cognoa

Deep learning for Developmental Delay Screening in Children

Ford Garberson, Cognoa Inc

cognoa

Introduction

- Cognoa App: ML algorithms screen for autism
  - Clinical validation: higher performance than standard screeners
    (See PAW healthcare talk)
  - Has been used by ~250,000 families
  - R&D into screeners for different conditions

- Outline:
  - Summarize our current algorithms
  - DL algorithms with audio + video in interactive stories
  - DL algorithms with cell phone phone videos uploaded by parents

Part 1: Introducing Cognoa

cognoa

The Cognoa autism screening App

Three sources for ML predictions of autism:

- Parental questionnaire
- Optional: short videos of child uploaded by parent
  - *Currently:* human analysts watch videos and answer questionnaire
- Optional: pediatrician questionnaire

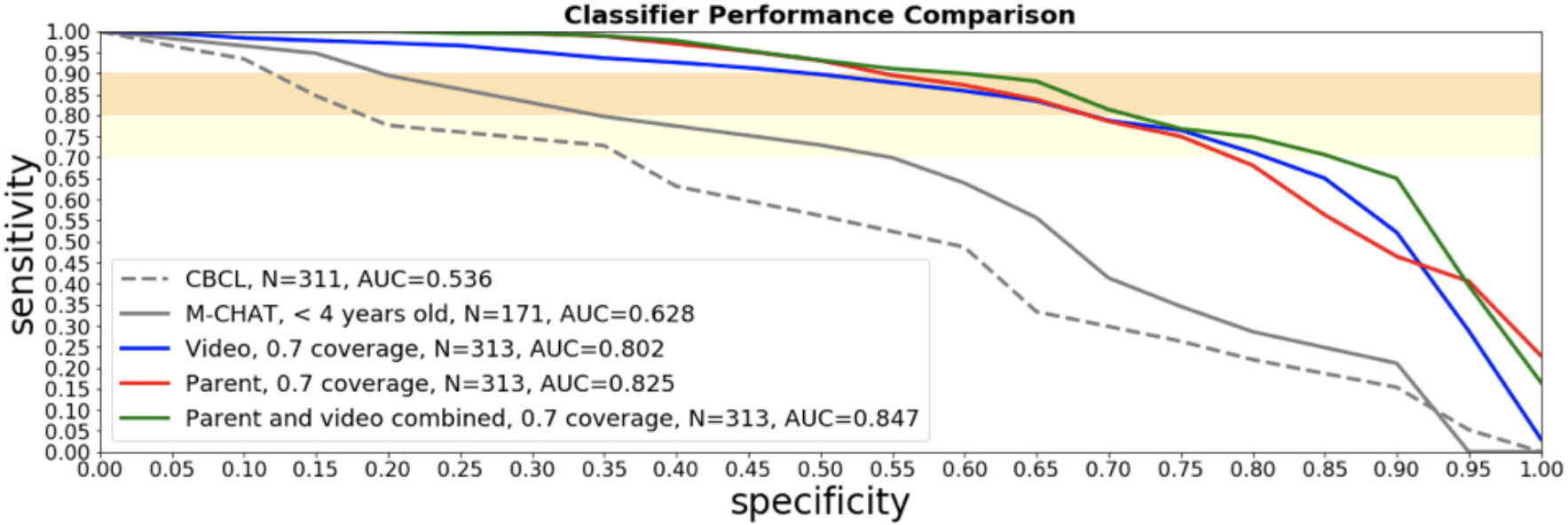
cognoa

How well does our screener perform?

Clinical study of 313 high risk children ages 1.5 to 6

- Gray curves: commonly used, non-ML based autism screeners
- Colored curves: our algorithms

More details in PAW Healthcare talk



Part 2: R&D for deep learning based interactive instrument

Beyond screening tools: interaction with children

All algorithms so far: passive questionnaires and videos

Interacting directly with the child: increase sensitivity

- Information no other instrument can access
- Probe new conditions
- Therapies?

