first few frames
  b) Then tracks the salient motion in the subsequent frames.
- In the following example, it first focuses on the eyes, mouth and then tracks the motion (variance) happening around them.
- Attention technique highlights the spatial and temporal information which has the positive contribution for the final prediction



Previous Year: Deception Cues vs. Model attention

Examples which fall into the **Deceiver** category but are more subtle

These promising results show: **eyes closed**, **fake smile**, **changes in lips**.



This Year: Data Observations and Challenges we addressed.

- The Deceiver/Truth-Teller visual cues are sparse:
  - Most of the time, Deceivers and Truth-Tellers have the same behavior
  - Annotating the facial difference between Deceivers and Truth-Tellers is almost impossible
  - We do not know the deception behavior location(s) and their duration in given video.
  - Last year we got encouraging recognition results on 6 homogeneous games:
    - 006AZ/ 009SB/ 011SB/ 012AZ/009ISR/011NTU
    - There are 43 players' videos in total, including 18 spies and 25 villagers
    - Total length is ~20700 seconds (~345min)
  - We need to discover directly the AUs present during deception and the location of deception without annotations.
  - C3D is not memory-efficient and is prone to identity overfitting.

This Year

We addressed the following three questions:

- Q1: How to build learning models that are modular, lightweight and independent of the identity of a person (separate facial geometry from expressions and behaviors)

- Q2: How do we perform retrospective analysis of deceptive behavior: Where and when does deception or Truth-Telling behavior occur in terms of AUs?
  - Instead of Pixel level analysis we do AU level analysis

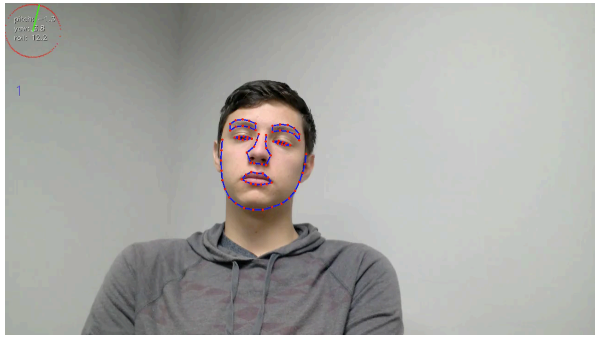
- Q3: Our method should be applicable to other behavior types such as trust, dominance

Part1: Analysis of players' faces and Geometric Setting

- Three angles for head pose are visualized on the left top corner; numbers (blue) indicates the frame ID, and 68 facial key points are plotted
  - The head pose angles are measured in 3D: pitch, roll and yaw (new videos)
- Player's expression in every frame are mapped into three categories: positive, neutral and negative

- The player's faces and facial landmarks can be detected automatically at the same time
  - We proposed a coupled-encoder and decoder network to achieve the two tasks [*published in journal Image & Video Computing 2019*]

Truth-Teller

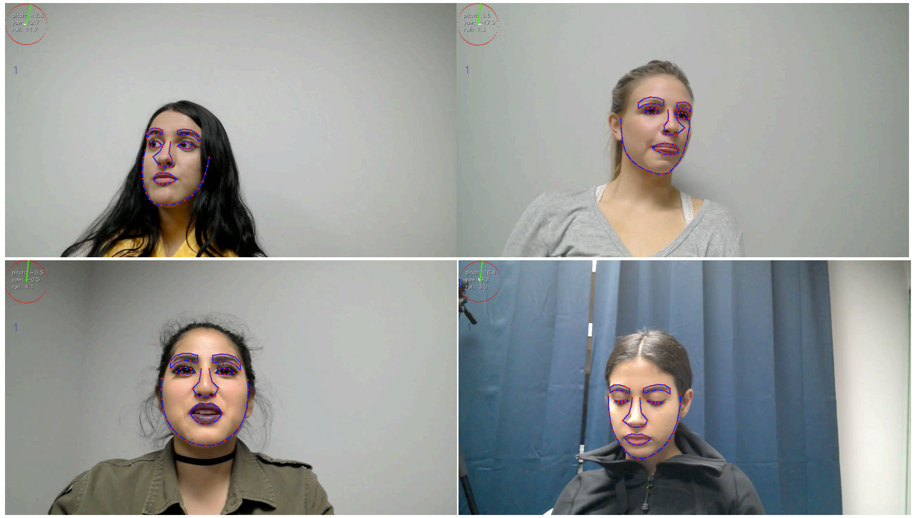


Deceiver

Part2: Interpreting who are Deceivers or Truth- Tellers

Deceivers

Truth-Tellers

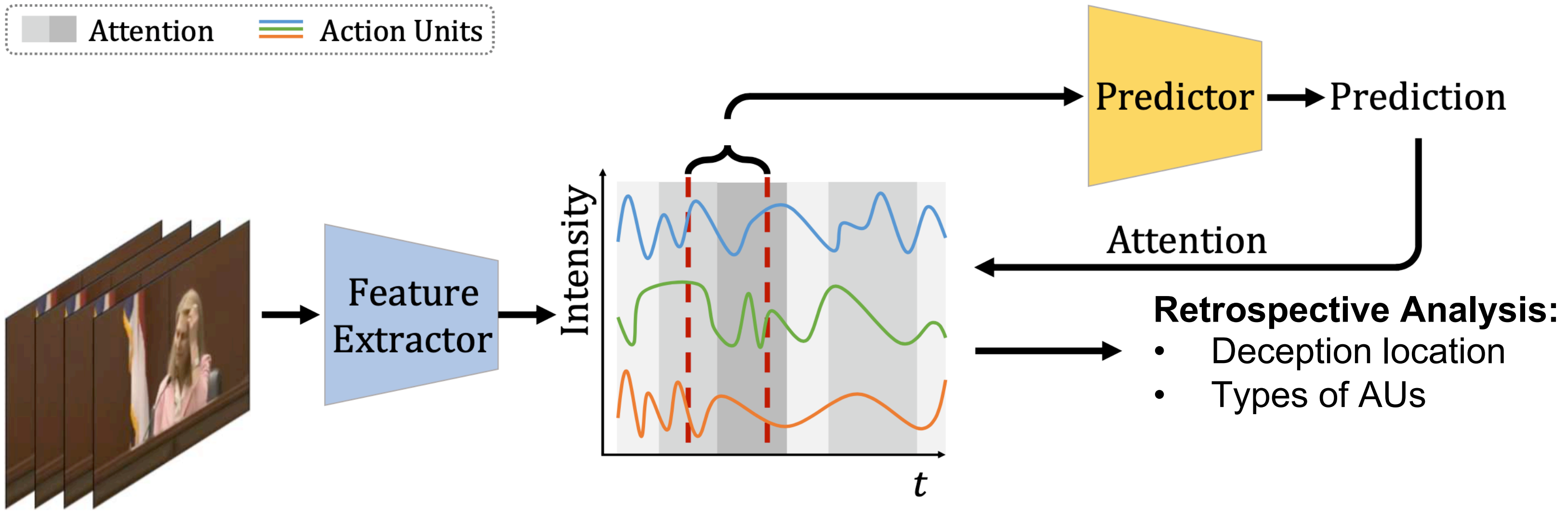


Question 1: Deceiver vs Truth-Teller Classification

- Proposed a two-stage approach
- Extract identity invariant and robust facial features (17 Facial Action Units, or FAUs, normalized with the parameters of the morphable model fitted to subjects' face; gaze angles, etc.)
- Those AUs define a set of 1-D signals (over time);
  Concatenate those 1-D signals channel-wise
- Feed input AU waveforms to a Temporal Convolution Network (TCN)
- Use labels to train the model for binary classification at the video level (no groundtruth framewise)

