Identifying Interesting Moments in Controllers Work Video via Dimensionality Reduction

Kristofer Krus
Tatiana Polishchuk
Valentin Polishchuk

LINKÖPING UNIVERSITY

Support:

TRAFIKVERKET
SWEDISH TRANSPORT ADMINISTRATION

ATCO workload

Plethora of work on

- taskload
- fatigue
- physical/physician's measurements
  - iris dilation
  - heartrate
  - ...
- complexity perception
- ...

Important subject
(see papers in ATM outlets)

To study it
better need to know better
what ATCOs do

And even if you **are** an ATCO,
a view from "outside" yourself
might be nice

The Holy Grail

Longer-term goal:

correlate

- "outside, objective"

observations/views

with

- "personal, subjective"

measurements/perceptions

of work ⇨

model of ATCO workload (?!?)

Wanted:

- "outside"
- observer
- "objective"

view of

- work place
- out-of-the window
- RWY (in airport)
- ...

Field experiment

Tower in Stockholm
- Bromma airport
- 4-27 movements/day

March 2019
- winter(!) day

Videotaped ATCOs work

- Less intrusive than heatrate sensing, VR, etc
- More info (ATCOs+ planes+snowblowers+...)

(thank you, LFV!)

Next 8 slides:
- look @
- photos of 4/5 cameras looking @
- ATCOs work looking @
- monitors and RWY

+

views from the 5 cameras

Top secret

Collected data (few hrs of recording x 5)

Done? No, need to *watch* through

ANSP side business: sleep helpers

- ~~Plane spotting~~
- ~~Pushing tin~~
- ~~Ground control~~
- ~~Die Hard 2~~
- ~~...~~

Appreciate ATCO fatigue:
(fatigue from watching ATCOs)

No offense:
most of the time

- nothing happens
- routine

Au contraire

- Boring is safe!
- Good job!

We want the opposite

Interesting moments

- rare
- off-nominal
- non-routine
- higher workload
- smth unusual
- …

Accuracy isn't an issue

- false-positive are fine
- false-negative are OK

May be interesting also for

- training
- incident investigation
- superivision
- ...

Our (quick) solution to handle the annoying bottleneck of video browsing?
(HInt: we're at a University)

Use the brains

C. Massinger and H. Willers

Analysis of mental workload for RATCO while handling traffic flows

B. Sc. thesis

Linkoping University, 2019

Supervisors:
T. Polishchuk and C. Schmidt

Manual

- scene analysis
- video synchronization
  (cameras slightly different)
- identification of
  interesting moments
- ...

~~Comme il Faut~~

Our contribution

Let computer watch the video

Artificial intelligence (AI)

Neural network (NN)

Convolutional NN (CNN)

- Automatic identification of unusual moments
- GUI for human fun & control (if desired)

Feature preview

- No training
- No human oversight
- Standard tools
- Open source
- ...

A vanilla ML paper

Split video data

- training
- testing
- (validation)
- …

Choose the type of NN
Create (initialize) NN

- train
- test
- validate

Report how good NN
vs humans
or earlier NNs

+

extras:

- count coffee drunk
- identify rats/cats in frames (if any)
- replace controllers faces
(deep fake)
- et al. reported successes of AI…

Our paper is different

Issue with standard approach:

to be trained,

NN needs *labelled* data

- interesting frames vs
- boring frames

or whatever the labels are in our case

Who'd label the frames?

Humans?

Begging the question :(

Our idea:

- stand on the shoulders of giants
- take *existing* NN
- for image classification
- fiddle with it

to make it work for our purposes

Next couple of slides:

technical details...

