Human Action Recognition summary

V. Dimaridou, D. Makrygiannis, N. Kilis, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr

Version 4.0

Artificial Intelligence &
Information Analysis Lab

Human Action Recognition

- **Human Action Recognition definition and data**
- Classical Human Action Recognition
  - Single view Human Action Recognition
  - Multiview Human Action Recognition
- Neural Human Action Recognition
- GCN Human Action Recognition
- 3D Human Action Recognition
- Human Action Recognition applications

Human Action Recognition

Human Action Recognition (***HAR***):

- To identify the action of a person.
- Action is an elementary ***human activity***.
- ***Input***: a single-view or multi-view video or a sequence of 3D human body models (or point clouds).
- ***Output***: An action label belonging to a set of N_A action classes (e.g., walk, run, jump, …) for each frame or for the entire sequence.

Video Based HAR – Problem Statement

"Video-Based Human Activity Recognition ***(HAR)*** aims to automatically recognize the actions of one or more persons given a series of frame sequences"

Video HAR – RGB Inputs (1)

RGB data

- Are easy to obtain, massive datasets available.
- Can be used as basis for feature extraction.
- Can be used on pretrained networks on massive datasets.

RGB data

- Do not protect user privacy.
- Skeletons extracted from color images are of lower accuracy.

Video HAR – RGB Inputs (2)

Example RGB dataset : UCF 101 [SOO2012].

- *13320 videos from 101 action categories.*

Video HAR – RGB Inputs (2)

Example RGB dataset : HMDB [KUE2011].

- *7000 videos from 51 action categories.*

Video HAR – Depth Inputs (1)

Depth data

- Protect user privacy.
- Highly accurate 3D skeletons can be extracted.

Depth data

- Difficult to obtain, depth cameras are more expensive / difficult for outdoor environments.
- SOTA CNNs mostly use RGB data.

Video HAR – Depth Inputs (2)

RGBD dataset: **UTKinect-Action3D Dataset** [SHI2017].

- *10 subjects performing 10 actions*

Downsides of RGBD datasets:

- Captured in laboratory environment.
- Include limited number of people.

Action recognition


run


walk


jump f.


jump p.


bend

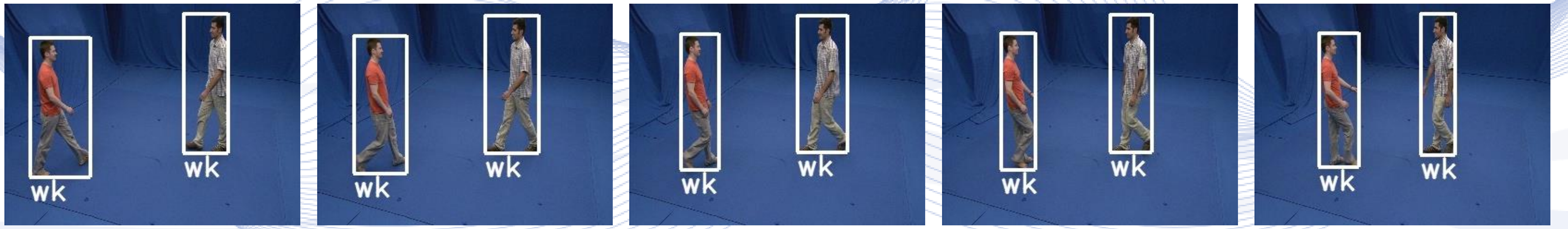

sit


wave

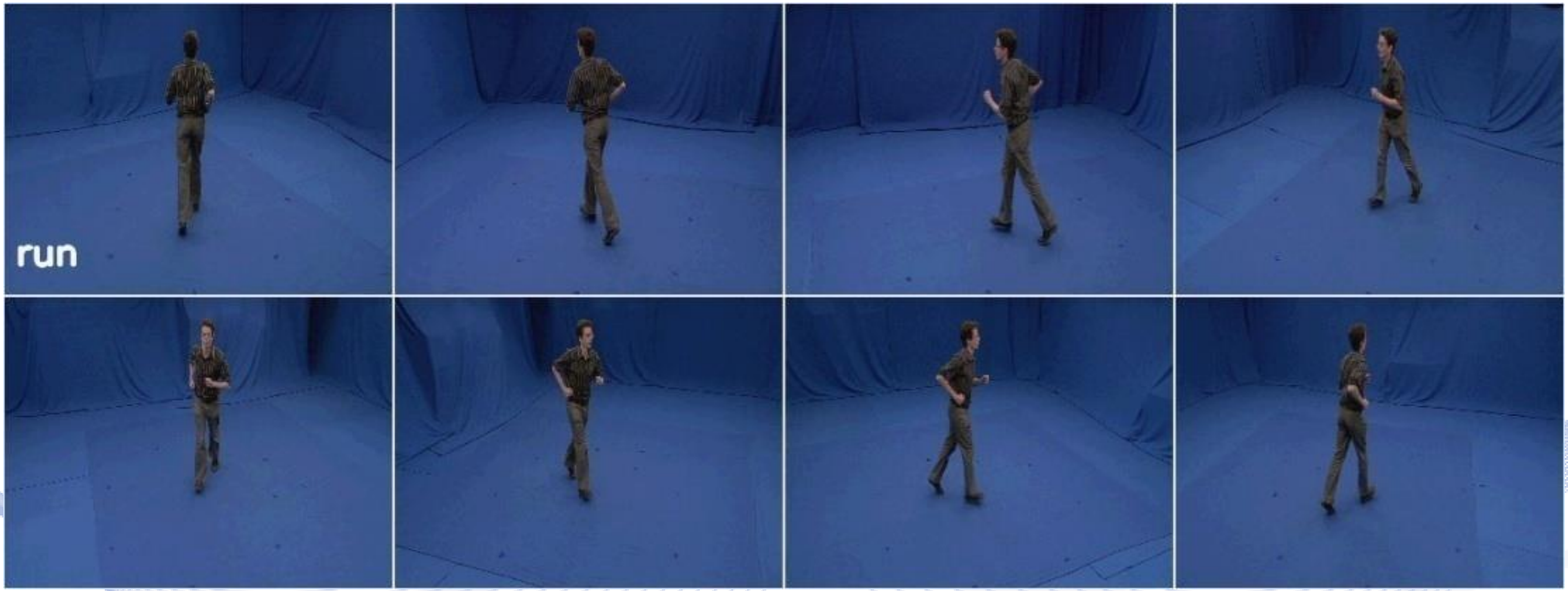

fall

Action recognition

- Applications:
  - Semantic video content description, indexing, retrieval.
  - Video surveillance.
  - ***Human – Computer Interaction*** (***HCI***).

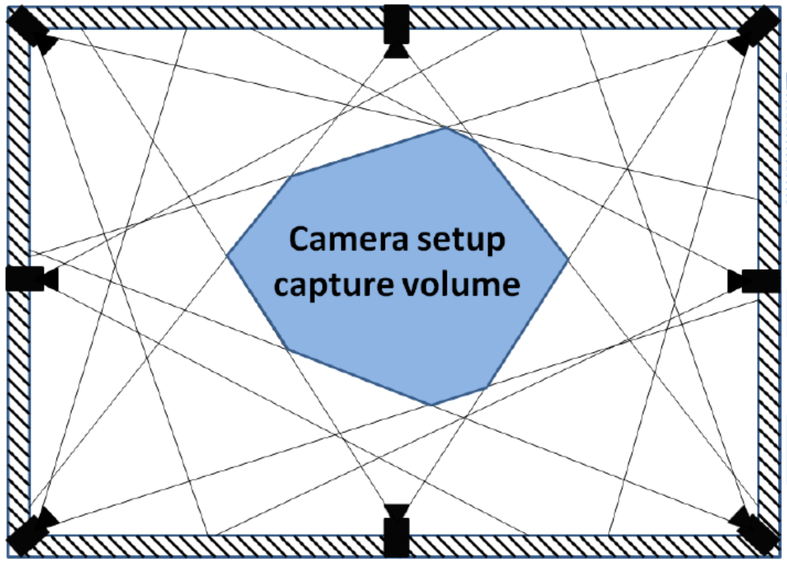


Action recognition



Action recognition methods categorization

- Single-view: methods utilizing one camera:
  - special cases of multi-view ones, i.e., for $N_C = 1$.
- Multi-view: methods utilizing multiple cameras forming a multi-camera setup.



An eight-view camera setup ($N_C = 8$).

Human Action Recognition

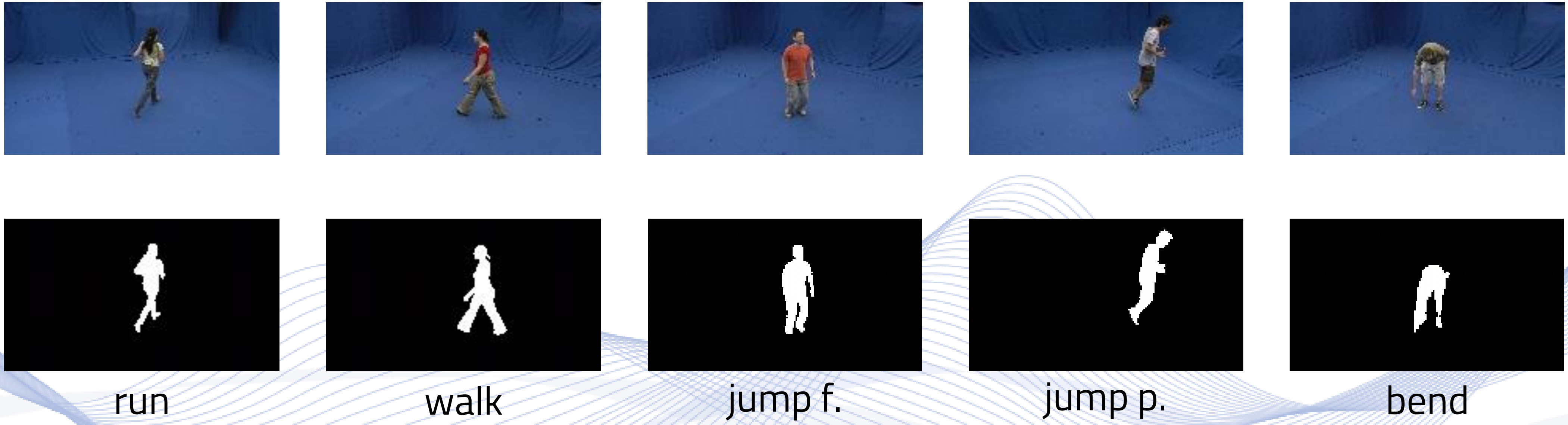
- Human Action Recognition definition and data
- **Classical Human Action Recognition**
  - **Single view Human Action Recognition**
  - Multiview Human Action Recognition
- Neural Human Action Recognition
- GCN Human Action Recognition
- 3D Human Action Recognition
- Human Action Recognition applications

Action recognition on video data

- Input feature vectors: ***human silhouettes***, i.e., binary human body images resulting from coarse body segmentation on each video frame.

- Segmentation techniques:
  - background subtraction,
  - chroma keying,
  - motion detection.

Action recognition on video data



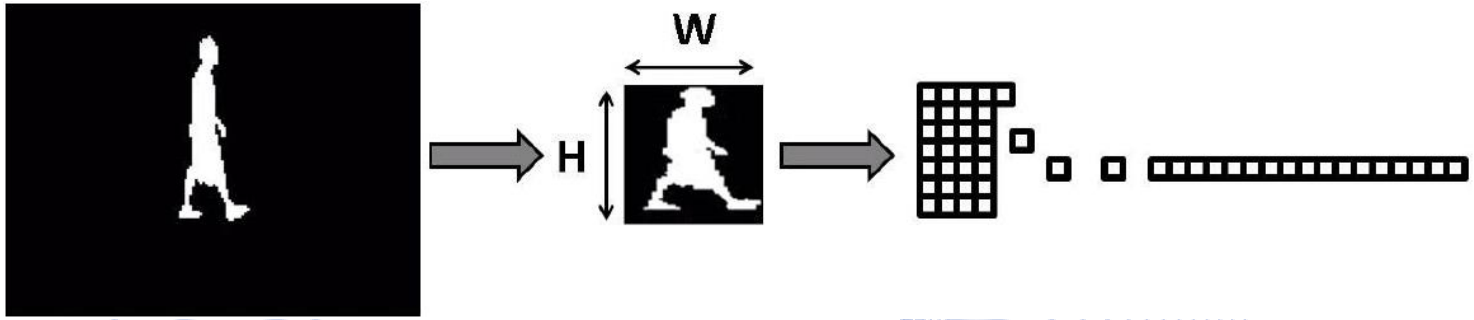
Action description

- A series of successive ***human body poses***:

- Human body poses are represented by binary posture images:

Action representation

- ***Posture vector*** creation:



Binary human body image

Posture image of fixed size

Posture vector

$\mathbf{p}_{ij} \in \mathbb{R}^{HW}$

frame index

action video index

Action representation

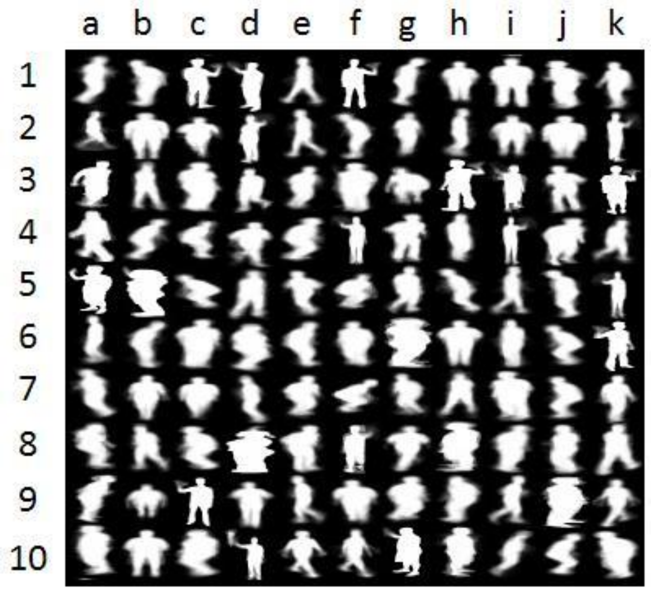
- ***Dyneme*** calculation: Cluster all the training posture vectors $\mathbf{p}_{ij}$ in m clusters without exploiting the available action labels.

- Clustering techniques:
  - K-Means.
  - Self Organizing Map (SOM).
- Dynemes can be considered as representative human body poses.

Dynemes Calculation

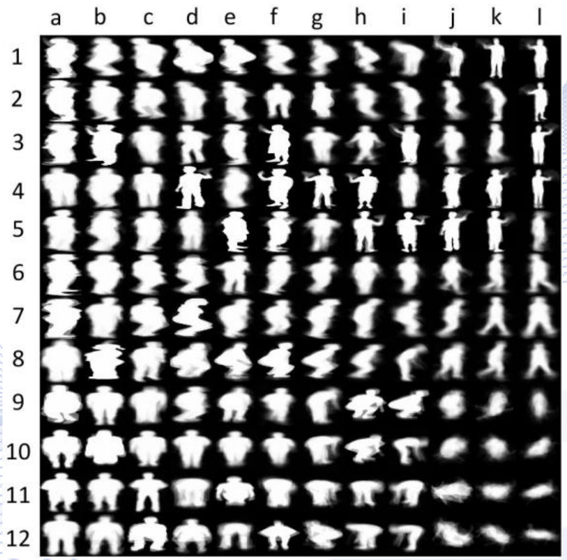
- K-Means: a fast clustering technique minimizing the intra-cluster variance.
- Dynemes are evaluated as the mean vectors of the resulting clusters (cluster centers).