Question 1: Deceiver vs Truth-Teller Classification

- Pipeline overview
  - 1-D FAU signals are extracted from video sequences.
  - A predictor temporal convolutional network is then trained on the extracted waveforms.
  - In this ML framework attentions are computed by backpropagating the trained predictor model to discover type of AUs and deception location

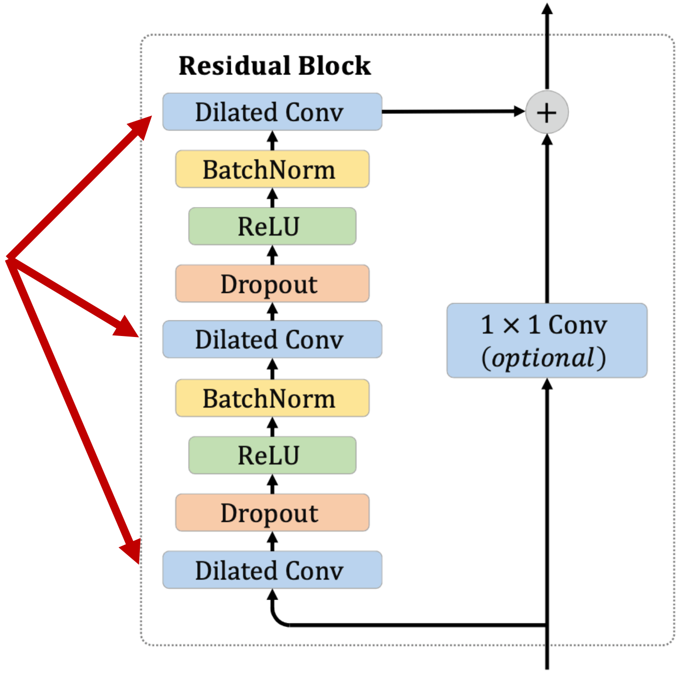


Temporal Convolutional Networks (TCN) for Video Classification

- Residual blocks used in the proposed temporal convolutional networks (TCN)

- In order to capture long-term dependencies in our input we use residual blocks with dilated convolutions.

- Finally, we do average pooling on the feature maps and apply a Fully Connected layer to get the final prediction.



Residual blocks used in the proposed video classification model.

Bayesian Ensemble for Improved Classification

- To further improve the results we develop a Bayesian neural network (BNN) variant of the proposed TCN
- Bayesian neural nets try to estimate the **posterior** over **network weights during training**
- Variational inference is commonly adopted to approximate the intractable posterior
- At testing time, the BNN prediction is approximated by Monte Carlo sampling (with K samples, K = 10),

$$P(y|x;\theta) = \frac{1}{K}\sum_{k=1}^{K} \mathrm{softmax}(\hat{f}^{(k)}(x)).$$

Evaluation and Results

- Evaluation benchmarks (We also tested our model on public datasets with ground truth at the **video level** and compared with other methods as baselines)
    1. Real-life Trial (RLT [1]): Consists of 121 videos from real-life court room trials. This is an easy benchmark since video clips are short (10 seconds 1 minute).
    2. Bag-of-Lies (BoL [2]): Consists of 35 subjects, each of whom is shown 6-10 images and then being asked to describe them. Each participant is free to describe the image honestly or deceptively and the answer is recorded in a video (3.5 seconds to 42 seconds). The total number of samples in the dataset is 325 (163 truth, 162 lie). They also have EEG data.
    3. **Resistance Game (our group):** contains homogeneous and heterogenous games, very long videos (average length 46min), thus very sparse supervision and the most difficult and closest to real life dataset.
- Quantitative evaluation:
    - Average classification accuracy (ACC)
    - Average area under the precision-recall curve (AUC)

[1] Perez-Rozas et al., **Deception detection using real-life trial data**, ICMI 2015
[2] Gupta et al., **Bag-of-lies: A multimodal dataset for deception detection**. CVPR workshops, 2019

Evaluation and Results

Results on RLT[1]

Methods	ACC (%)	AUC (%)
TSN [2]	77.5	81.78
DDiV [3]	-	83.47
FFCSN [4]	89.16	91.89
Ours	**92.36**	**97.27**

- FFCSN [3] is the state-of-the-art deception detection method.
- For DDiV [2] and FFCSN [3], we use the numbers reported in the original paper.
- For TSN [1], we use the official implementation and run on RLT dataset.

[1] Perez-Rozas et al., **Deception detection using real-life trial data**, ICMI 2015
[2] Limin Wang et al., **Temporal Segment Networks,** ECCV 2016
[3] Wu et al., **Deception Detection in Videos**, AAAI 2018
[4] Mingyu Ding et al., **Face-Focused Cross-Stream Network for Deception Detection in Videos**, CVPR 2019

Evaluation and Results

Results on BoL[1] Most recent data on deception.

- **Limitation**: Their videos are short (max 42secs) not as revealing of deception and was meant to be combined with EEG. Only Indians i.e. single culture.

Method	ACC (%)	AUC (%)
LBP [2]	55.12	55.32
TSN [3]	56.94	57.62
Ours	**64.47**	**67.08**

- For LBP [1], we reimplemented the baseline. LBP serves as a naïve baseline.
- For TSN [2], we use the official implementation and run on BoL dataset.
- Accuracy lower than RLT since in their data deception behavior are not as prominent only from visual cues. Also Univ Students describing pictures and not a good choice for deception behavior not a natural interaction among people and not long video.

[1] Gupta et al., **Bag-of-lies: A multimodal dataset for deception detection**. CVPR workshops, 2019

Methods used in BOL paper in 2019

[2] Ojala, Timo, et al., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." *IEEE Transactions on pattern analysis and machine intelligence* 24.7 (2002): 971-987.

[3] Limin Wang et al., **Temporal Segment Networks,** ECCV 2016

Evaluation and Results

- Results on Resistance Game (ours)

Method	ACC (%)	AUC (%)
LBP [1]	49.56	49.56
TSN [2]	51.15	51.15
Ours	**71.08**	**71.08**

- For LBP [1], we reimplemented the baseline. LBP serves as a naïve baseline.
- For TSN [2], we use the official implementation and run on our Resistance Game dataset.
- Accuracies are higher than C3D last year (67.03%) and is slightly better than humans (around 70%).