CNN: a recent breakthrough

EfficientNet is one af the fame
in fact, we took smallest version
EfficientNet-80

- among best image classifiers
  > human
- pretrained on ImageNet dataset
  state-of-the-art collection
- multi-$$-to-train
  freely available

Little issue:

- routine situation in ATCO work
- unusual moment in ATCO work

~~among~~ the many labels
(cats, guitars, mountains, ppl, ...)
of ImageNet

The glorious EfficientNet classifier
never saw an ATCO (I suspect)
(shame on NN community)

BlackBox use: not quite what's needed

CNN classifier

~~car~~

~~piano~~

...

mountain

~~runner~~

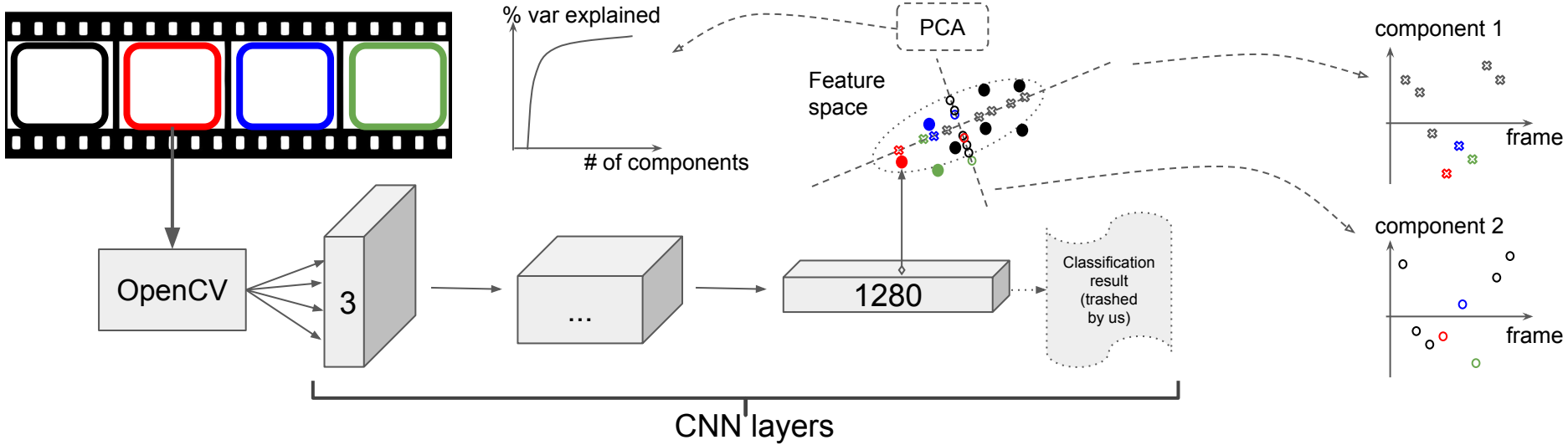
...

so we go deep inside the deep CNN (to fulldepth-1 layer)...