- $m = 110$ dynemes resulted by clustering the posture vectors of the i3DPost eight-view database.



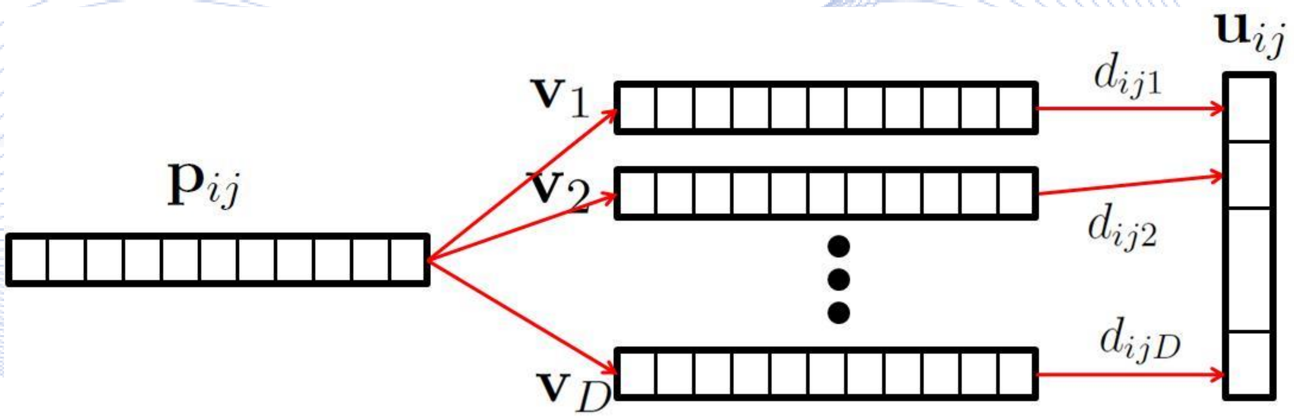
Dynemes Calculation

- SOM: a self organising neural network resulting to a topographic map (lattice) of the input posture vectors.
- A 12×12 lattice ($m = 144$) resulted by clustering the posture vectors of the i3DPost eight-view database.
- Dynemes are determined to be the obtained SOM neurons $\mathbf{v}_k \in \mathbb{R}^{HW}$.



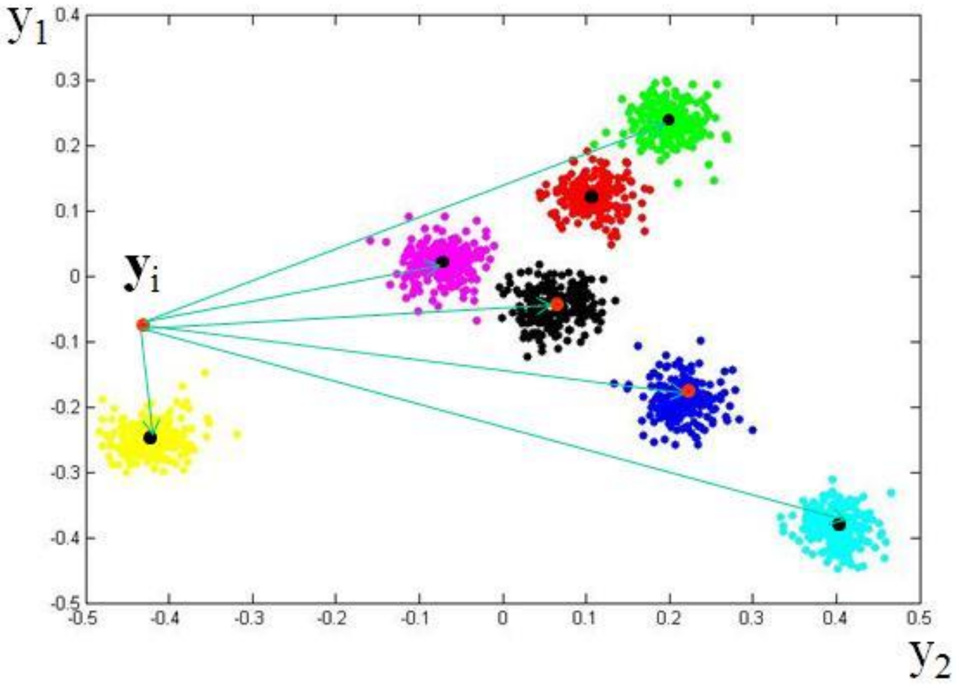
Action representation

- After dynemes calculation each posture vector $\mathbf{p}_{ij}$ is mapped to the so-called membership vector $\mathbf{u}_{ij} \in \mathbb{R}^m$.
- Membership vector $\mathbf{u}_{ij}$ encodes the similarity of posture vector $\mathbf{p}_{ij}$ with all the dynemes.



Action classification

- Classification based on the smallest Euclidean distance from all the mean action class vectors.



Action classification

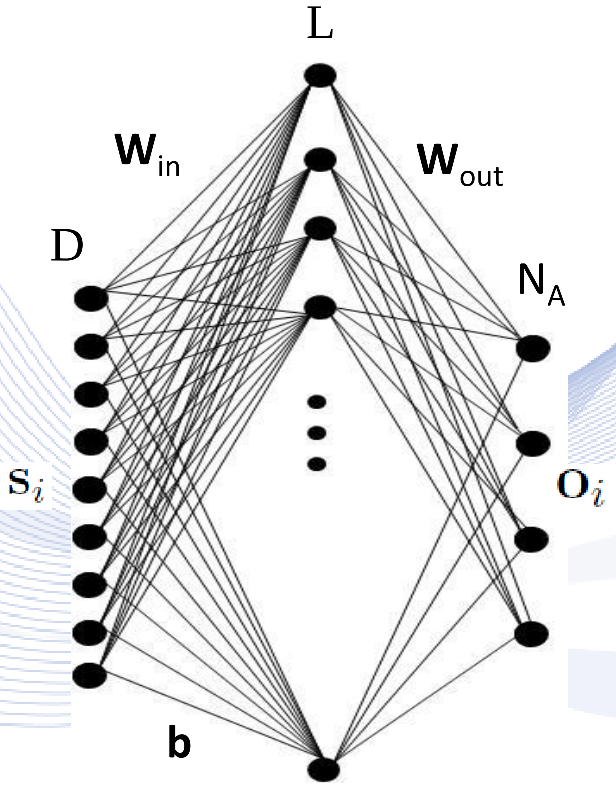
- Calculation of the optimal, in terms of Fisher ratio minimization, projection matrix $\mathbf{W}_{opt}$:

$$\mathbf{W}_{opt} = \underset{\mathbf{W}}{\operatorname{argmin}} \frac{trace\{\mathbf{W}^T \mathbf{S}_w \mathbf{W}\}}{trace\{\mathbf{W}^T \mathbf{S}_b \mathbf{W}\}}.$$

- $\boldsymbol{S}_w, \boldsymbol{S}_b$: ***within–class*** and ***between–class scatter matrices***.
- Each action vector $\mathbf{s}_i$ is mapped to the corresponding discriminant action vector $\mathbf{z}_i \in \mathbb{R}^d$ by $\mathbf{z}_i = \mathbf{W}_{opt}^T \mathbf{s}_i$.

Action classification

- Artificial Neural Networks based action vector classification:
  - Classification based on a SLFN.
  - Network topology: $m \times L \times N_A$ neurons.
  - Randomly chosen input weights $\mathbf{W}_{in} \in \mathbb{R}^{m \times L}$ and bias vector $\mathbf{b} \in \mathbb{R}^{L}$.
  - Analytically calculated output weights $\mathbf{W}_{out} \in \mathbb{R}^{L \times N_A}$.



Action classification

- Artificial Neural Networks based action vector classification.

Binary mask	Description	Walk	Run	Jump in place	Jump forward	Bend	Sit	Fall	Wave one hand
	Walk 90°	0.634	-1.158	-0.442	-1.370	-0.975	-0.812	-0.923	-0.835
	Walk 0°	0.104	-0.319	-0.793	-1.280	-1.007	-0.862	-0.967	-0.902
	Run 0°	-0.727	0.543	-0.190	-1.563	-1.003	-0.915	-1.045	-0.920
	Run 315°	0.613	0.624	-0.799	-1.527	-0.972	-0.815	-1.018	-0.903
	Jump in place 45°	-0.922	-1.231	0.645	-0.222	-0.936	-1.213	-1.015	-1.016
	Jump forward 45°	-1.194	-0.124	-1.039	0.296	-1.044	-0.843	-0.932	-0.920
	Bend 180°	-1.799	-0.469	-1.714	-1.794	1.657	-0.624	-0.624	-1.192
	Sit 225°	-1.010	-0.926	-1.101	-1.307	-1.120	1.225	-1.112	-1.078
	Fall 0°	-1.061	-1.615	-0.684	-0.592	-0.966	-0.964	0.706	-0.986
	Wave one hand 45°	-0.985	-1.199	-1.150	-0.640	-1.014	-1.137	-1.046	1.064

Network outputs for single posture images.

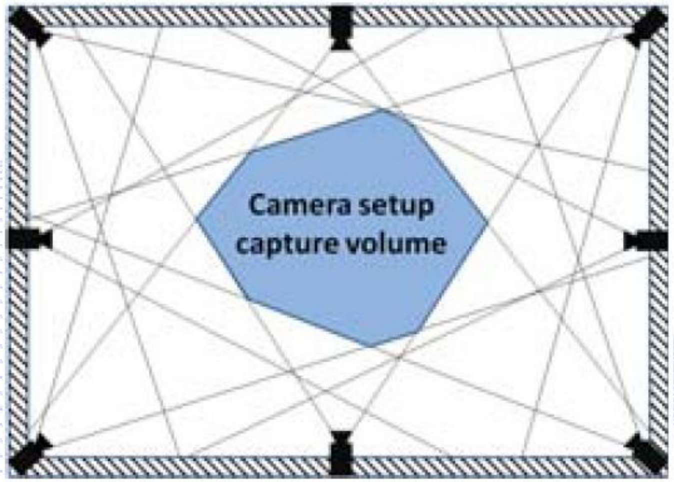
Human Action Recognition

- Human Action Recognition definition and data
- **Classical Human Action Recognition**
  - Single view Human Action Recognition
  - **Multiview Human Action Recognition**
- Neural Human Action Recognition
- GCN Human Action Recognition
- 3D Human Action Recognition
- Human Action Recognition applications

Multi-view action recognition

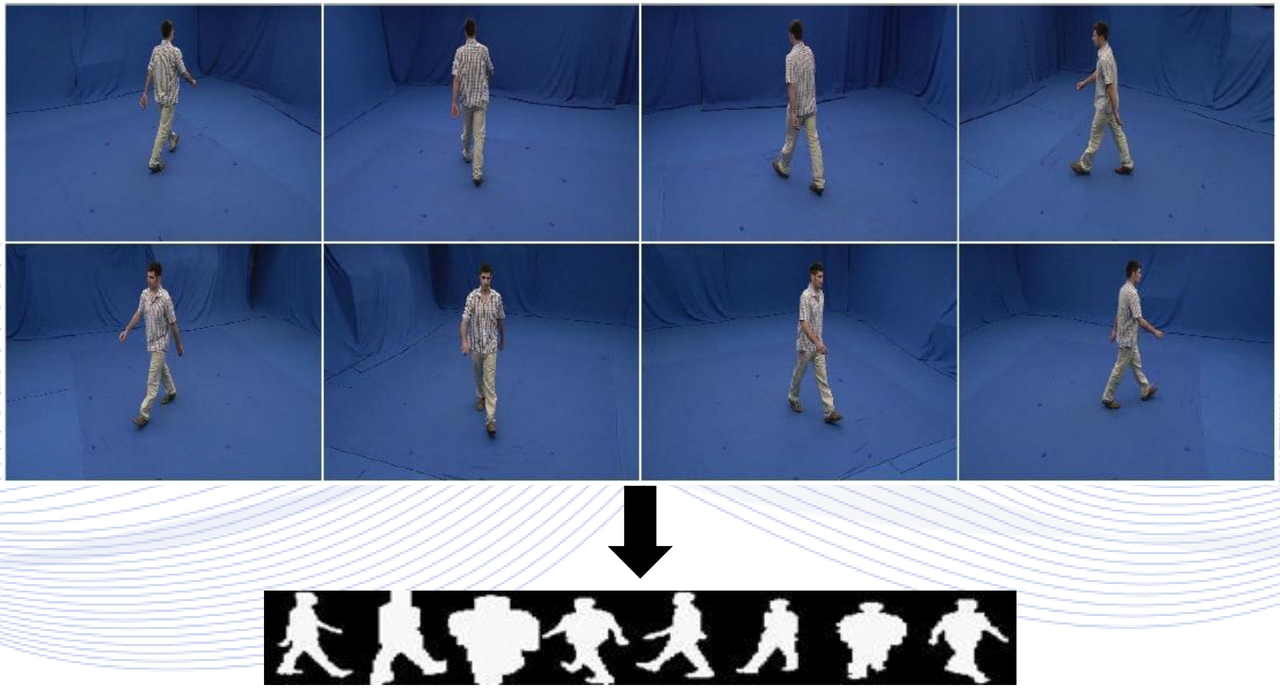
How to fuse information coming from multiple views?

- Creation of multi-view human posture images and proceed with the classification.
- Classification of action videos coming from all the available cameras and combination of the obtained classification results.



Multi-view posture images creation

- Concatenation of posture images according to the known camera labels.



Viewpoint identification problem

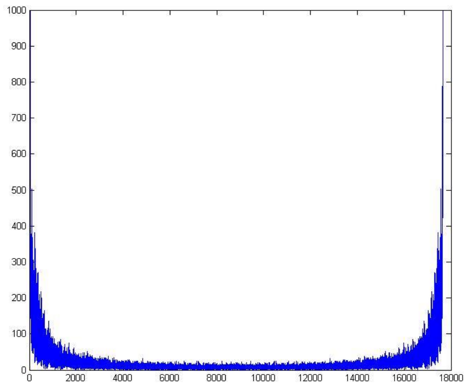
- The person can freely move inside the cameras capture volume. This affects the viewing angle he/she is captured from.
- This will affect the action recognition performance.
- Need for view invariant human body representation.

Multi-view posture images resulting from different movement direction with respect to the camera setup coordinate system.