[1] Ojala, Timo, et al., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." *IEEE Transactions on pattern analysis and machine intelligence* 24.7 (2002): 971-987.
[2] Limin Wang et al., **Temporal Segment Networks,** ECCV 2016

Question 2: Localizing and Interpreting Deception Behavior

- Adapt Grad-CAM [1] to find the find the attention of the model in the time domain (this way AUs and their durations/locations can possibly be detected)
- For positive (deception) samples we can compute the key time-steps for the decision of the detection model
- Utilize the gradient of the model w.r.t. a feature layer and Aus
- Interpret Deception based on presence of AUs

[1] Selvaraju et al., **Grad-Cam: Visual explanations from deep networks via gradient-based localization,** ICCV 2017

Attention Method

- Attention mechanism for TCN:
  - A novel Grad-CAM based attention -- flexible and does not require architectural modification to the model

- Attention based on gradients of network output Y^c (Deceiver or Truth-teller) wrt AUs at a particular level of the network with k features F^k at a time t. We also combine with attention at a given channel level (gray curve in the results)

$$A_{ch}^{c} = \frac{1}{Z} ReLU\left(\sum_{k}\sum_{t} ReLU\left(\frac{\partial Y^c}{\partial F_t^k}\right) F^k\right)$$

- Positive gradient at a specific location implies increasing the intermediate feature value in F^k results in a positive impact on the prediction score

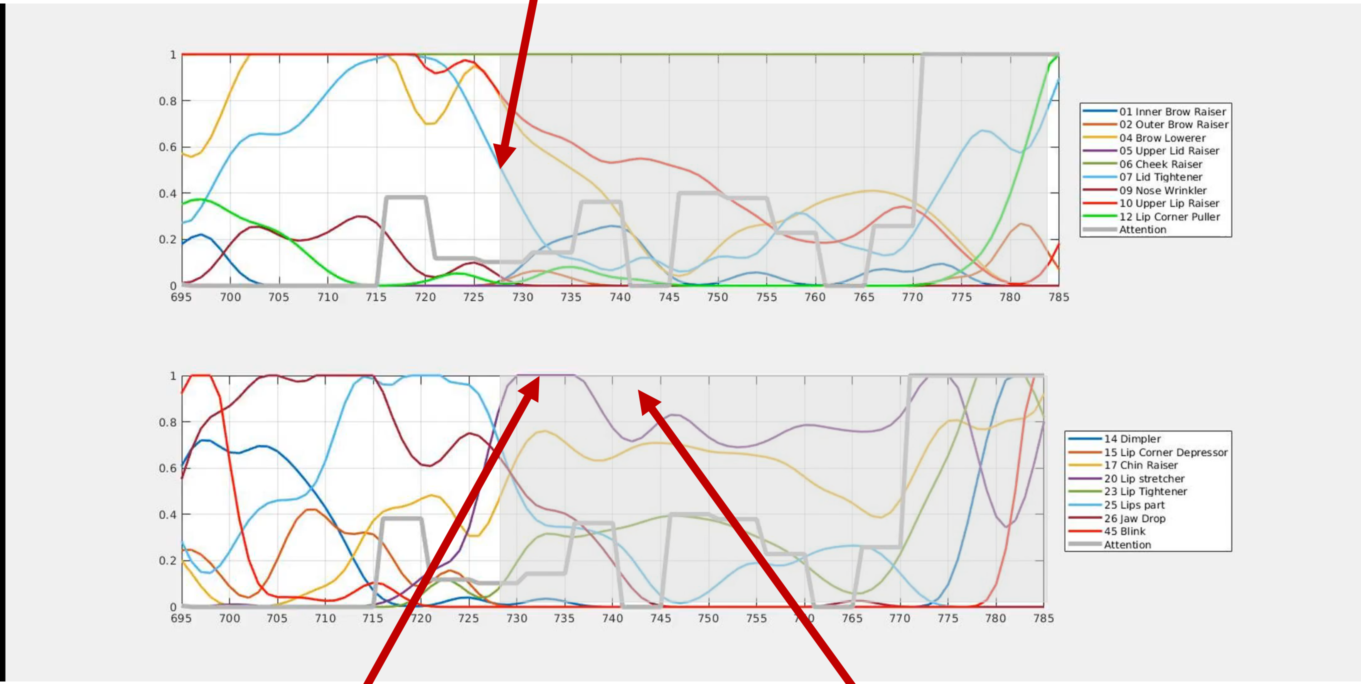
$$A_{AU}^{c} = \frac{1}{Z} ReLU\left(\sum_{k}\sum_{t} ReLU\left(\frac{\partial Y^c}{\partial F_t^k} \frac{\partial F_t^k}{\partial AU_t}\right) AU\right)$$

where Y^c is network output for class c (deceiver vs truth-teller), AU is an action unit. We pick those Action units above a threshold and when the given level A^c is also high.

Attention continued (RLT)