Our workflow

- adjust frame sizes
- unify aspect ratio from different cameras
- etc …

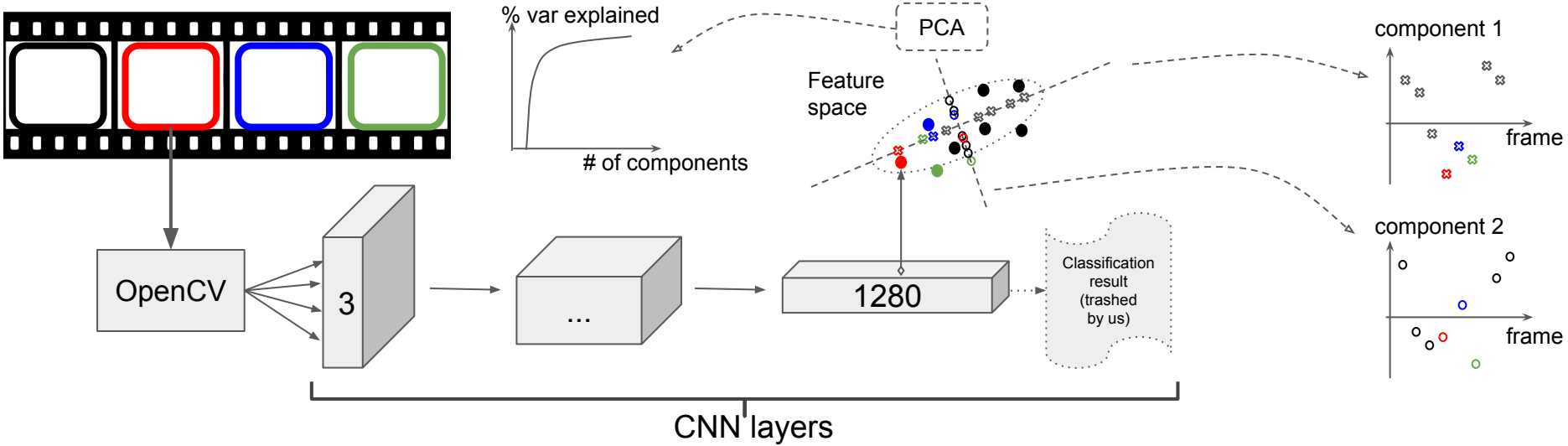
Start from minor preprocessing



The crux: CNN learned to (thank you, the Trainers!)

- compress info
- extract "features"
- classify by looking in feature space

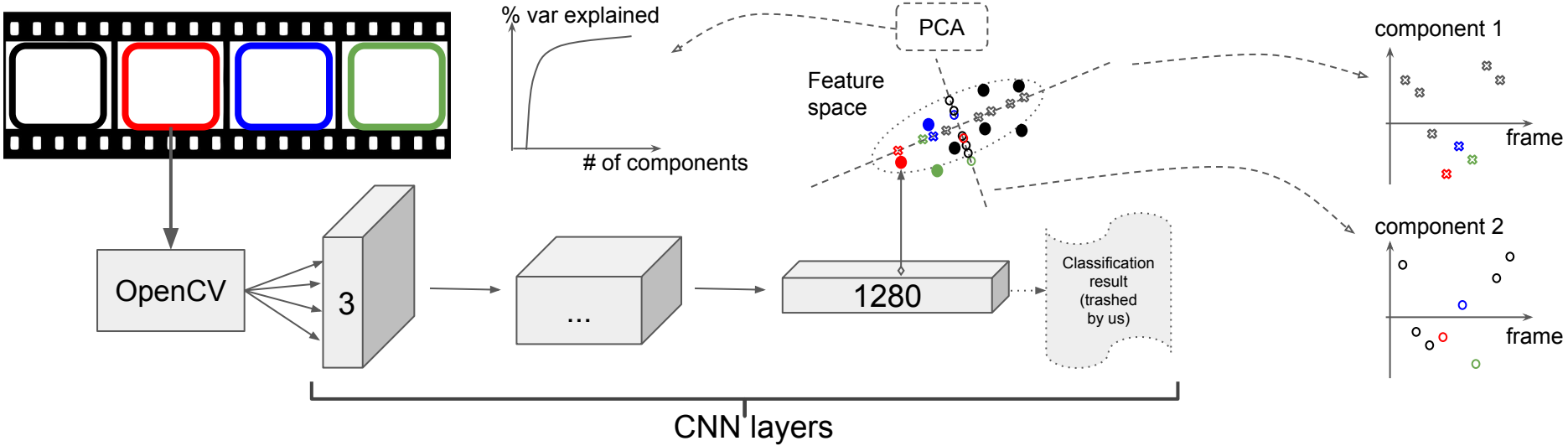
(that's what experts believe)



We "cut CNN open"

- Remove the last layer
  (the classifier itself)
- Record CNN last-1 layer's output
- Get the "features"

We can now work
in the feature space
(and classify, or whatever!)
ourselves!