DFT multi-view posture vectors creation

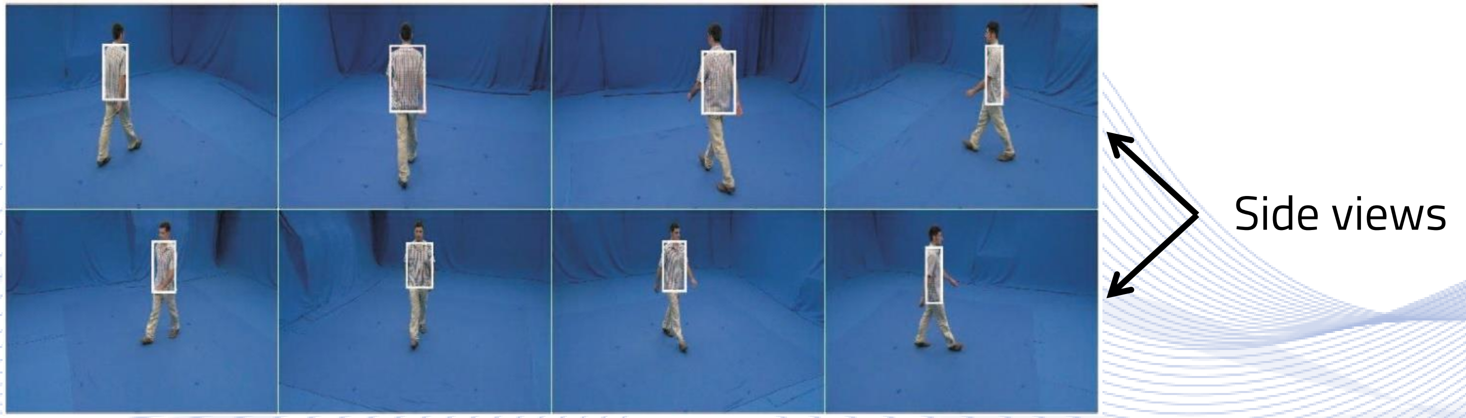
- ***View-invariant*** action recognition by:
- Exploiting the circular shift invariance property of the magnitudes of the Discrete Fourier Transform (DFT).
- Each posture vector $\mathbf{p}_{ij}$ is mapped to a vector $\tilde{\mathbf{p}}_{ij} \in \mathbb{R}^{W \times H}$ containing the magnitudes of the DFT.

Circular shifting of views results to the same DFT magnitude vector.

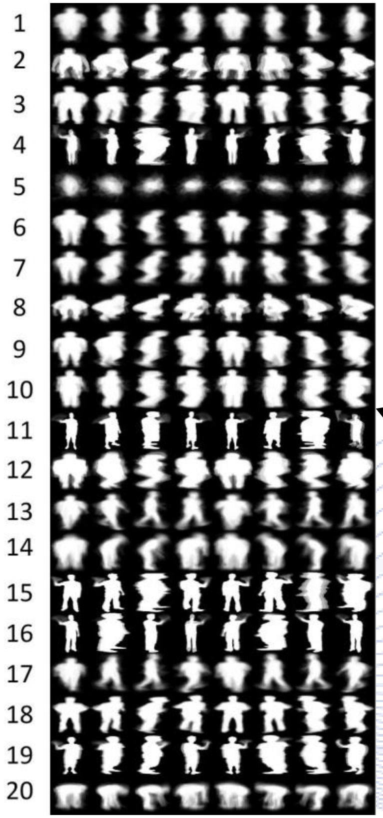


Re-arrangement of multi-view posture images

- View-invariance by:
  - Automatically re-arranging the single-view posture images.

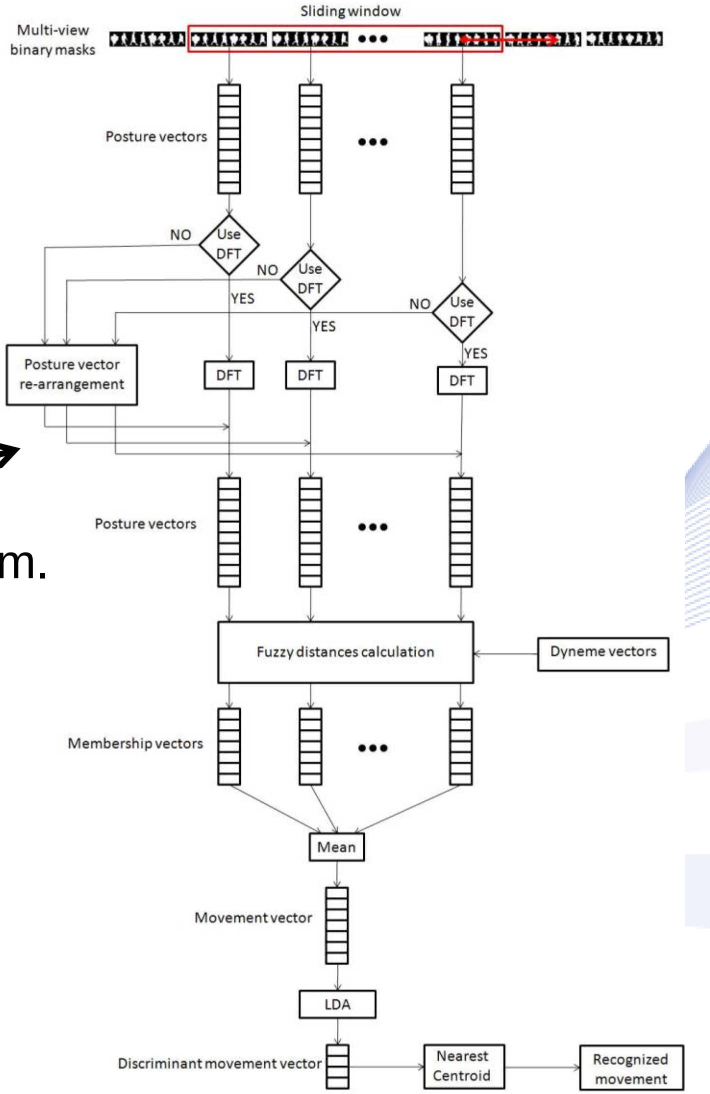


Action recognition by multi-view posture images creation



20 eight-view dynemes.

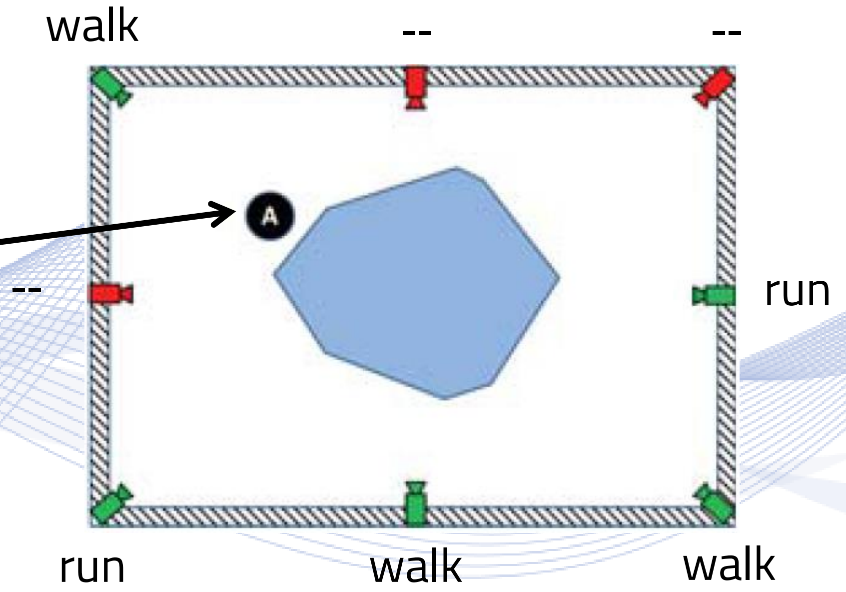
Test phase block diagram.



Action recognition by combining single-view classification results

- Classification of all the available single-view videos independently.

A person performing an action captured by $N \leq N_C$ cameras resulting to the creation of N test action vectors $\mathbf{s}_{test,i}$.



Action recognition by combining single-view classification results

- Combination of the single-view action classification results.

Simple combination (majority voting):

walk:	3
run:	2

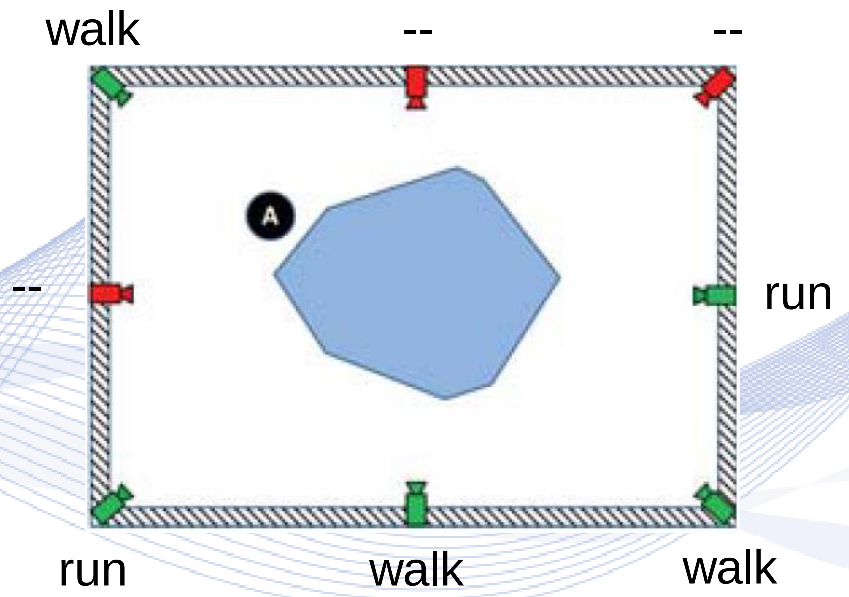
walk

Bayesian combination:

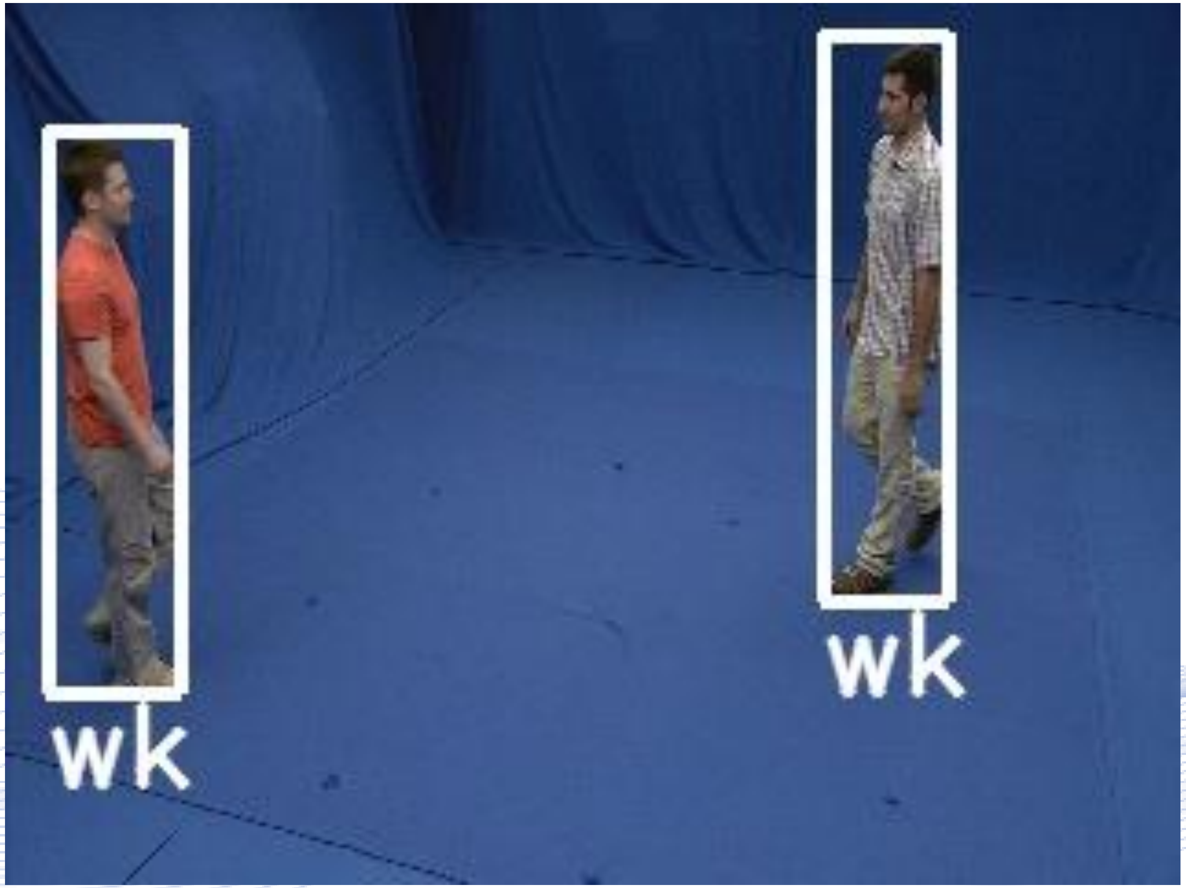
walk:	72%
run:	26%
jump:	2%
bend:	0%
sit:	0%