- Trial lie 011

Gray curve indicates attention score



AU20 Lip stretcher activated

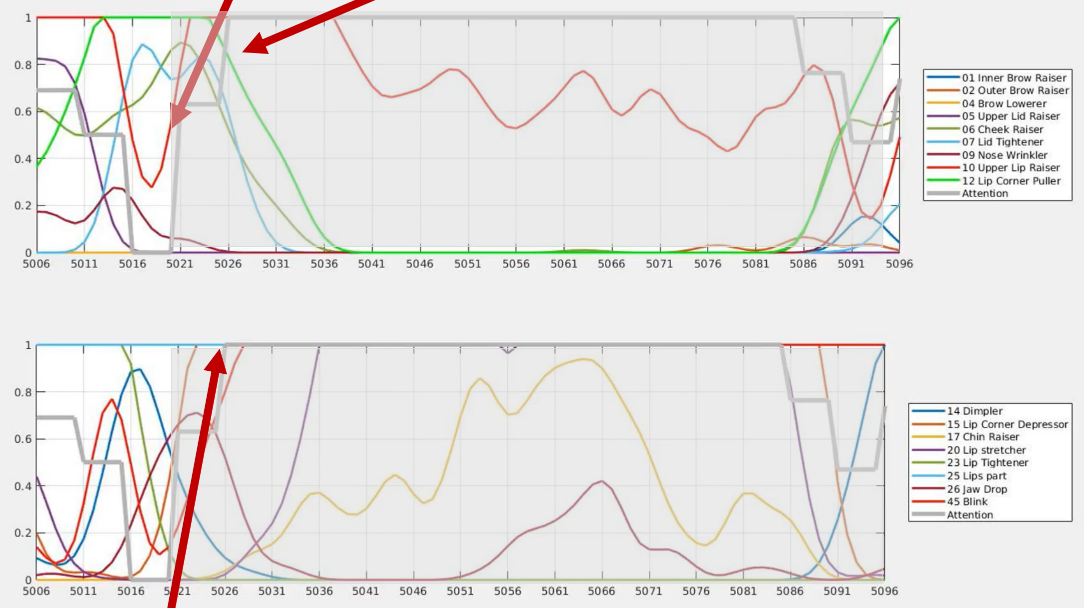
AU45 Blinks activated

Attention based on our Data. Location: US, AZ

- AZ-005 1

Gray curve indicates attention score

AU10 Upper Lip Raiser, both activated



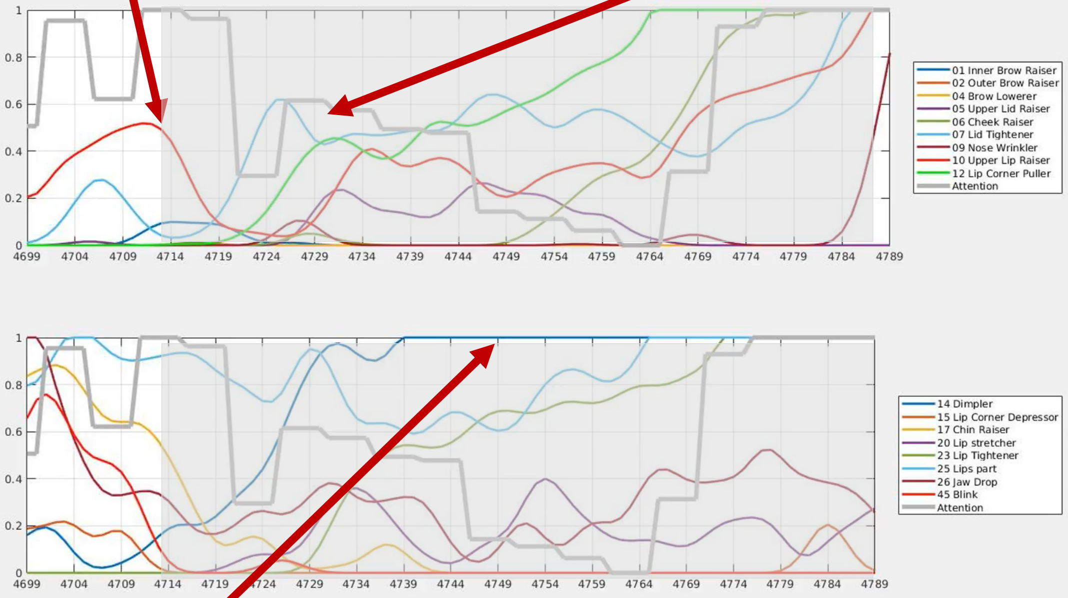
AU45 Blinks activated

Attention based on our Data. Location: US, AZ

- AZ-005 1

Gray curve indicates attention score

AU09 Nose Wrinkler
AU10 Upper Lip Raiser, both activated

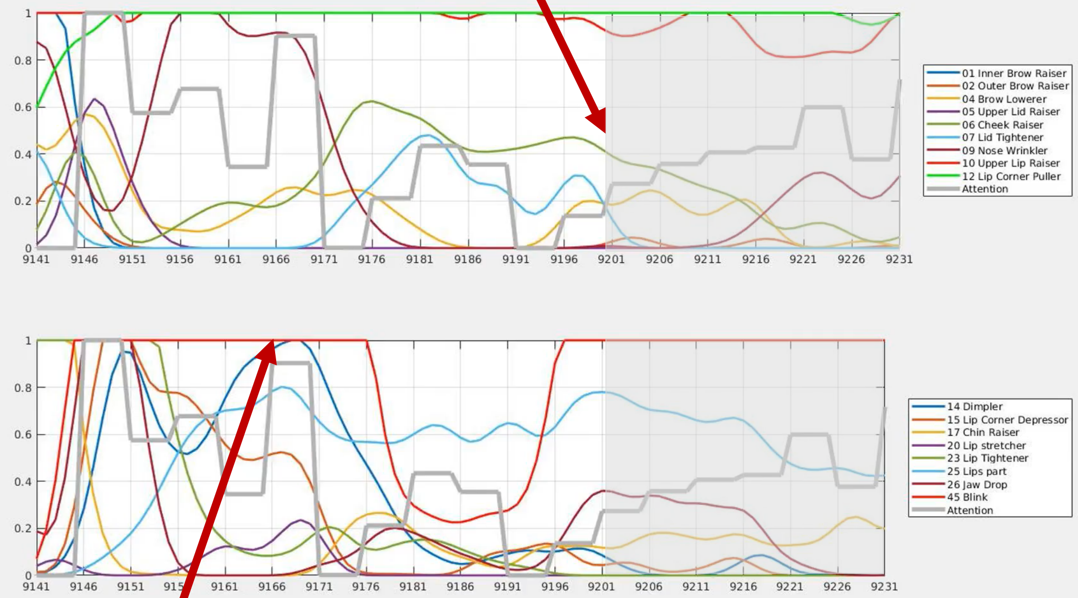


AU45 Blinks activated

Attention based on our Data. Location: US, AZ

Gray curve indicates attention score