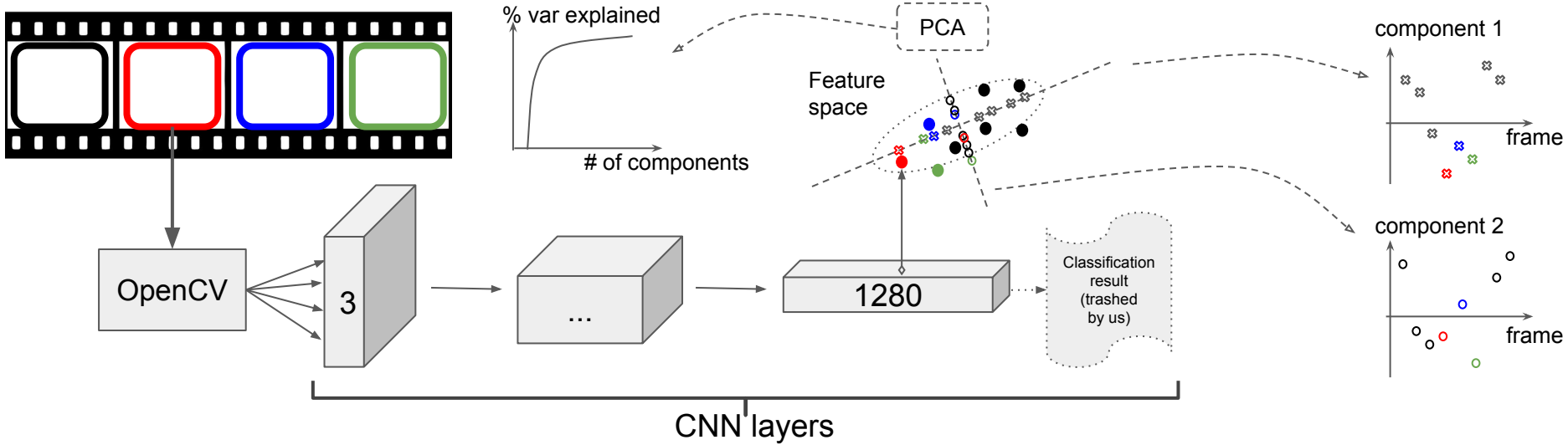
Issue: feature space is too large

For us…

the Trainers used it to classify into many classes (overkill for us)

Need to

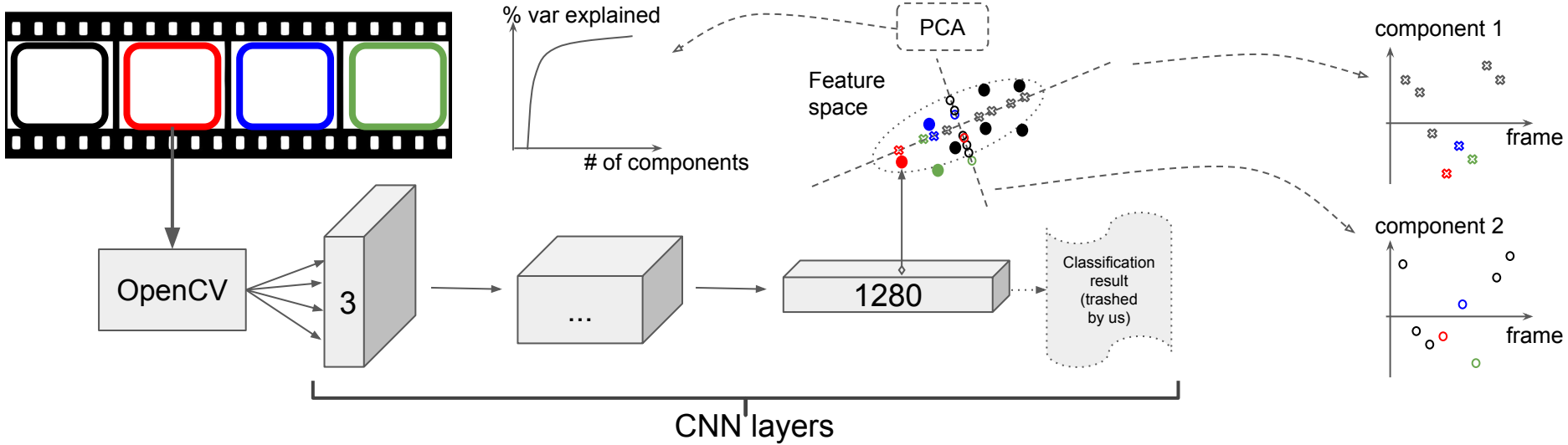
reduce data dimension in the feature space…