……

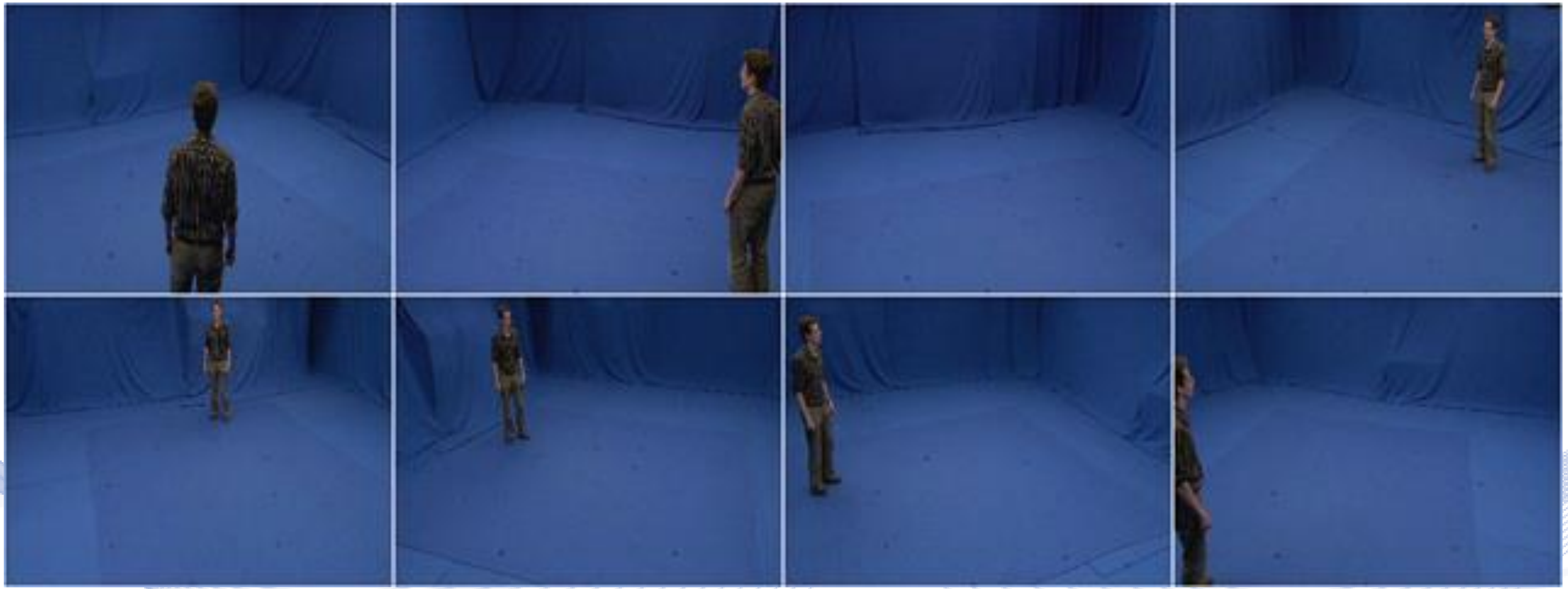
walk



Action recognition examples



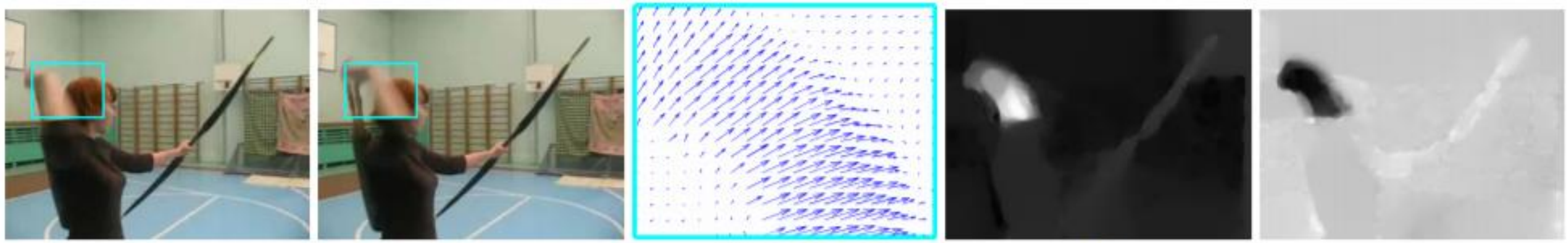
Action recognition examples



Human Action Recognition

- Human Action Recognition definition and data
- Classical Human Action Recognition
  - Single view Human Action Recognition
  - Multiview Human Action Recognition
- **Neural Human Action Recognition**
- GCN Human Action Recognition
- 3D Human Action Recognition
- Human Action Recognition applications

Temporal features using optical flow



[SIM2014]

The stacked optical flow images [SIM2014] are defined as follows:

$$I_\tau(u, v, 2k-1) = d^x_{\tau+k-1}(u, v),$$

$$I_\tau(u, v, 2k) = d^y_{\tau+k-1}(u, v), u = [1; w], v = [1; h], k = [1; L].$$

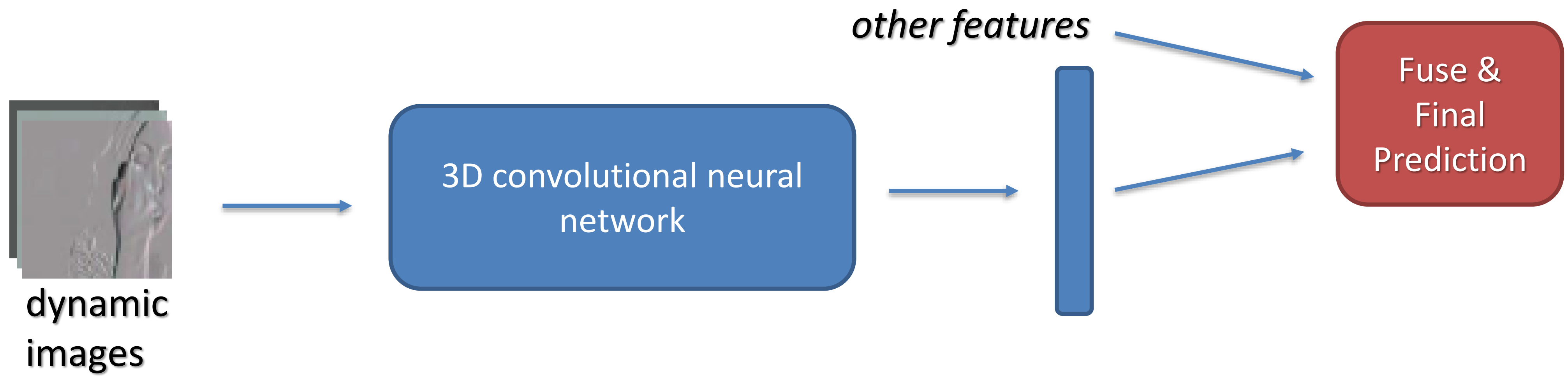
- w, h: video width and height,
- d^x_τ, d^y_τ: horizontal and vertical components of the displacement vector field d_τ,
- L: number frames stacked In one optical flow 3D image.

*The CNN inputs for the temporal channel are the **stacked optical flow images** I_τ, in the x and y directions, for every arbitrary frame τ.*

- Optical flow constructs short-term temporal features
- Is computationally expensive thus not used in real-time information.

Artificial Intelligence & Information Analysis Lab

Temporal features using dynamic image



A ***dynamic image*** [JIN2017] is defined as follows:

$$\rho(I_t, \ldots, I_T; \psi) = \sum_{t=1}^{T} \alpha_t \psi(\mathrm{I}_t),$$

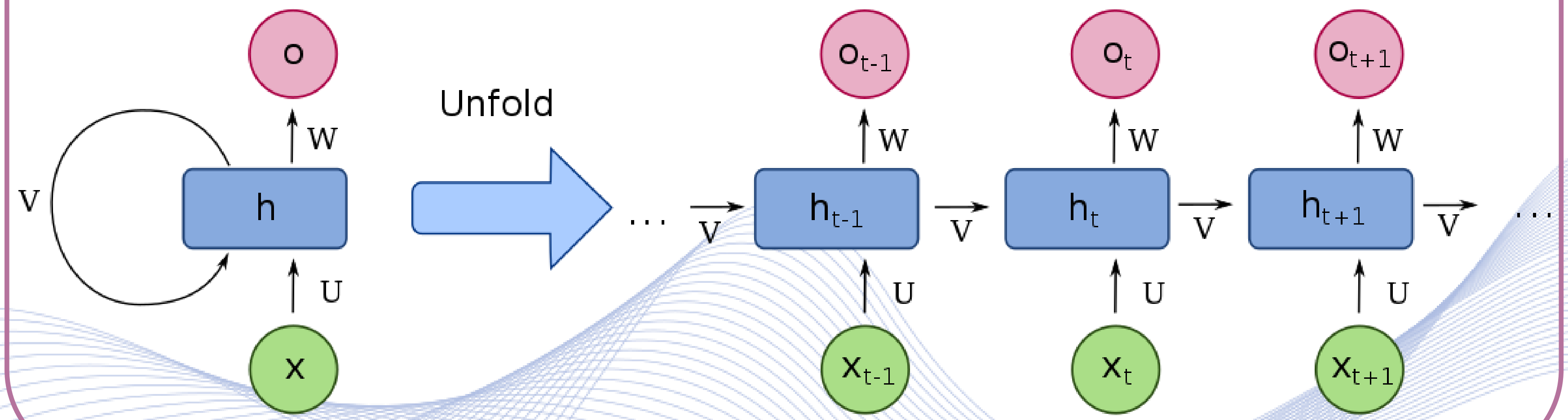
$$\alpha_t = 2(T - t - 1) - (T + 1)(H_T - H_{t-1}).$$

- *$\psi(\mathrm{I}_t)$ is the pixels of the t-th frame,*
- *T is the length of the original video,*
- *$H_t = \sum_{i=1}^{t} \frac{1}{t}$ is the t-th Harmonic number and $H_0 = 0$.*

Dynamic images are fast to compute & manage to capture long-term temporal information of the video.

Video HAR – Which classifiers are used (2)?

- ***LSTM networks*** are widely used in action recognition problems, due to their strength in modeling the dependencies and dynamics in sequential data.
- ***Fused CNN and LSTMs networks***: detect the appearance of an object with the CNN and then analyze its motion using the LSTM[2]



[COLAH]

Skeleton Based Video HAR

"Skeleton-based human action recognition with global context-aware attention LSTM networks" [LIU2017]

Objective:

- Recognize the action performed in video sequences using 3D skeleton data

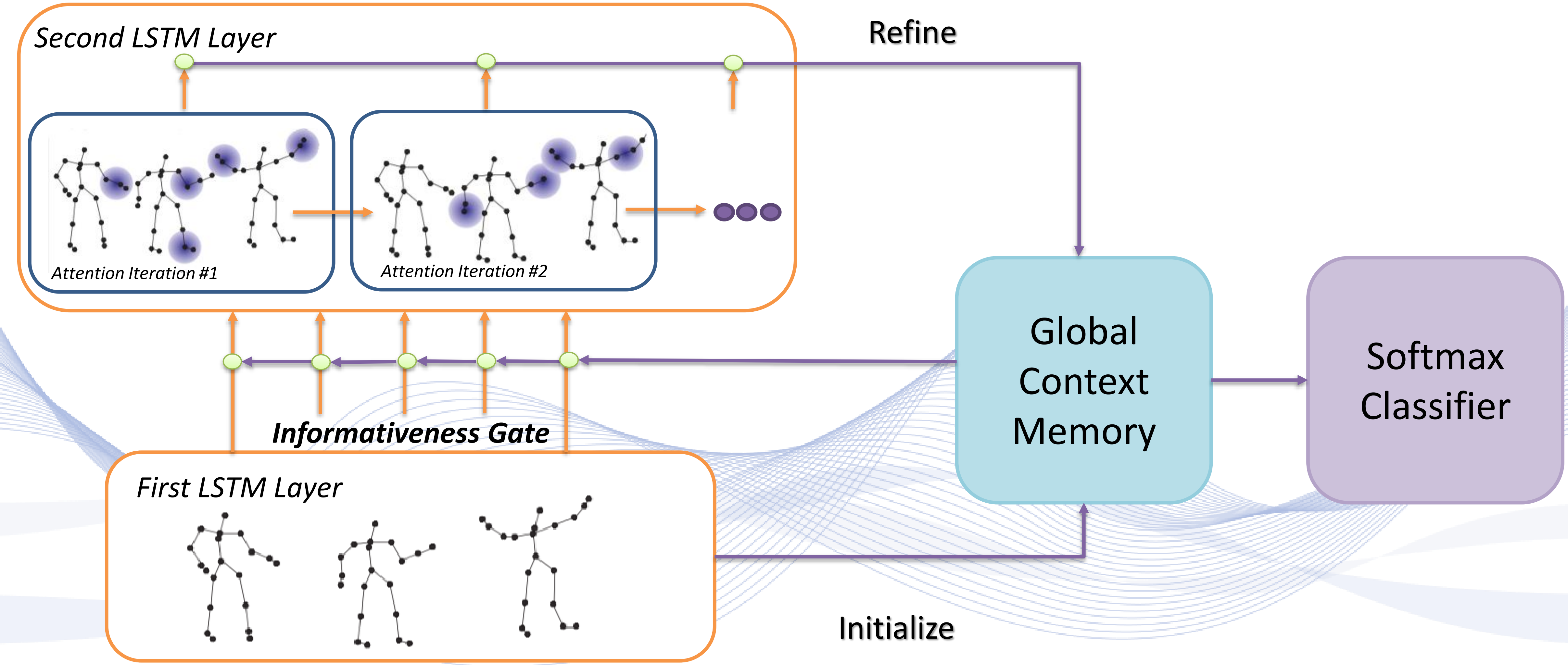
Methodology:

- 2 LSTM models and one ***Global Context-Aware cell*** (***GAC***) are used
- The first LSTM layer initializes the GAC
- The second LSTM performs attention over the inputs by using the global context memory cell to achieve an attention representation for the sequence.
- The attention representation is used back to refine the global context. Multiple attention iterations are performed to refine the global context memory progressively.
- The refined global context information is utilized for classification.

Experiments & Accuracy:

- Tested on : NTU RGBD (84%), SYSU-3D (78.6%), UT-Kinect (99%), SBU-Kinect Interaction and Berkeley MHAD (94.9%)

“Skeleton-based human action recognition with global context-aware attention LSTM networks” [LIU2017]



Multi-stream networks (Quick background)

- Multi-stream networks are implemented using model architectures (e.g. CNNs for image classification tasks) which are trained separately.
- Their softmax scores are combined by late fusion considering different fusion methods, such as averaging or training multi-class classifiers (e.g. SVM) on stacked L_2-normalized softmax scores as features.

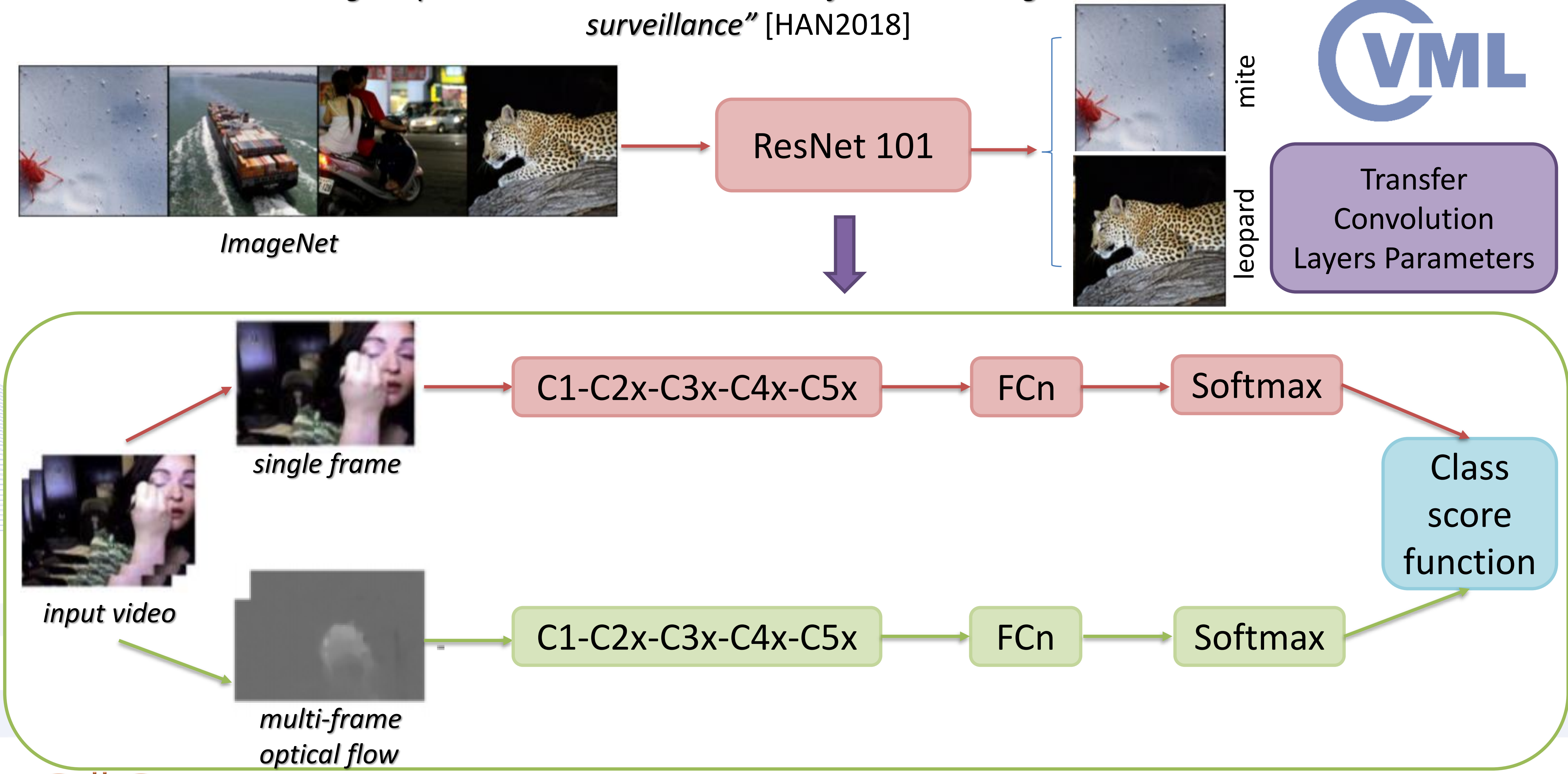
Human visual cortex contains two pathways:

1. the ventral stream (which performs object recognition),
2. the dorsal stream (which recognizes motion).