- AZ-015 2

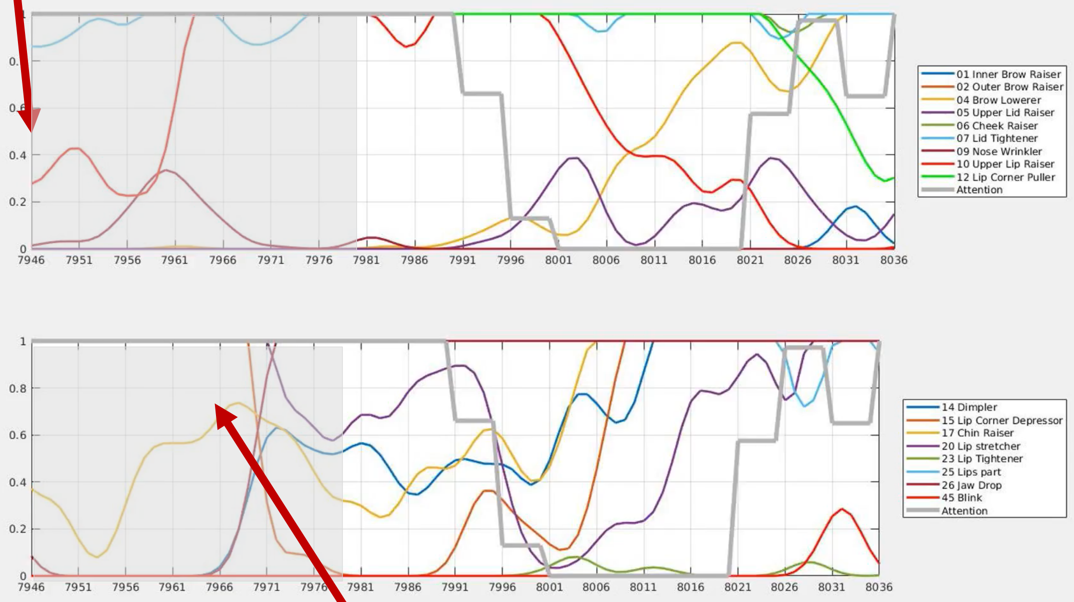


AU45 Blinks activated

Attention based on our Data. Location: Singapore, NTU

- NTU-003 2

Gray curve indicates attention score

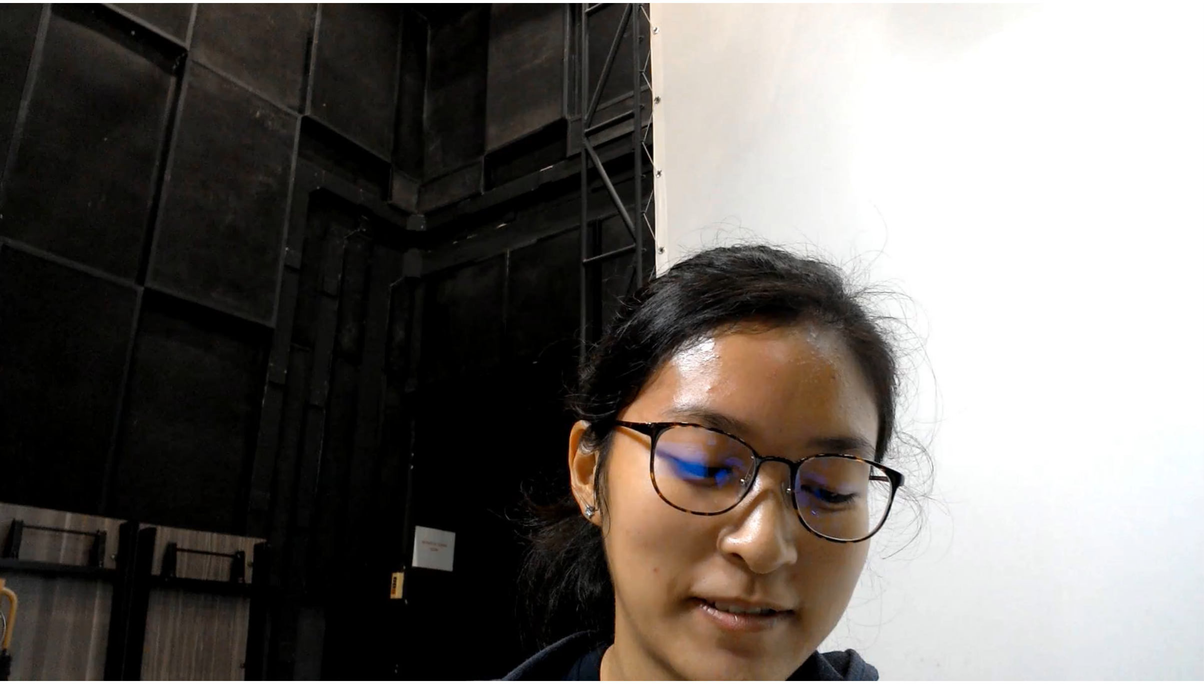


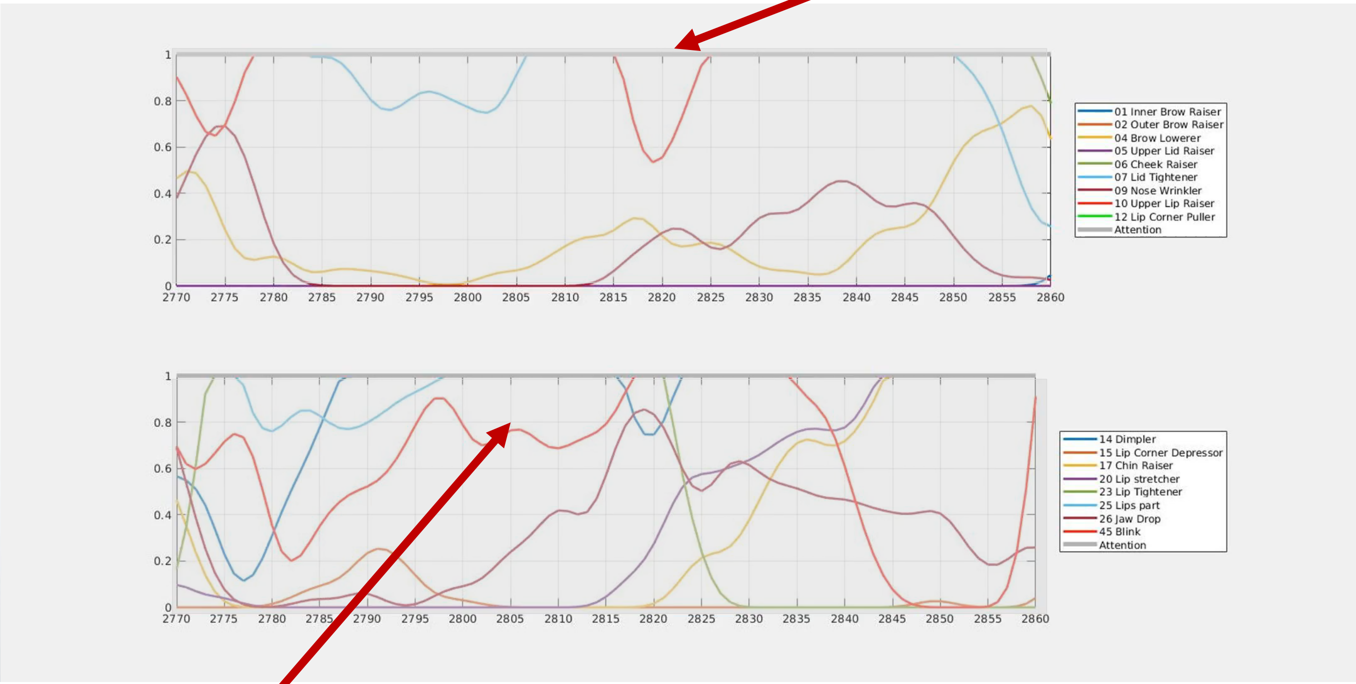
AU20 Lip stretcher activated

Attention based on our Data. Location: Singapore, NTU

- NTU-003 4

AU12 Lip Corner Puller





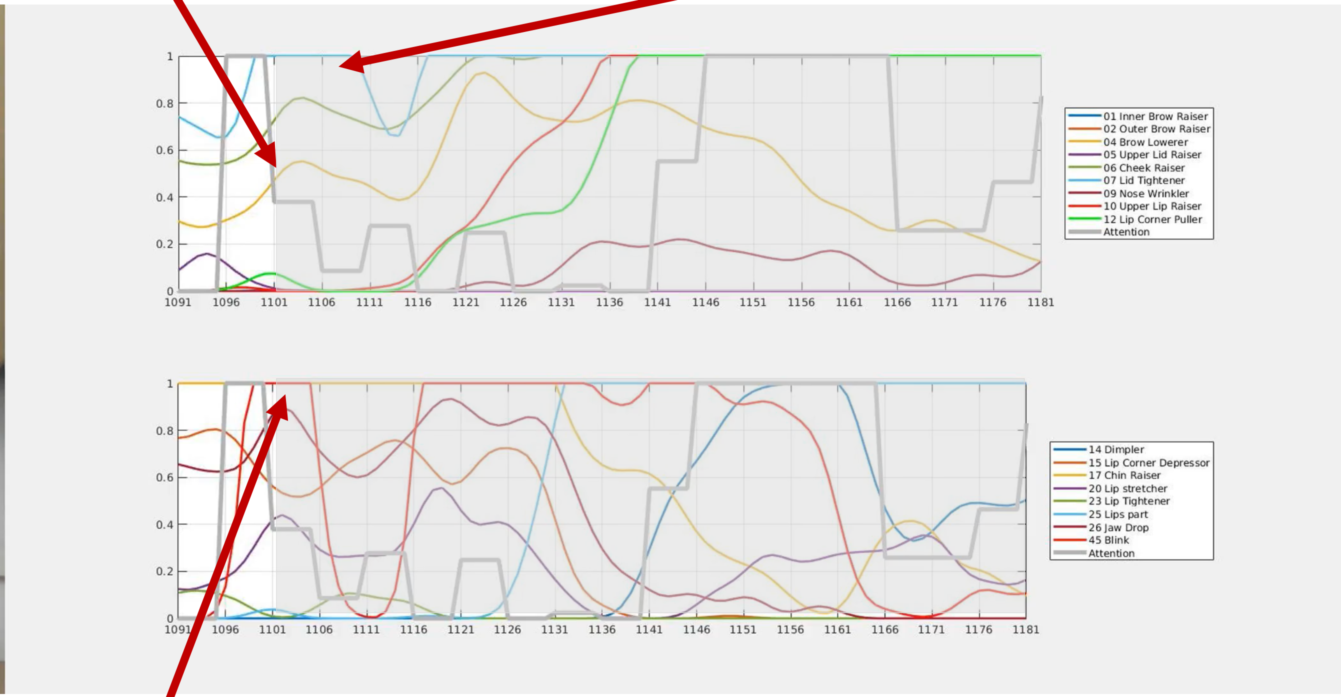
AU45 Blinks activated

Attention based on our Data. Location: Israel

Gray curve indicates attention score

AU12 Lip Corner Puller

- ISR-001 2

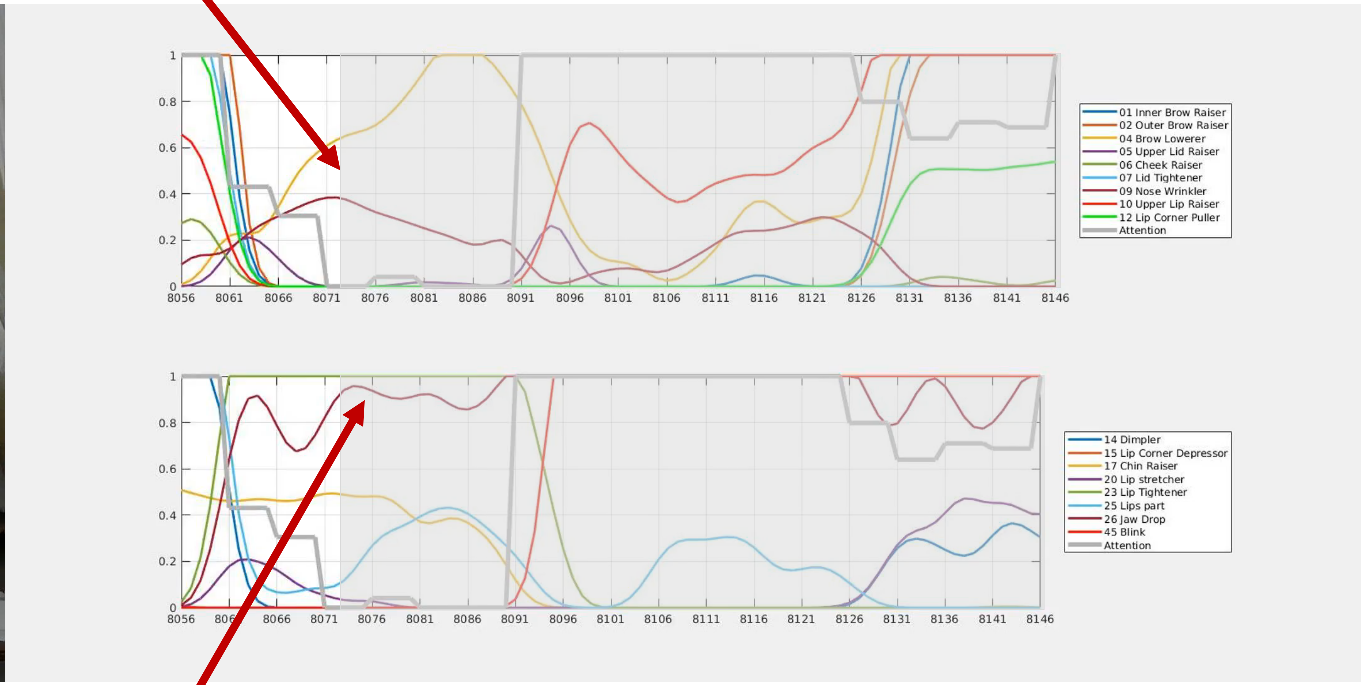


AU45 Blinks activated

Attention based on our Data. Location: Israel

Gray curve indicates attention score

- ISR-001 3



AU45 Blinks activated

Summary of Collaborative Research

- We have developed a General ML Framework for Detecting, Localizing, and Interpreting Deceptive behavior from Video

  - Providing high-level information (AUs) to the model helps

  - Do not model pixel-level nuances

- Framework for retrospective analysis of deception Based on Attention

- Our Framework is General and can be use for other tasks such as Trust

- Our Framework will also be extended to include other cues such as audio and can be integrate with the Dartmouth-Stanford research.