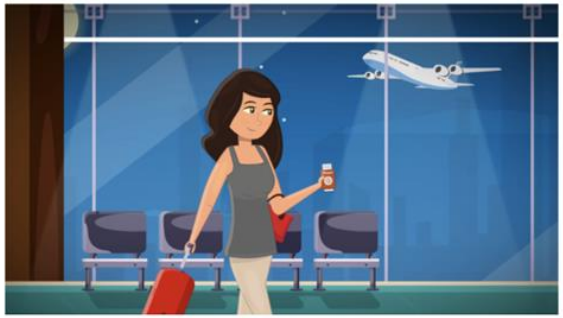
cognoa

Tablet storytime

Child watches cartoon story on tablet at home

- Frequent pauses to ask child questions
- Record the *audio* and *video* streams from the tablet for analysis

"Hello, my name is Susan. Today I am going to tell you a really cool story. Are you ready for it? Tell me as loud as you can, are you ready?! Okay, let's listen to it now!"



One night, a young woman was walking through the airport...

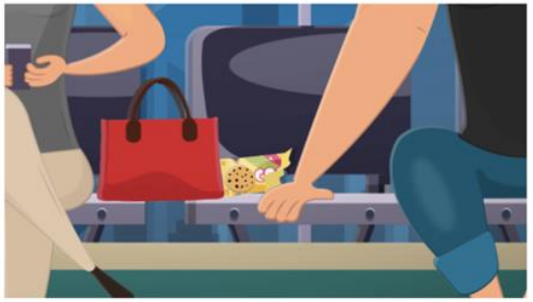
"I hope you enjoyed the story!! Let's see if you remember what happened. Where did the story happen?"

cognoa

Storytime comprehension

Can child remember the story and repeat it?

"Alright. Now tell the story again to your mom or dad! You can do it!"



cognoa

Storytime creativity

Can child make up their own story?

"Here's a picture for you to tell your own story. Where are they and what are they doing?"

cognoa

POC trial

76 children 4 to 6 years old from general population

- 11 with speech/language impairments
- 60 normally developing

By-hand analysis:

- Can discriminate between kids with language problems and normal kids

cognoa

Deep learning algorithm potential

Progress tracker for normal children:

- Ability to focus, vocabulary & grammar sophistication, ...
- Comprehension, short term memory, creativity, ...

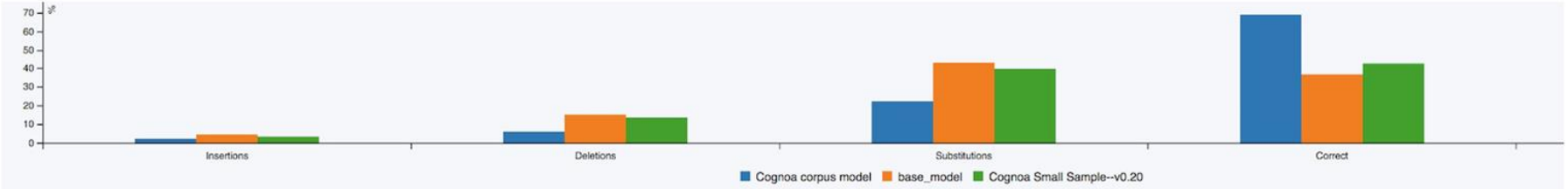
When more data for children with developmental problems available:

- Identify speech and language disabilities (prosody analysis)
- Autism and ADHD are challenging
    - But each have distinguishable features

Improved speech to text algorithm

Partner's STT algorithm: IDs ~35% of words, many mistakes

- Update corpus to child's vocabulary: performance ~2x



Cognoa corpus model ***With new corpus***

MEDIA FILE	SENTENCES	WORDS	INSERTIONS(%)	SUBSTITUTIONS(%)	DELETIONS(%)	CORRECT(%)	WORD
P1_01.wav	9.0	46.0	2.2	19.6	2.2	78.3	23.9
P1_03.wav	19.0	117.0	4.3	22.2	8.5	69.2	35.0
P1_02.wav	22.0	131.0	1.5	24.4	6.1	69.5	32.1
Totals	**50**	**294**	**2.38**	**22.45**	**6.12**	**69.05**	**31**

base_model ***Base model***

MEDIA FILE	SENTENCES	WORDS	INSERTIONS(%)	SUBSTITUTIONS(%)	DELETIONS(%)	CORRECT(%)	WORD
P1_01.wav	9.0	46.0	6.5	41.3	17.4	41.3	65.2
P1_03.wav	19.0	113.0	7.1	47.8	13.3	38.9	68.1
P1_02.wav	22.0	131.0	2.3	41.2	16.0	42.7	59.5
Totals	**50**	**290**	**4.83**	**43.45**	**15.17**	**36.55**	**63**

cognoa

Future storytime plans

Immediate: keep improving speech to text algorithm

Build deep learning features from:

- Vocabulary and grammar analysis
- Prosody analysis of speech patterns (clues of impairments)
- **Video analysis**:
  - Where is child's attention?
  - Eye tracking: strong signals for autism

Part 3: Deep learning analysis of videos

cognoa

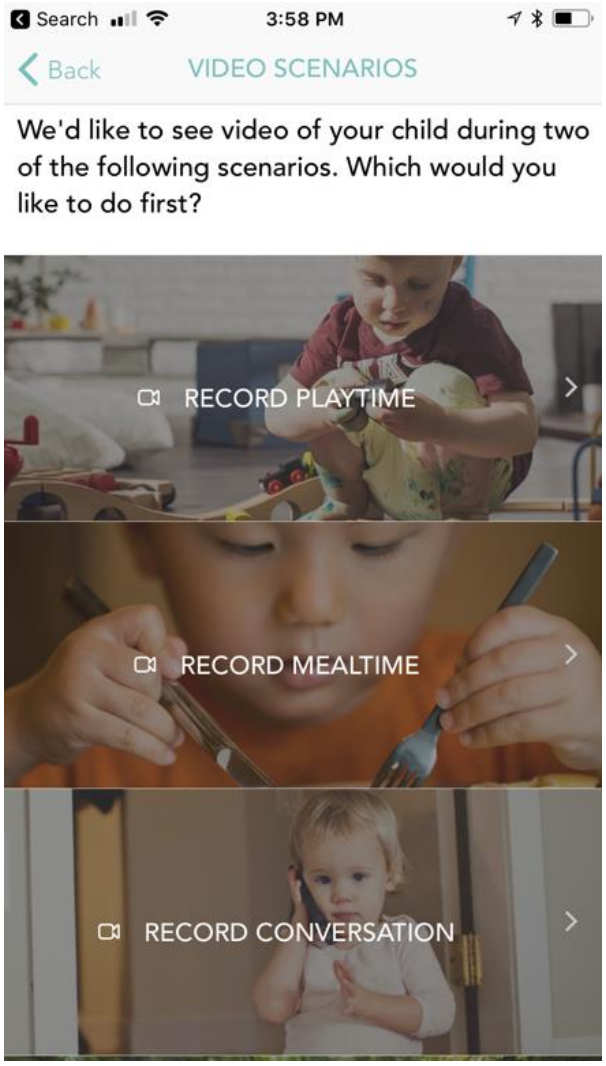
The cognoa video screening algorithm

Parents upload a two or three ~2 minute videos of child

Multiple analysts watch videos

- Analyst answers to questionnaire pass to ML algorithm

Key challenge: does not scale well



cognoa

How could deep learning help us?

Option 1: do tasks of video analysts

- Algorithm uses videos to answer questionnaire
- Preprocessing step to our autism prediction ML algorithm

Option 2: end-to-end identification of autism

- Algorithm uses videos to directly predict autism

Focus on the simpler option 1 in this talk

cognoa

Test on an example feature

Does the child make appropriate eye contact with others?

- Can we train deep learning algorithm to answer this?