First stream: ***spatial stream*** performs object recognition on still images.

Second stream: ***temporal stream*** conveys motion information using features like optical flow.

"Going deeper with two-stream ConvNets for action recognition in video surveillance" [HAN2018]



3D CNNs – Quick background

- ***3D CNNs*** can be applied where temporal (e.g., AR) or volumetric context (e.g., Medical Imaging) is important
- Difference between 3D CNN & 2D CNN is that the first applies 3D convolution by using 3D kernels to 3-dimensional data producing 3-dimensional maps

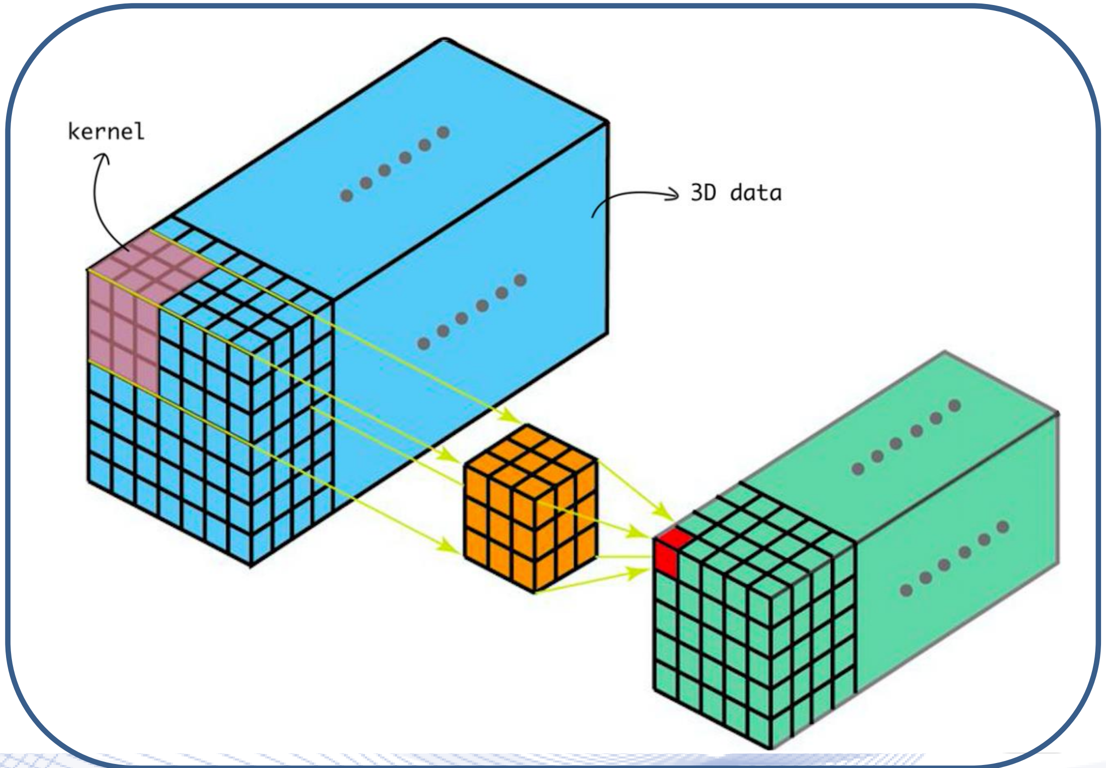
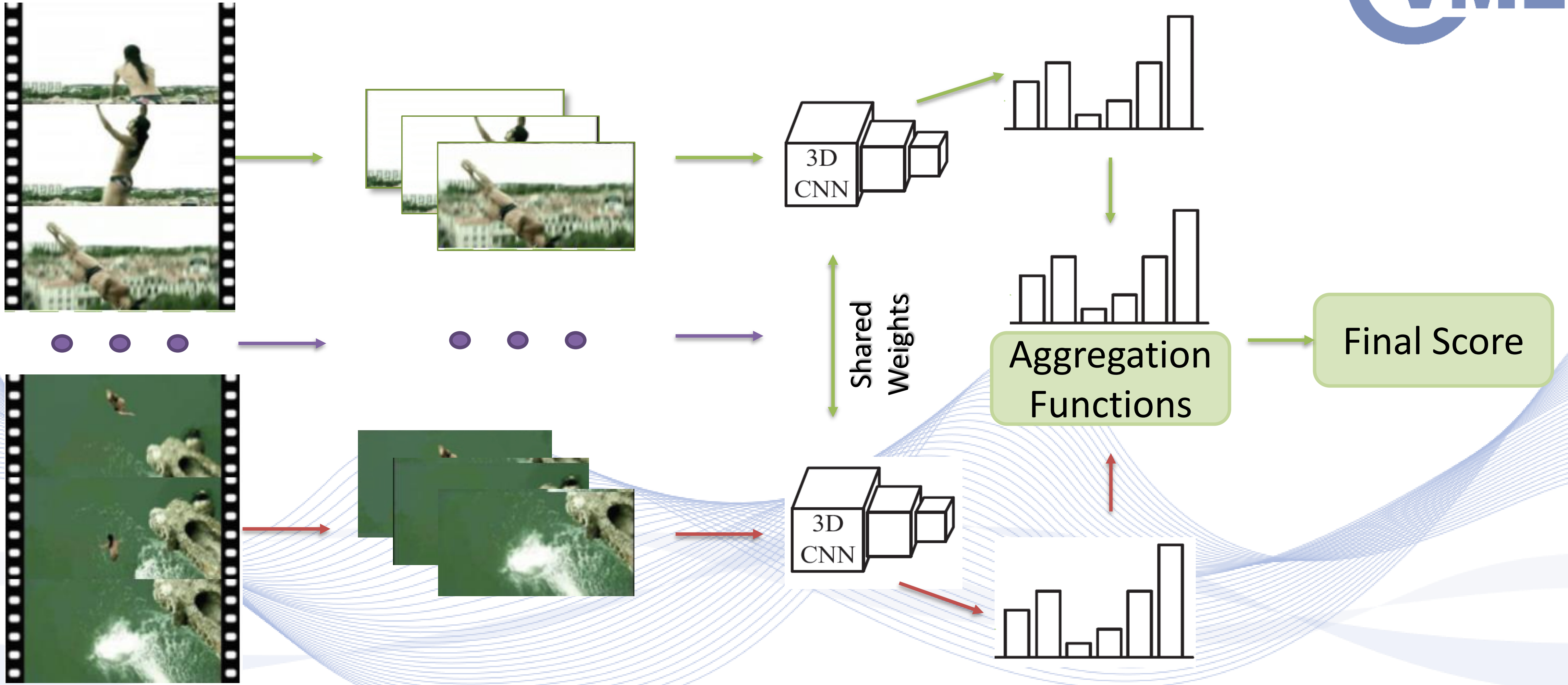


image from https://towardsdatascience.com/understanding-1d-and-3d-convolution-neural-network-keras-9d8f76e29610

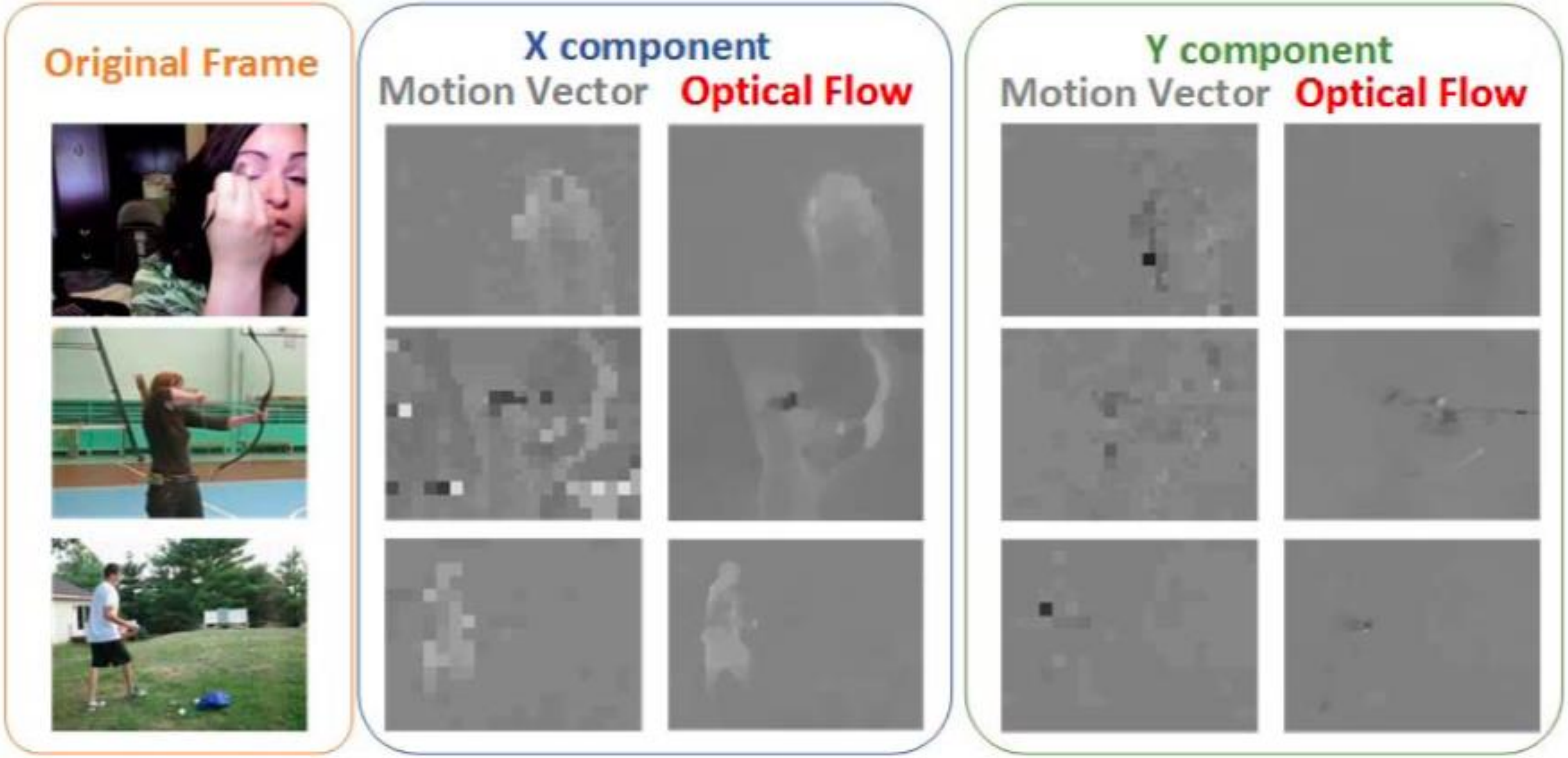
In the HAR case the 3D based CNN can learn spatio-temporal features from raw frame sequences without using complex hand-crafted features or multi-stream architectures.

Artificial Intelligence & Information Analysis Lab

"T-C3D: temporal convolutional 3d network for real-time action recognition" [LIU2018].



Real-Time Action Recognition With Deeply Transferred Motion Vector CNNs



[ZHA2018] Comparison of motion vector and optical flow in X and Y components.

Real-Time Action Recognition With Deeply Transferred Motion Vector CNNs

Structure for real-time action recognition system. In spatial and temporal CNN, F stands for kernel size and S means stride step. O represents for output number and P is pad size.[ZHA2018]