Publications

- **Publications related to attention-based facial analytics**:
  - Wang, L., Bai, C., Bolonkin, M., Burgoon, J.K., Dunbar, N., Subrahmanian, V.S, and Metaxas, D., ‘Attention-based Facial Behavior Analytics in Social Communication’, the British Machine Vision Conference (BMVC), September 2019 [**BMVC 19**]
  - Wang, L., Yu, X., Bourlai, T., & Metaxas, D. N. (2019). A coupled encoder–decoder network for joint face detection and landmark localization. Image and Vision Computing, 87, 37-46.
- **Publications related to facial analysis**
  - We propose face data augmentation which generate faces in different views, which is able to generate more data for training.
    - Anastasis Stathopoulos, Ligong Han, Norah Dunbar, Judee K. Burgoon, and Dimitris Metaxas. "Robust Facial Features for Deception Detection in Videos." *Future Technologies Conference (**FTC**) 2020.*
    - Zhao, Long, et al. "Towards Image-to-Video Translation: A Structure-Aware Approach via Multi-stage Generative Adversarial Networks." ***International Journal of Computer Vision*** (2020).
    - Han, Ligong, et al. "Robust Conditional GAN from Uncertainty-Aware Pairwise Comparisons." ***AAAI***. 2020.
  - We improve the face tracker system to have more stable face detection and landmark localization, which is the important preprocessing steps for facial behaviour analysis.
    - Wang, L, Xiang, Y,. Thirimachos, B., and Metaxas, D., “A coupled encoder–decoder network for joint face detection and landmark localization.” Image and Vision Computing 87 (2019): 37-46. **[journal paper @ IVC’19]**

Other Activities

- **Joint Team Course**
  - 2 Lectures on Faces and our ML Methods fro Deception Detection

- **Invited Talks (Universities, Major Conferences ECCV 2020)**

- **Chapter in a Springer Book by our Group**

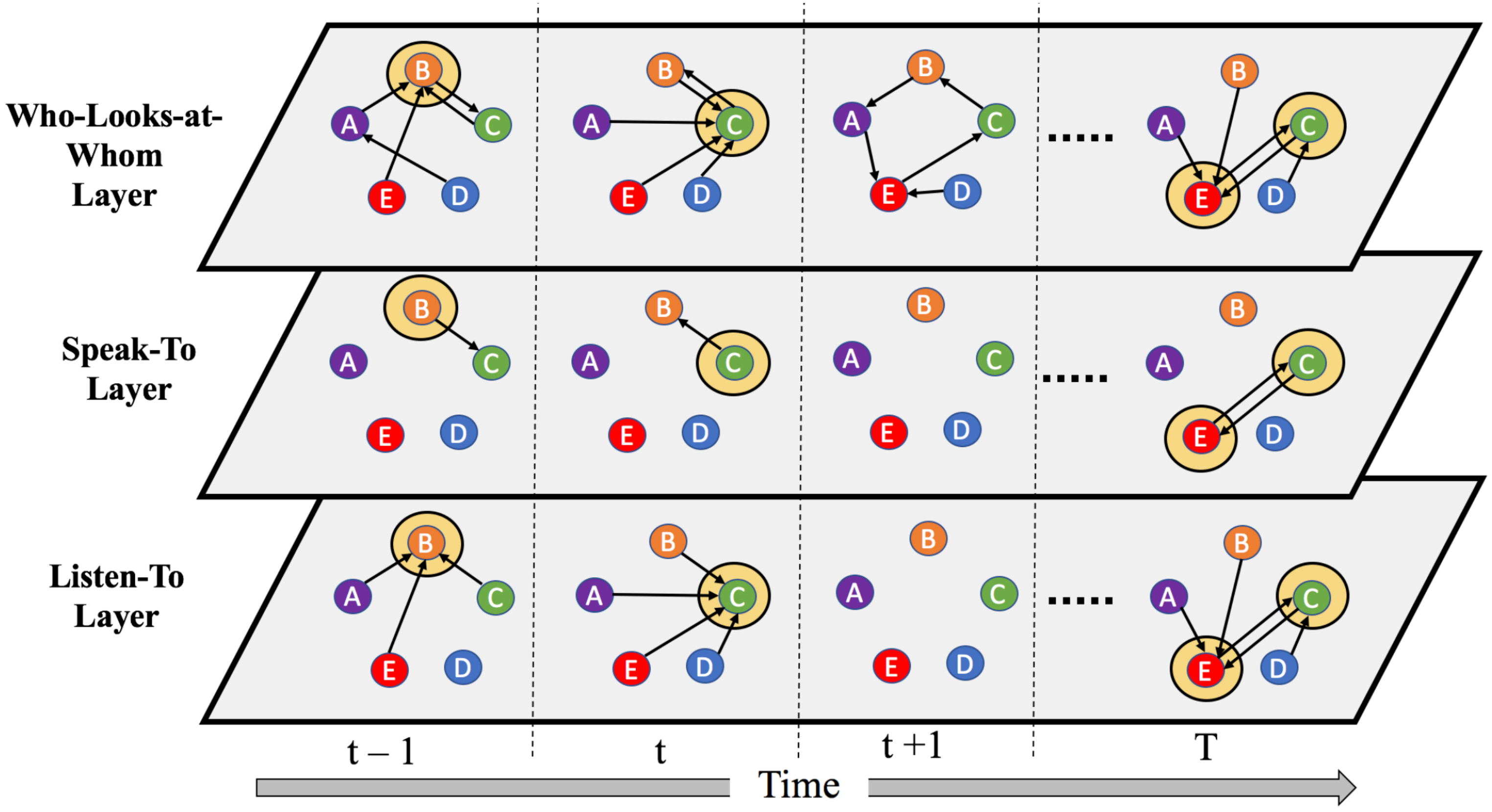
- **Graduated PhD Students: Zachary Daniels May 2020, SRI Princeton, NJ**

Deception detection in Face-to-face Dynamic Networks

Srijan Kumar, Georgia Tech
Chongyang Bai, Dartmouth College
Jure Leskovec, Stanford University
V.S. Subrahmanian, Dartmouth College

Deception detection in Face-to-face Dynamic Networks

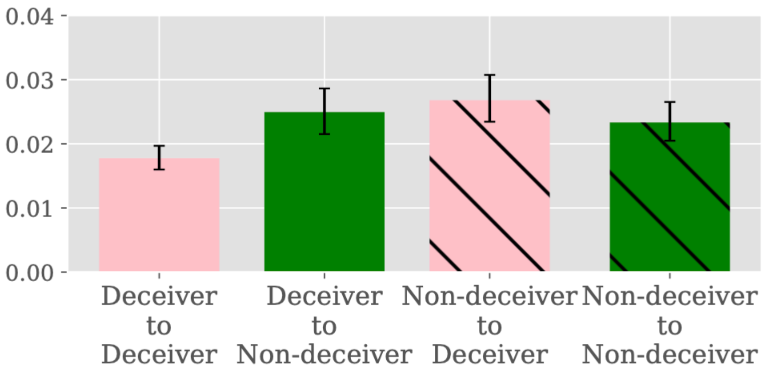
- **Look-at**
  - Edges: who-look-at-who in each time step
  - Weights: $E_G^t(u, v)$. Probability of u looking at v.
- **Speak-to**
  - Edges: look-at edges a speaker.
  - Weights: probability of a speaker looking at v.
- **Listen-to**
  - Edges: Incoming look-at edges towards a speaker.
  - Weights: probability of looking at a speaker.