PCA: good old friend

- Did we lose steam
  coming up with insights? ;(
- Good old tool worked,
  so why reinvent the wheel ;)

PCA looks at how data is "layed out"
identifies "major directions"
of data variability

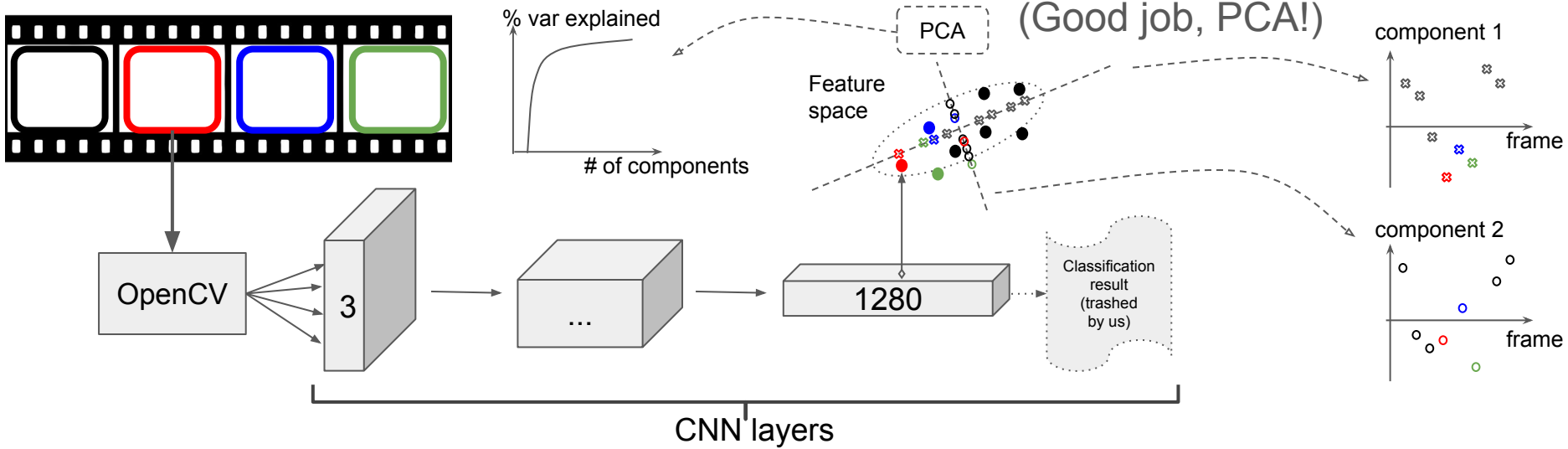


R^2: another classics

- coefficient of determination
- how much of data variability is explained

Computed R^2: high-% of var explained by few principal components

⇨ success measure in statistics: extracted statistically meaningful info (Good job, PCA!)