Available videos with proper labels: 1500 negative, 1100 positive

- 80/20 train/test split

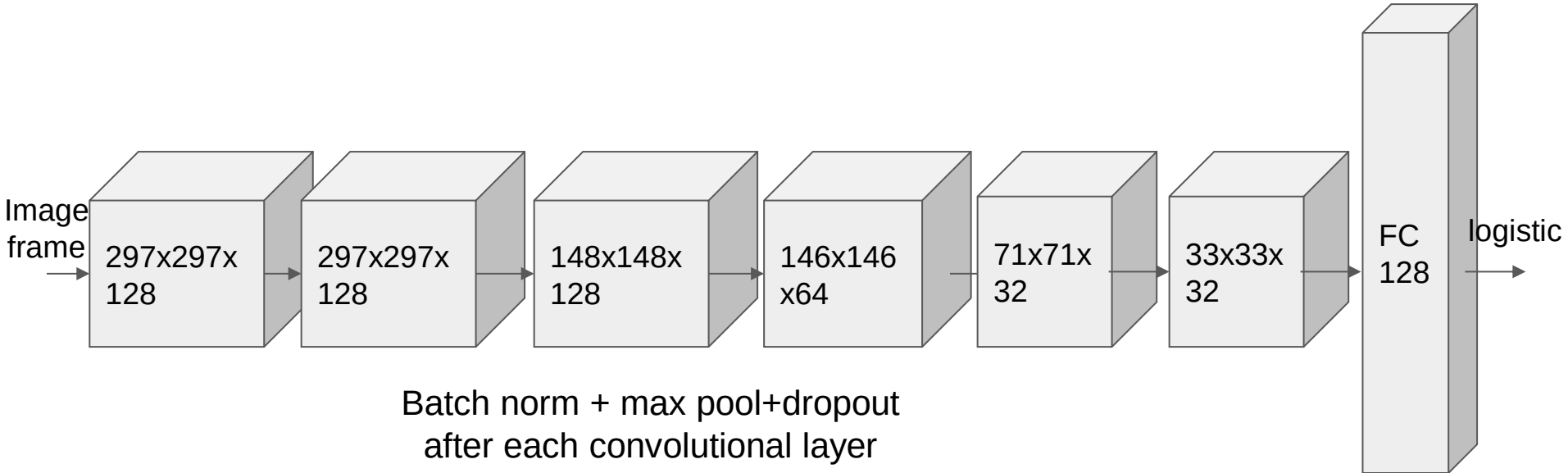
Experiment with different algorithms to predict eye contact

cognoa

Method #1: 2D CNNs

Split each video into 75 frames

- Train CNN on individual frames



cognoa

Method #1: 2D CNNs

Split each video into 75 frames

- Train CNN on individual frames
- Average performance over each frame

Test set accuracy: ~72% in no eye-contact case, ~62% in eye-contact case

(Only using single frames in algorithm)

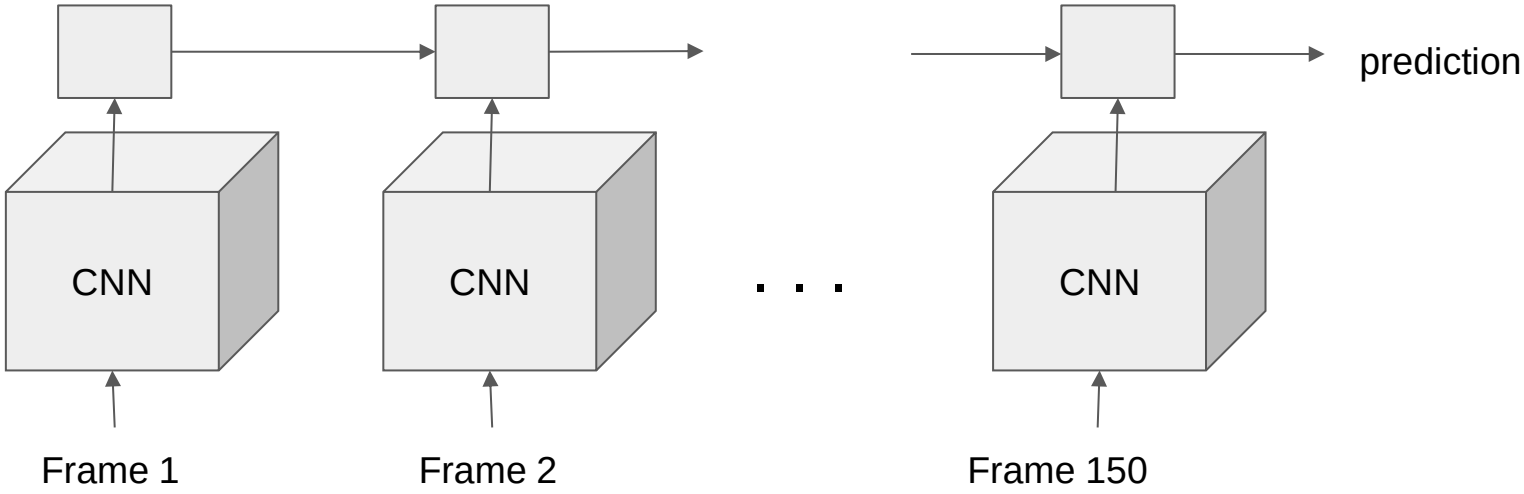
cognoa

Method #2: CNN + LSTM

Example: Donahue et al: https://arxiv.org/abs/1411.4389

Split video into 150 frames

- Each frame fed to 2D CNN
- Output to LSTM that crosses all frames



Method #3: 3D convolutional neural network

Naturally incorporates time dimension

Drawbacks:

- Models for transfer learning?
  - Probably have to project models into 3D
- Very slow to train

cognoa

Method #4: Two-stream networks

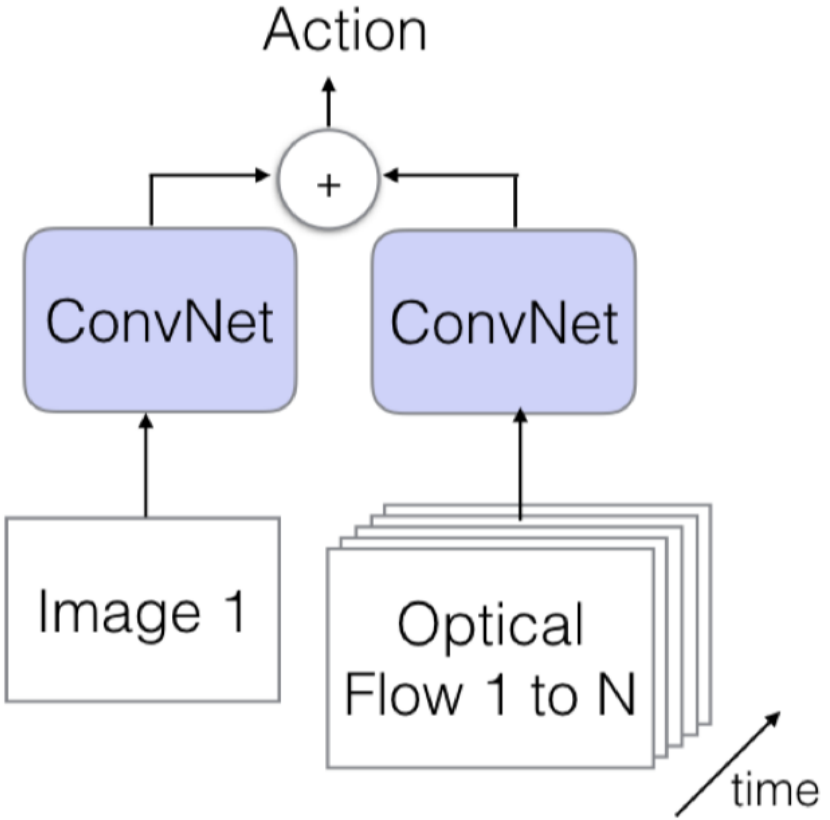
Snapshots of videos + optical flow

- J Carreira and A Zisserman: https://arxiv.org/pdf/1705.07750.pdf

Stream 1: CNN on snapshot

Stream 2: snapshot + optical flow → many images → CNN

Aggregate predictions from each stream



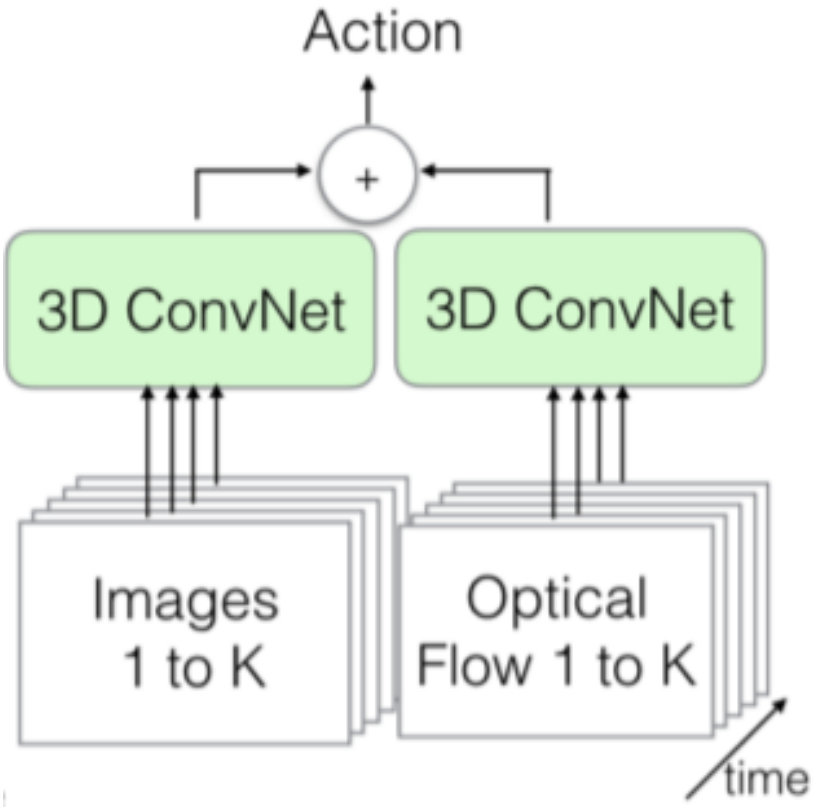
cognoa

Method #5: Two-stream inflated 3D CNNs

Promote 2D CNN to 3D

- Extrapolate from 2D CNN for transfer learning
- One stream: snapshots of real images
- Other stream: optical flow to improve those snapshots

Aggregate predictions from each stream

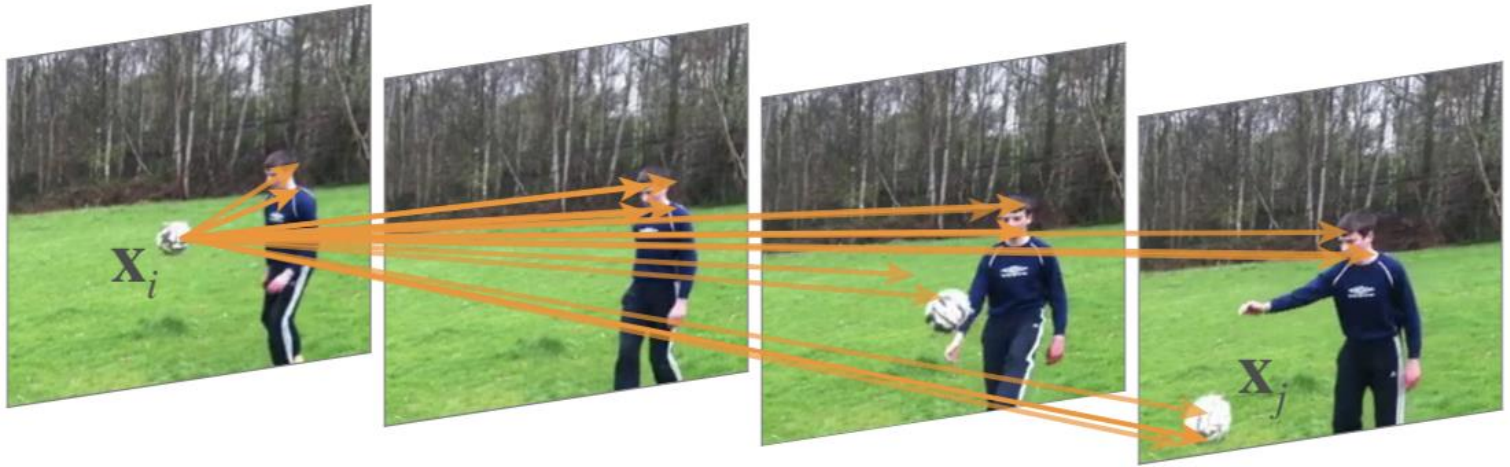


cognoa

Possible Method #6: Non-local neural network layers

Calculations between every combination of points in video spacetime

- Great for long-range dependencies
- Insert as a layer to any CNN
- Surprisingly fast



From X Wang et al: https://arxiv.org/pdf/1711.07971.pdf

cognoa

The future

Cognoa ML algorithms already identify autism better than conventional screening tools

- But many new challenges lie ahead

Our R&D with deep learning has potential to build

- Revolutionary screeners
- Progress tracking tools
- Hopefully (therapies)

Bonus slides

cognoa

Why autism?

- Autism rates rising
  - 1.7% as of 2014, but maybe higher today
  - Some studies: higher rates in children of tech and academic workers

- Delayed diagnosis: treatments miss critical time when child's brain is developing
  - Autism can be diagnosed at 1.5 years, but …
  - Average wait time for diagnosis: 1.5 years
  - Average diagnosis age: 4.2 years