Real-Time Action Recognition With Deeply Transferred Motion Vector CNNs

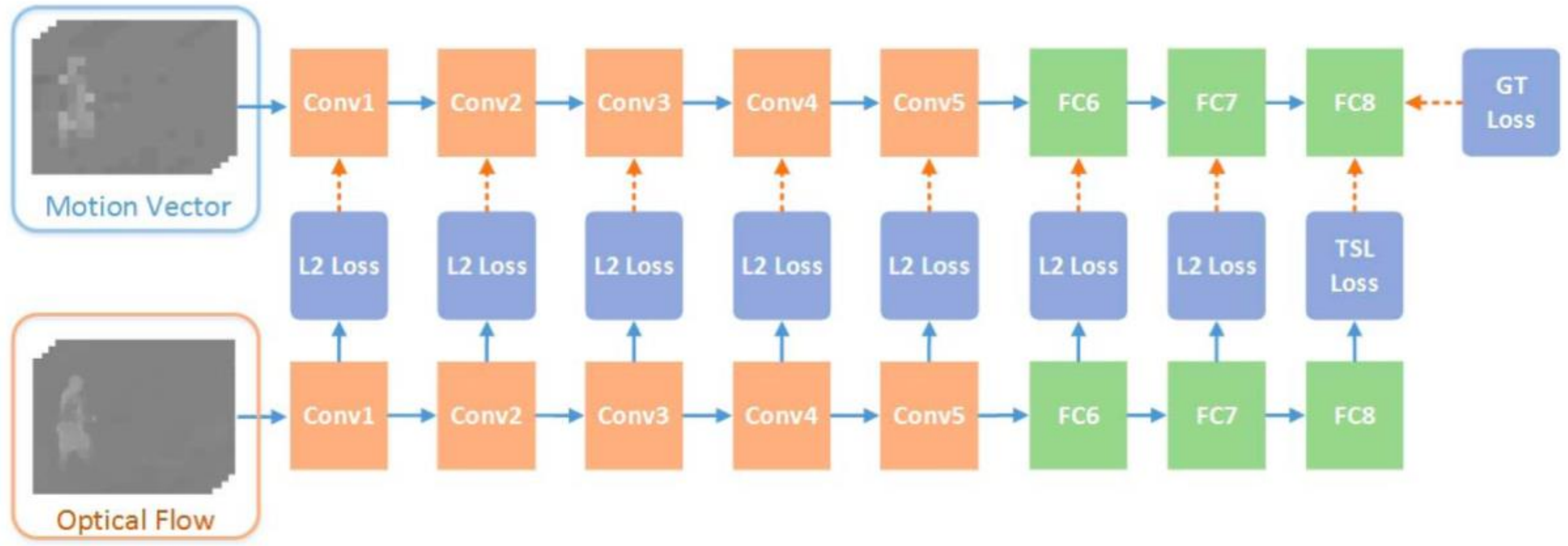
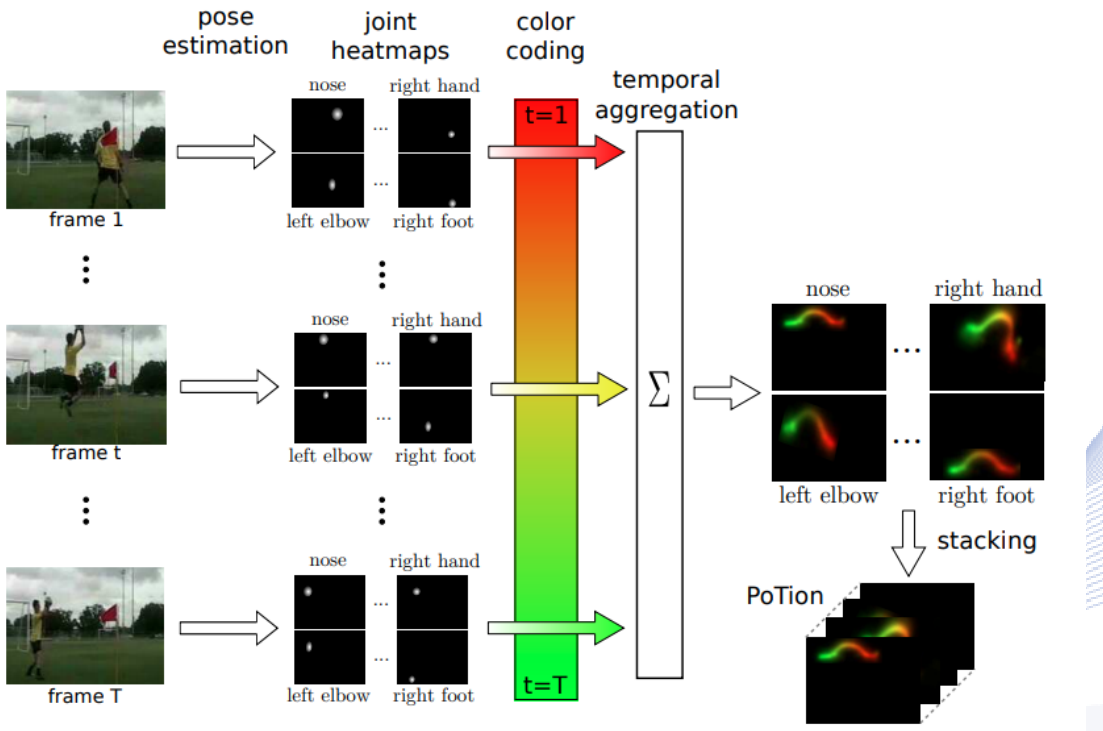


Fig. 4. Structure for Deeply Connected Transfer. Blue lines represent the feed forward process of CNN, while the orange dash line means the back propagation for DTMV-CNN. It should be noticed that the OF-CNN is only utilized during training and the weight for OF-CNN is frozen.

[ZHA2018]

PoTion: Pose MoTion Representation for Action Recognition

- The experimental evaluation shows that PoTion outperforms other SOA pose representations.
- When combining PoTion with the recent two-stream I3D approach , SOA performance is obtained on the JHMDB, HMDB and UCF101 datasets.



[CHO2018] Illustration of PoTion representation. Given a video, joint heatmaps are extracted for each frame colorized using a color that depends on the relative time in the video clip.

PoTion: Pose MoTion Representation for Action Recognition

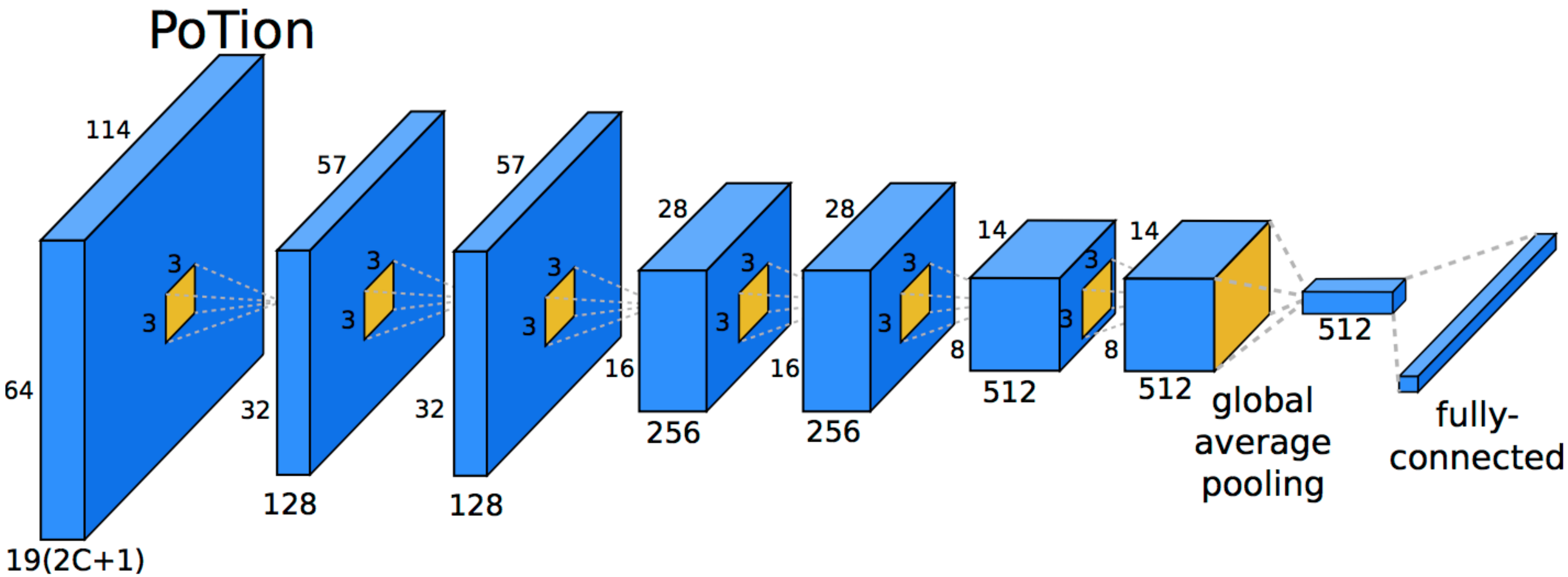


Figure 4. Architecture of the classification network that takes as input the PoTion representation of a video clip.

[CHO2018]

Human Action Recognition

- Human Action Recognition definition and data
- Classical Human Action Recognition
  - Single view Human Action Recognition
  - Multiview Human Action Recognition
- Neural Human Action Recognition
- **GCN Human Action Recognition**
- 3D Human Action Recognition
- Human Action Recognition applications

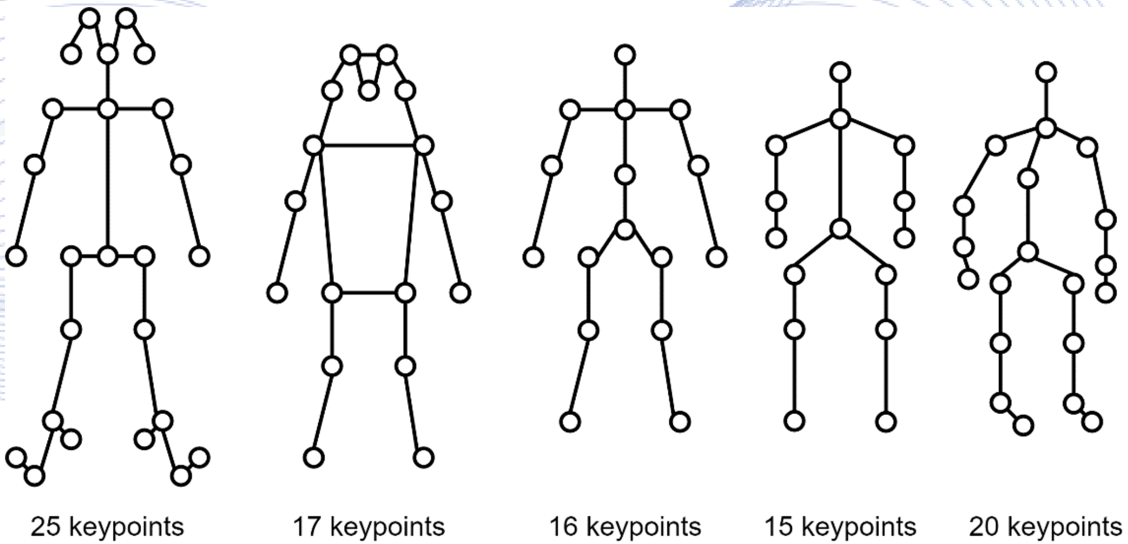
Spatio-Temporal GCN

- Proposed by [YAN2018].
- Applied in skeleton-based ***Human Action Recognition*** from video frames:
  - Important topic in Computer Vision,
  - Identification of ***actions*** that take place in a video:
    - Primitive action, elementary body part motion (e.g., Hand raising).
    - Action, incorporates multiple temporally organized primitive actions (e.g., Running).
    - Activity, high-level motion that includes several actions (e.g., Playing tennis).
  - Other applications: Robotics, Medicine, Supervised physical training, Human-computer interaction.

Spatio-Temporal GCN

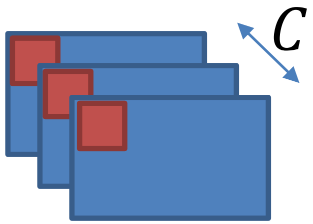
- Human skeleton:
  - Keypoints: Nodes in the Graph,
  - Connections: Edges in the Graph.

- Representation with graphs:
  - ***Invariant to view point and appearance***.

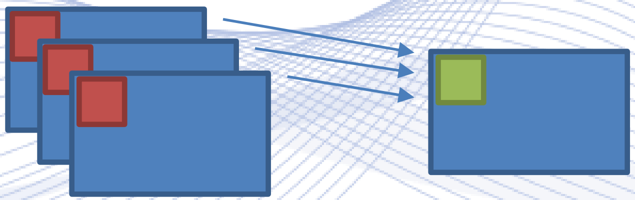


Spatio-Temporal GCN

- ***Spatial Convolution block***:
  - Uses $[1 \times 1]$ kernel, that ensures that features from a frame do not overlap with other frames.



  - Sums all the values from the C channels and returns a single value for each node.

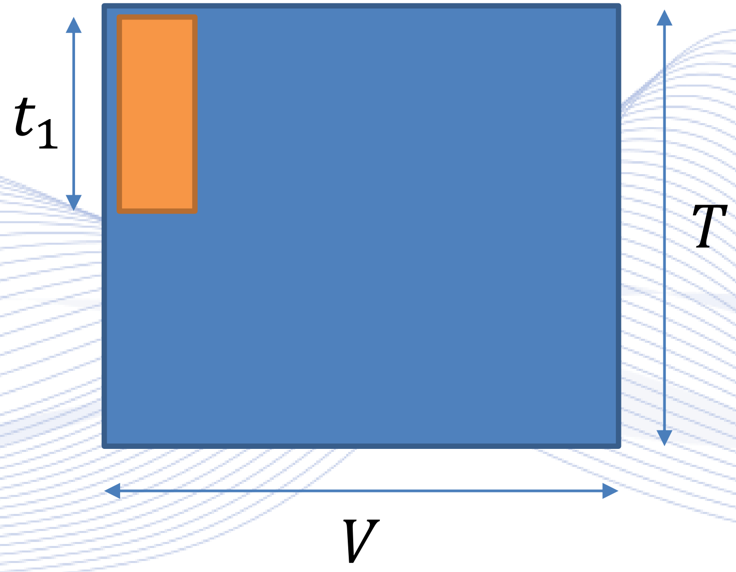


- The spatial convolution output is then ***multiplied with the Adjacency matrix***.

Spatio-Temporal GCN

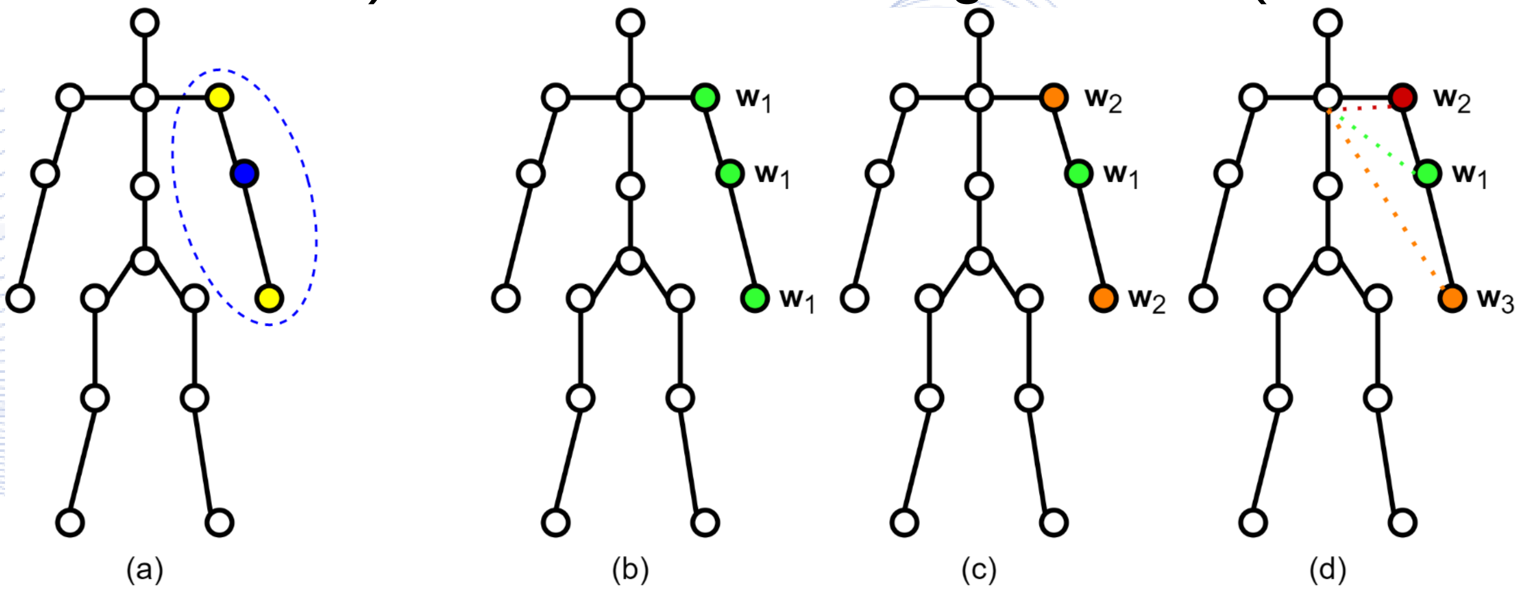
- The multiplication output is fed into a ***Temporal Convolution block***.