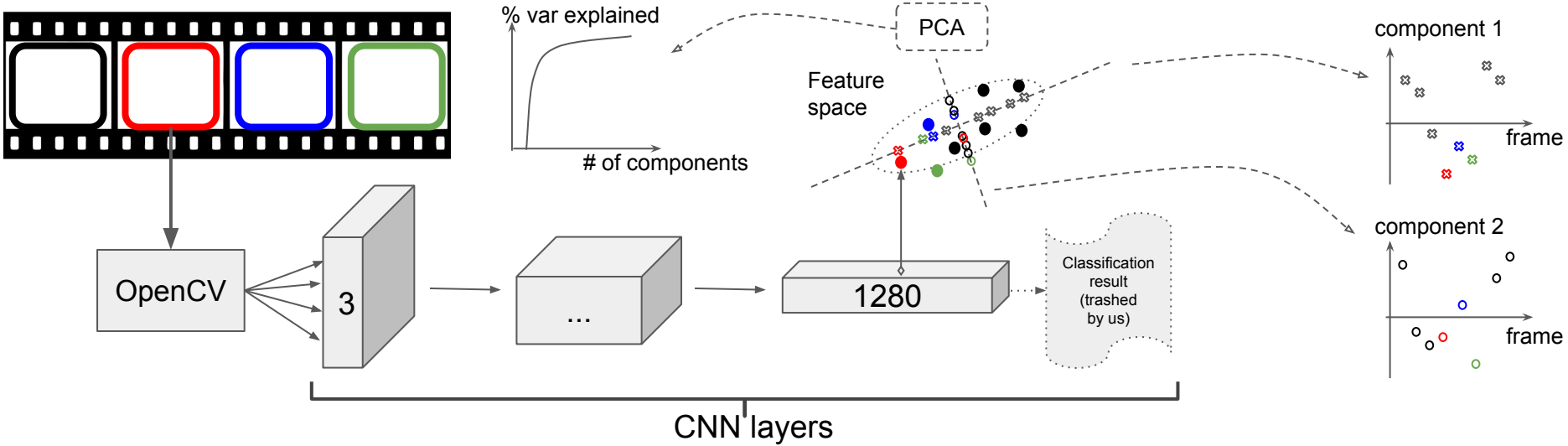


Final output: components/frame

A value

- per each component
- per each frame (time)

For each component:
graph how it changes with time



What are the components?

What do they represent?