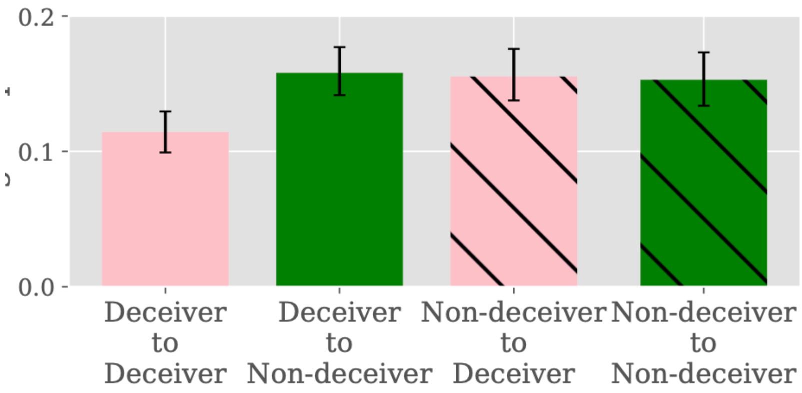
Deception Behavior Analysis

- RQ: Do deceivers interact differently with other deceivers *vs* non-deceivers?
  - Deceivers avoid non-verbal interactions with other deceivers (left and middle).

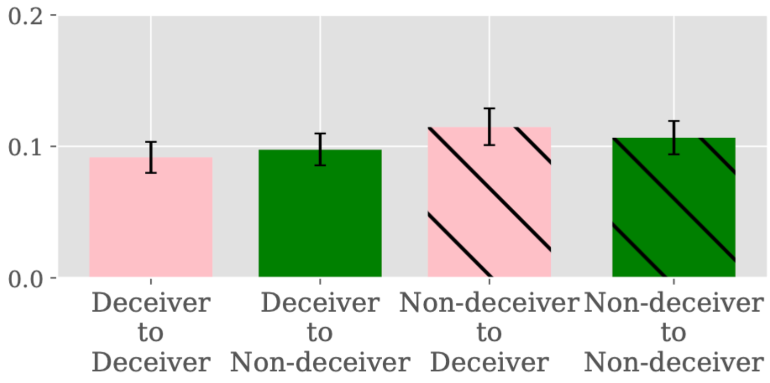
Reciprocity



Probability of listening to speaker



Probability of speaking to



Deception Detection – Negative network

- Deceivers avoid non-verbal interactions with other deceivers.
  - Negative network

$$N^{t,-} = (V, E^{t,-})$$

$$E^{t,-} = \{1 - w_{u,v,t} | \forall u, v \in V, u \neq v\}$$

Deception Detection – Negative network

- Deceivers avoid non-verbal interactions with other deceivers.
  - Features: verbal and non-verbal behaviors of people. Normalized to [-1,1].
    a) fraction of speaking
    b) average entropy of looking
    c) average in-degree
    d) average in-degree while speaking
- Define the **deceptive score** s(u) = 4 – ((a) + (b) + (c) + (d))

From our analysis,
If u is a deceiver, s(u) tends to larger!

BPNN: Belief Propagation on Negative Network

Algorithm 1 Belief Propagation on Negative Network

Input: Dynamic network $N = (V, E^1, E^2, ..., E^T)$, initial deceptive scores $s(V)$, β, convergence threshold τ, maximum iteration number M

Output: Final deceptive scores $s(V)$

```
iter = 0, dif = 1;
while dif > τ and iter < M do
    s^t(V) = s(V), ∀t = 1 ... T;
    for t = 1 ... T do
        foreach v ∈ V do
            r^t(v) = β Σ_{(u,v)∈E^t} s^t(u)(1 − w_{u,v,t}) + (1 − β)s^t(v);
    r(V) = Σ_t r^t(V)/T;
    r(V) = r(V)/‖r(V)‖_2;
    dif = ‖(r(V) − s(V))‖_2;
    iter = iter + 1;
    s(V) = r(V)
return s(V)
```

- Compute initial deceptive score s
- Belief propagation in negative network until convergence
  - Update s(u) by the weighted average of u's neighbors.
  - Average over time.

Deception Detection - Experiments

Method	Performance	% Improvement Over Baseline
Computer Vision Baselines		
Emotions (Bai et al. 2019a)	0.538	39.9%
Movements (Baltrusaitis et al. 2018)	0.549	37.2%
FAUs (Demyanov et al. 2015)	0.569	32.3%
LiarRank (Bai et al. 2019a)	0.590	27.6%
Multimodal (Wu et al. 2018)	0.594	26.7%
Graph Embedding Baselines		
TGCN on Look-At (Liu et al. 2019)	0.550	36.9%
TGCN on Speak-To (Liu et al. 2019)	0.538	39.9%
TGCN on Listen-To (Liu et al. 2019)	0.541	39.2%
Ensemble Baseline		
Late fusion of all above	0.623	20.9%
Proposed Method		
BPNN	**0.753**	-

- Setup:
  - Training and test sets have different games.
- Our BPNN: Logistic regression using
  - ➢ **Converged Deceptive score.**
  - ➢ 1 - fraction of speaking
  - ➢ 1 - average entropy of looking
  - ➢ 1 - average in-degree
  - ➢ 1 - average in-degree while speaking

[Bai et al., 2019a] Automatic Long-Term Deception Detection in Group Interaction Videos
[Baltrusaitis et al. 2018] Openface 2.0: Facial behavior analysis toolkit.
[Demyanov et al. 2015] Detection of deception in the mafia party game.
[Wu et al., 2018] Deception detection in videos.
[Liu et al., 2019] Characterizing and forecasting user engagement with in-app action graph: A case study of sna

Ongoing Collaborative work

- **Semantic analysis, Multi-modality (in collaboration with UA & UCSB & Dartmouth)** on the MURI transcripts.

  - Turn-at-talk transcripts of 40 games for improved deception detection

  - Preliminary analysis: it is possible to discriminate game roles with word counts (bag-of-words), simple Logistic regression model achieves accuracy 0.70 and F1 score 0.65

- Integrate Work of Dartmouth-Stanford to the Rutgers-UA-UCSB-Dartmouth framework as added features

- Initial Work on Trust and Dominance (collaboration with the other groups)

- Further collaboration with ARO and researchers at Army Research Labs (e.g., Aberdeen)

Contact Information

Dimitris N. Metaxas
Director CBIM Center
Department of Computer Science
Rutgers University
dnm@cs.rutgers.edu
www.cs.rutgers.edu/~dnm
Tel: 848-445-2914