- The Temporal Convolution uses a $[t_1 \times 1]$ kernel:



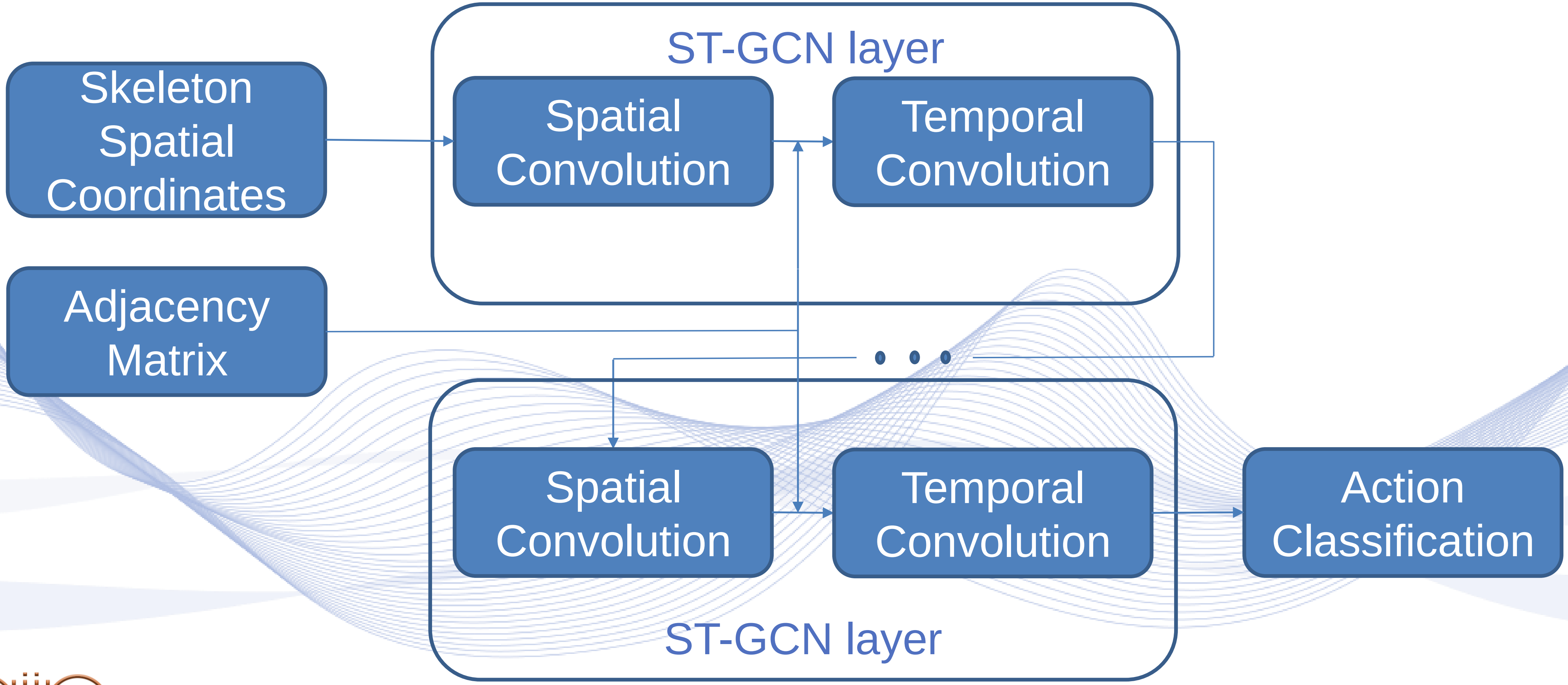
Spatio-Temporal GCN

- Deal with absence of node ordering, introduced by [NIE2016]:
  - Partition Strategies to create subsets:
    - ***Uni-labeling***, all nodes in a neighborhood are treated the same.
    - ***Distance based***, 1st subset: root node, 2nd subset: 1-hop neighborhood.
    - ***Spatial location based*** , 1st subset: root node, 2nd subset: centripetal nodes (closer to center than root), 3rd subset: centrifugal nodes (further away).



Artificial Intelligence & Information Analysis Lab

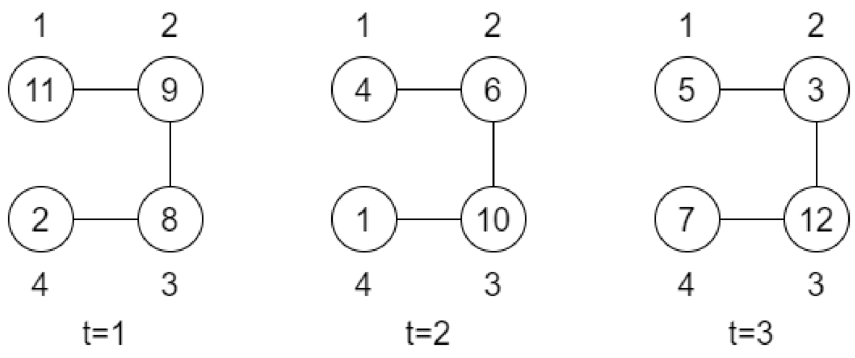
Spatio-Temporal GCN

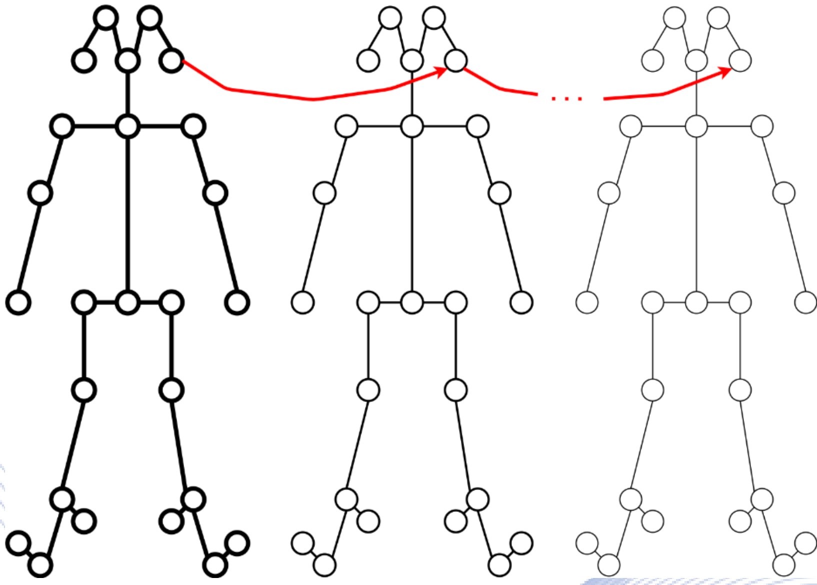


Artificial Intelligence & Information Analysis Lab

Spatio-Temporal GCN

- A sub-graph example of 4 joints and 3 frames:





11	9	8	2
4	6	10	1
5	3	12	7

⇨

0	0	0	0
11	9	8	2
4	6	10	1
5	3	12	7
0	0	0	0

padding

$\otimes$

1
2
3

kernel

=

34	36	8	46
34	30	64	25
14	12	34	15

*

0.5	0.33	0	0
0.5	0.33	0.33	0
0	0.33	0.33	0.5
0	0	0.33	0.5

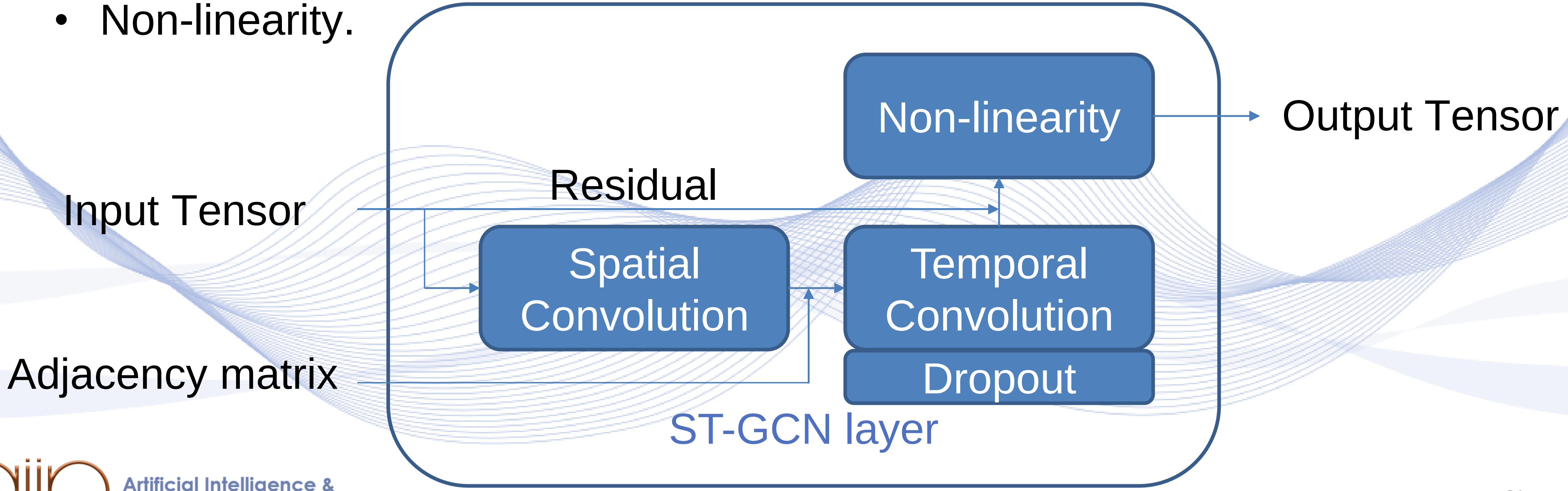
=

35	26	30	27
32	42	40	44
15	20	20	24

$[t_1 \times 1]$

Spatio-Temporal GCN

- The ***ST-GCN layer*** is also equipped with:
  - A Residual mechanism,
  - Dropout,
  - Non-linearity.

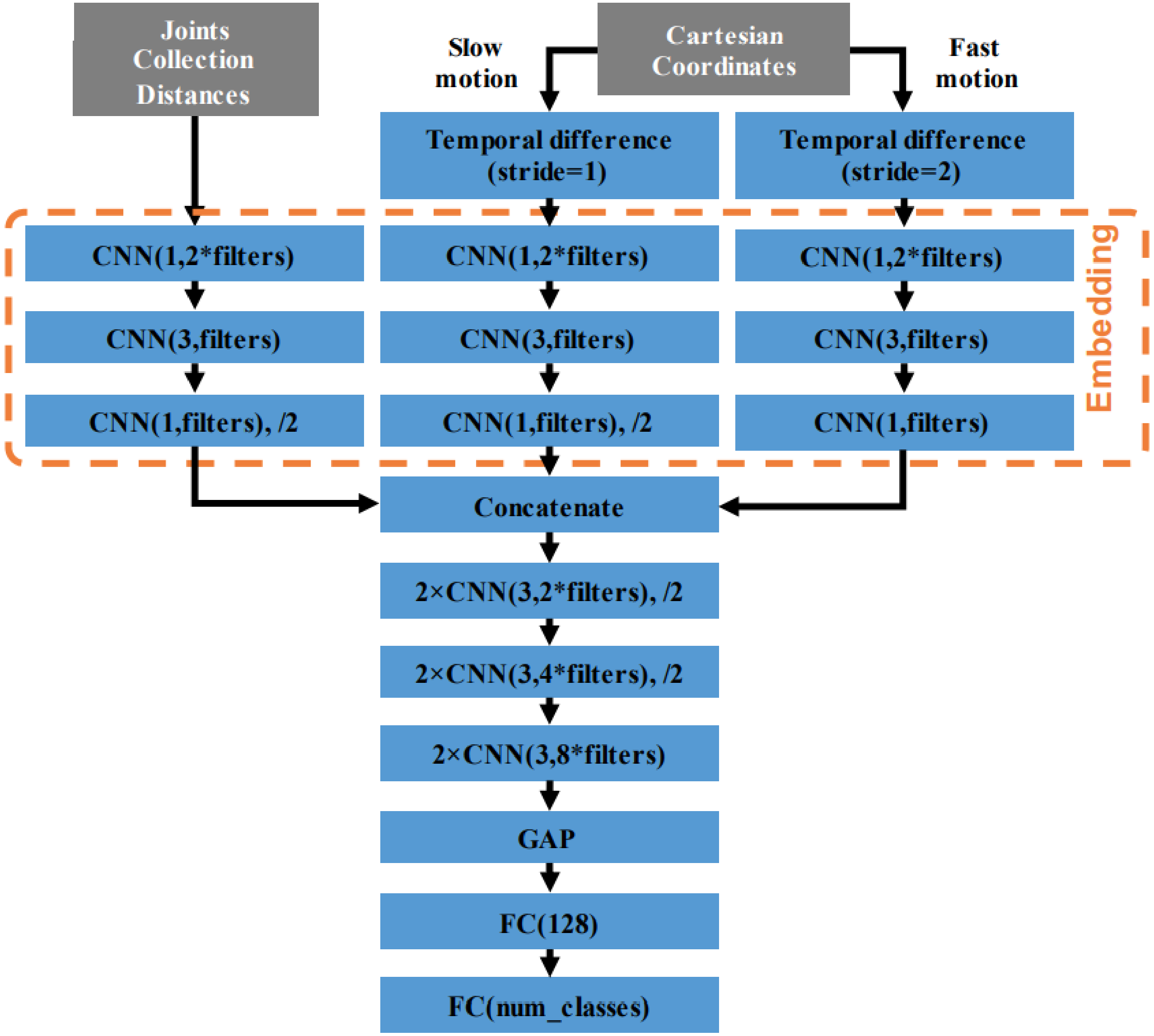


DD-Net Skeleton-based Action Recognition

[YAN2017]Most of the existing methods for skeleton-based action recognition may suffer from a large model
size and slow execution speed.

To address this: analyzed skeleton sequence properties to propose a Double-feature Double-motion Network (DD-Net) for skeleton-based action recognition.

[ZHA2018]



[YAN2017]The network architecture of DD-Net. GAP denotes Global Average Pooling. FC denotes Fully Connected Layers (or Dense Layers).

DD-Net Skeleton-based Action Recognition

Due to the simplicity of DD-Net, many possibilities exist to enhance/extend it for broader studies.