No clear idea… principal components

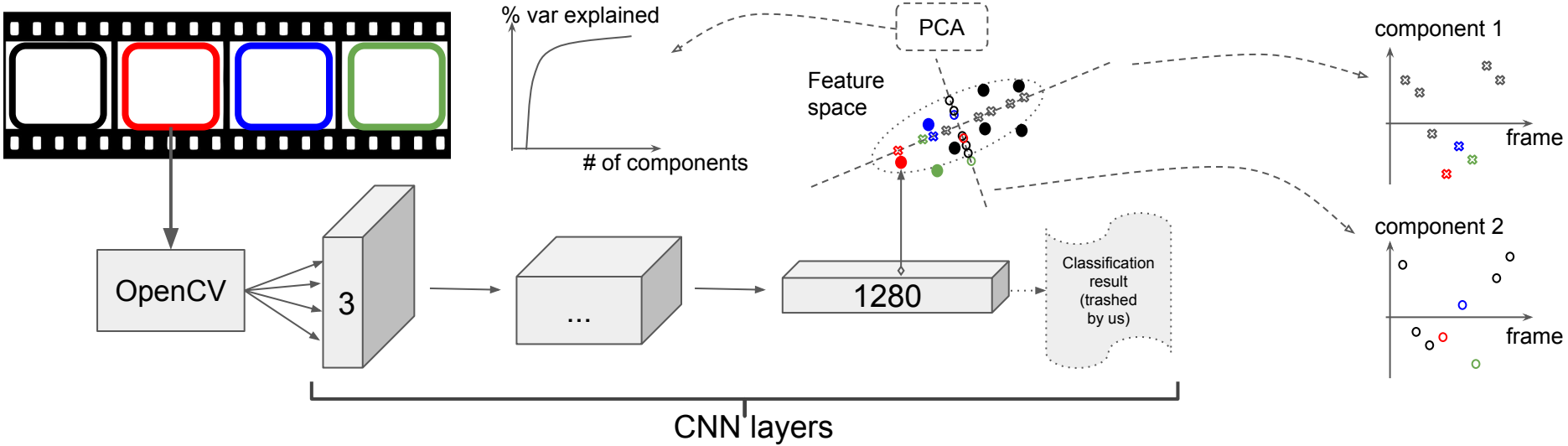
in the feature space

mystery inside mystery

But the magic works:

Let's look at the results

on the real data

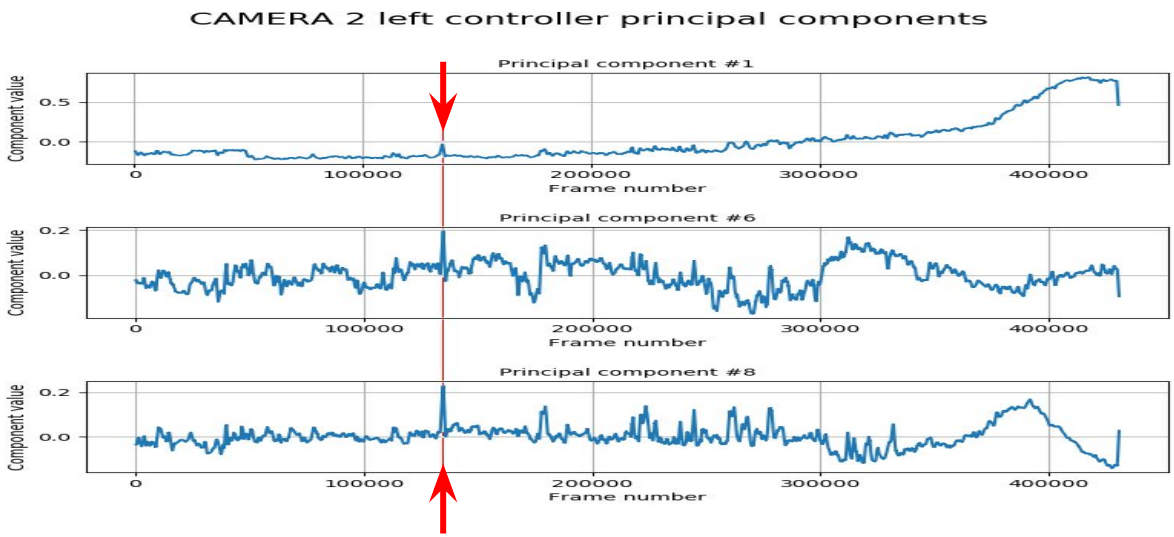


False-positive

Person blocking

Camera 2

- unusual view
- nothing interesting



Snowblowing

NB: peak direction

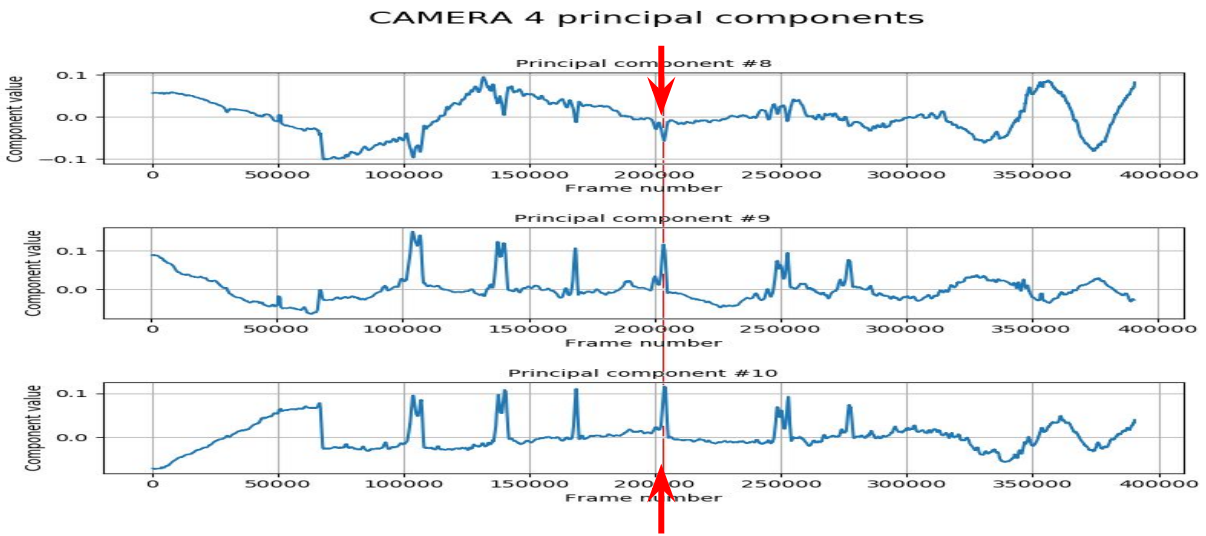
- up
- down

irrelevant:

directions of

components

are rnd



More components and frames in the GUI!

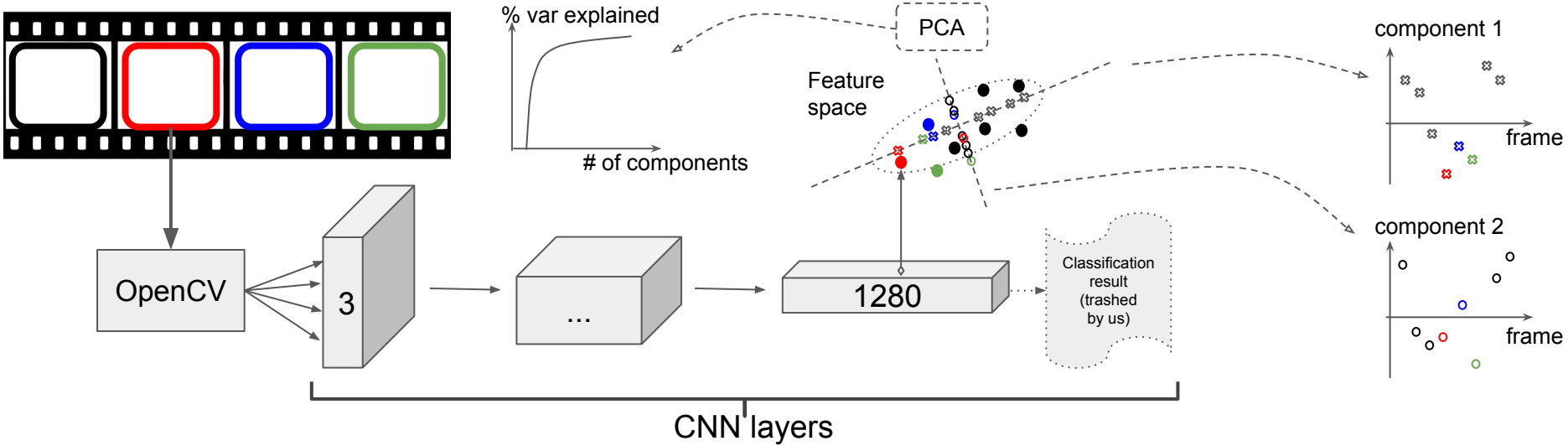
GUI screenshot on this slide

and switch to demo

Summary

Tweaked off-the-shelf state-of-the-art
image classification CNN
to point at unusual scenes
in ATCO work videos

Conclusion:
this thing works! :)



Future work

- Define measures of accuracy
- Compare with other methods
- Embed into workload evaluation