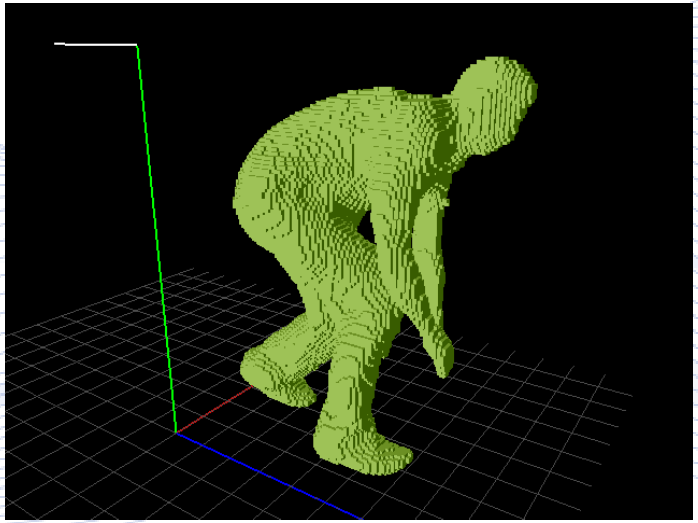
For instance, online action recognition can be approached by modifying the frame sampling strategies; RGB data or depth data could be used with it to further improve the action recognition performance; it is also possible to extend it for temporal action detection by adding temporal segmentation related modules.

Human Action Recognition

- Human Action Recognition definition and data
- Classical Human Action Recognition
  - Single view Human Action Recognition
  - Multiview Human Action Recognition
- Neural Human Action Recognition
- GCN Human Action Recognition
- **3D Human Action Recognition**
- Human Action Recognition applications

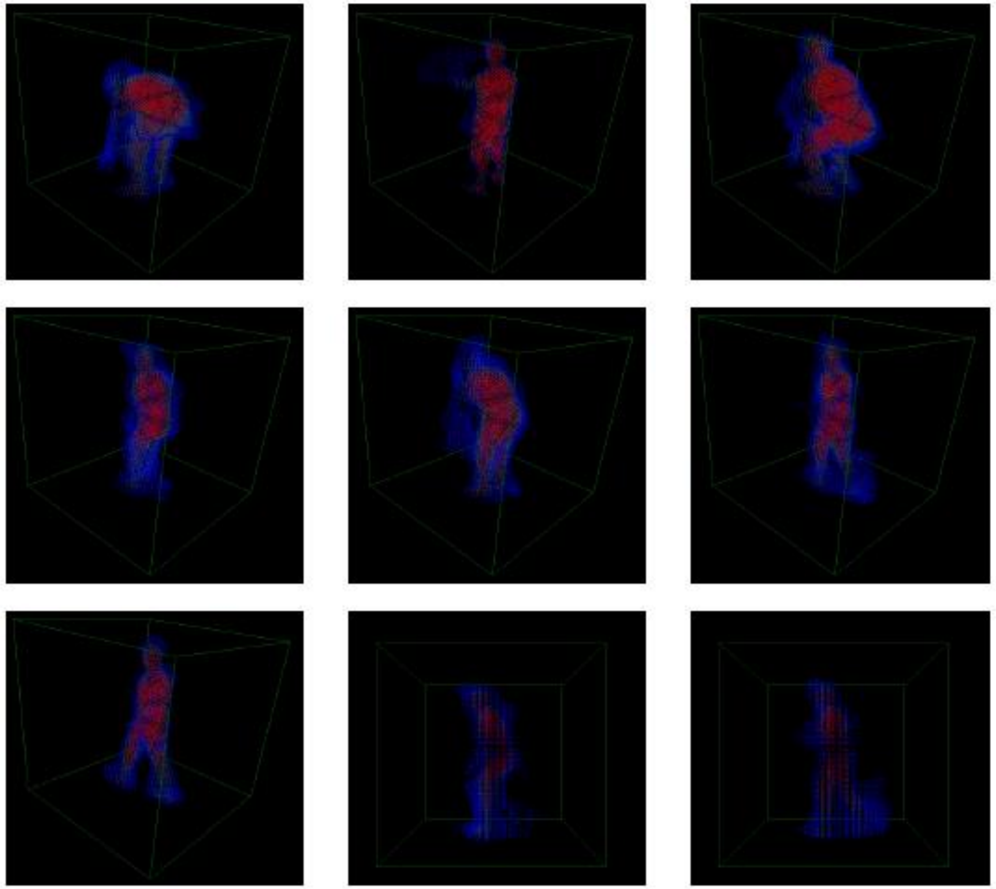
3D action recognition

- Action recognition on 3D data:
  - Extension of the video-based activity recognition algorithm.
  - Input to the algorithm: binary voxel-based representation of frames.



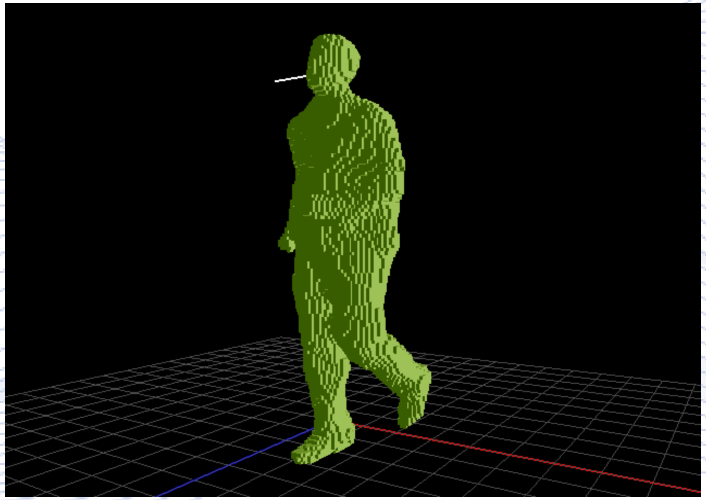
3D action recognition

- 3D action characterization based on 3D "dynemes" (representative poses) derived through clustering along with LDA.



3D action recognition

- Issue: Bodies should be consistently oriented in 3D space.
- Solution: use a body-attached coordinate system.
  - Vertical body axis, axis pointing from the user forward.



Human Action Recognition

- Human Action Recognition definition and data
- Classical Human Action Recognition
  - Single view Human Action Recognition
  - Multiview Human Action Recognition
- Neural Human Action Recognition
- GCN Human Action Recognition
- 3D Human Action Recognition
- Human Action Recognition applications

Special HAR cases

ASSISTED LIVING

GAIT RECOGNITION

CROWD ANALYSIS

Eating/drinking activity recognition

Very important in ***assisted living***.

- ***Eating/drinking/apraxia*** can be recognized based on the relative hand/face motion.
- Color image segmentation can be used to segment hands and face regions.

Fall Detection

"Video-based Human Fall Detection in Smart Homes Using Deep Learning"
[SHO2018]

Objective:

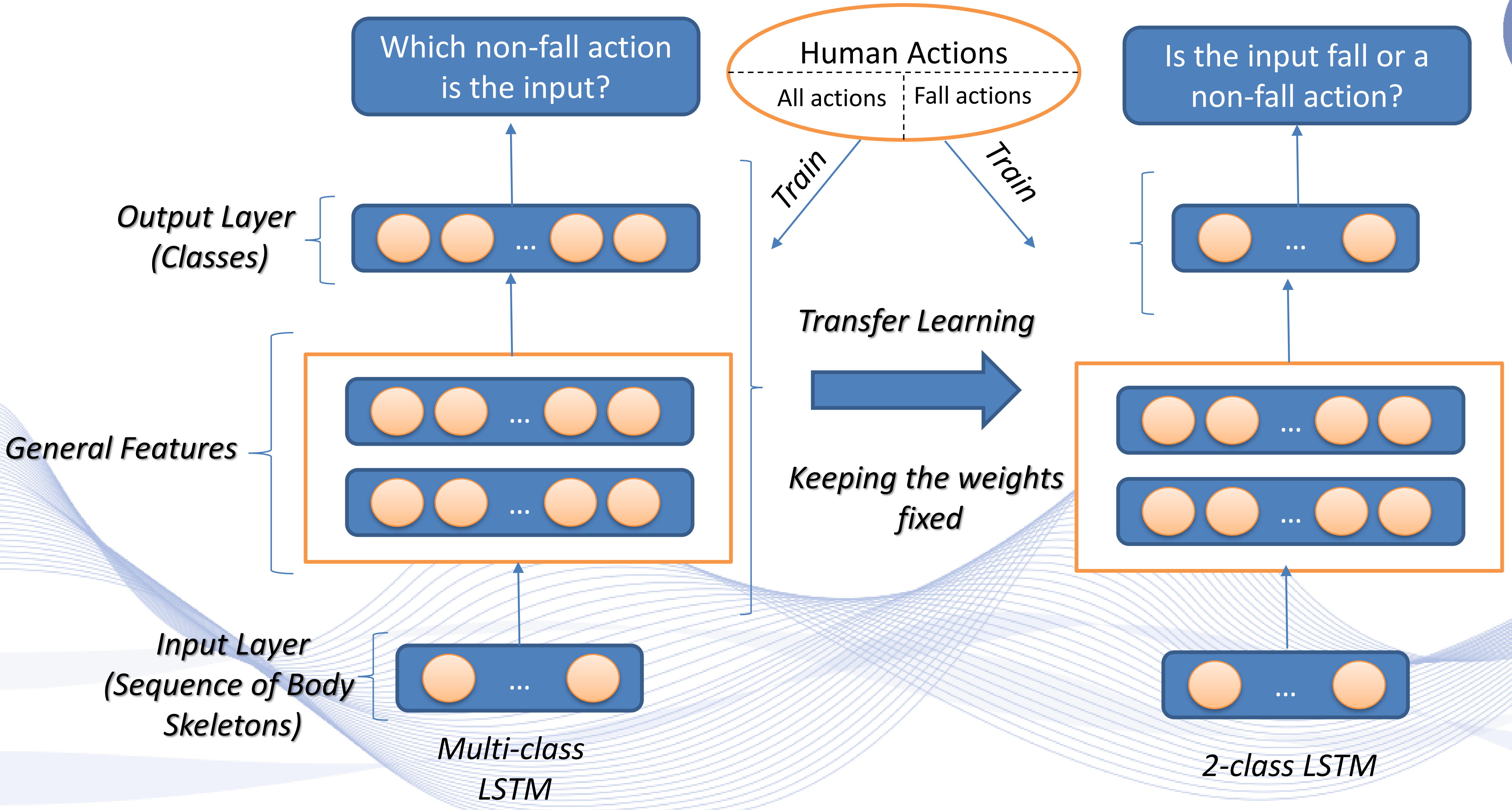
- Fall detection system using skeleton data.

Methodology:

- ***Transfer learning*** is proposed by training on massive generic action data and then fine tune in fall detection datasets.
- Only depth maps are used from which they obtain the skeleton data.
- The training is using LSTMs. *The LSTM + skeleton data is a popular choice.*

Experiments & Accuracy:

- NTU RGB+D Action Recognition Dataset (accuracy 0.9323%)


Which non-fall action is the input?
Human Actions
All actions
Fall actions
Is the input fall or a non-fall action?
Train
Train
Output Layer (Classes)
Transfer Learning
General Features
Keeping the weights fixed
Input Layer (Sequence of Body Skeletons)
Multi-class LSTM
2-class LSTM

Gait Recognition

The usage of gait as a biometric is a relatively new area of study that is gaining attention. ***Why is that?***

1. ***It can be required from a very long distance***

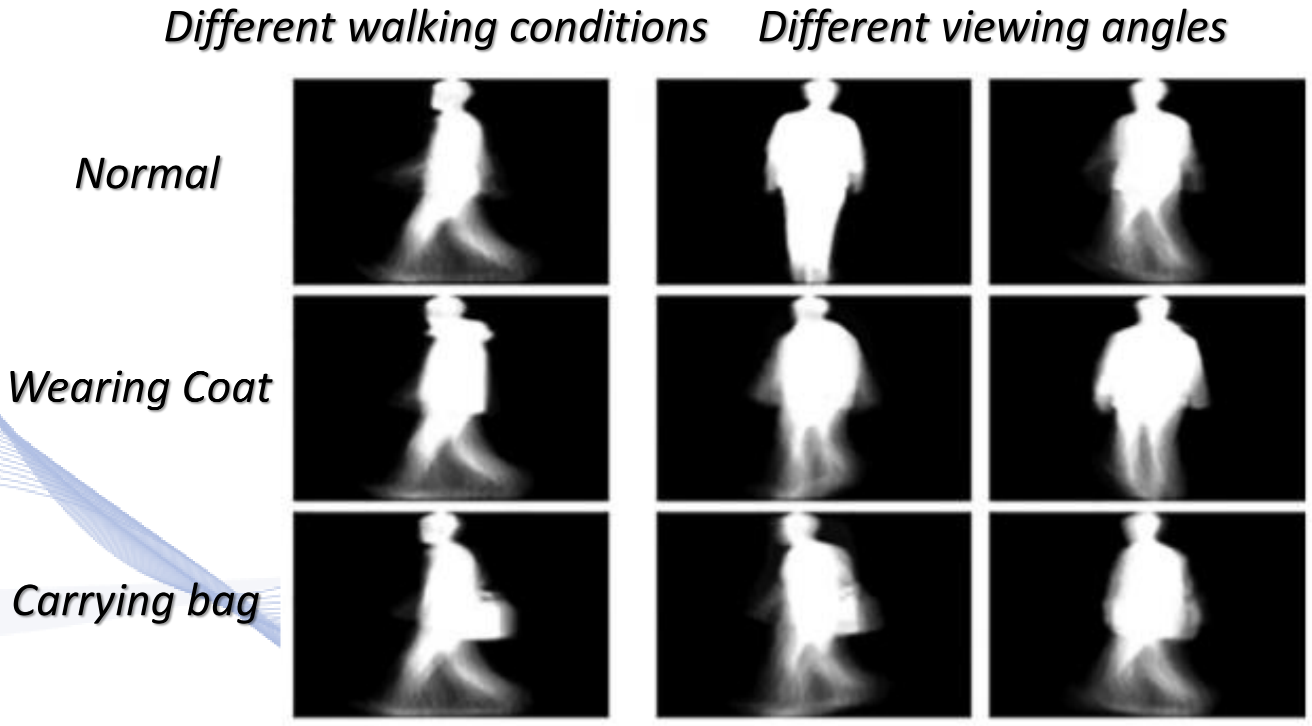
Unlike other biometrics such as face recognition, which need sufficient face size.

2. ***It is difficult to steal or fake***

Deep learning is powered by data & continue to increase their performance as new training data are given.

However, the gait of a person is not invariant to the capturing viewpoint and it can vary due to the clothes, footwear, walking surface, walking speed or emotional condition of the subject in discussion.

Gait Energy Image (GEI)



$$GEI(x, y) = \frac{1}{s} \sum_{t-1}^{s} F^t(x, y),$$

- s: total number of frames that represent one gait cycle,
- $F^t(x, y)$: binary silhouette of the subject in time t.

[ALO2017].

What is Crowd Analysis?

Crowd analysis is frequently used. In general, the attributes of crowd to be considered for this analysis are:

- crowd counting
- crowd motion detection
- crowd tracking
- crowd behavior understanding → abnormal event detection

Abnormal Detection Example Work

“Abnormal event detection in videos using generative adversarial nets”
[RAV2017]

Objective:

- Detect abnormal events in crowds using GANs.

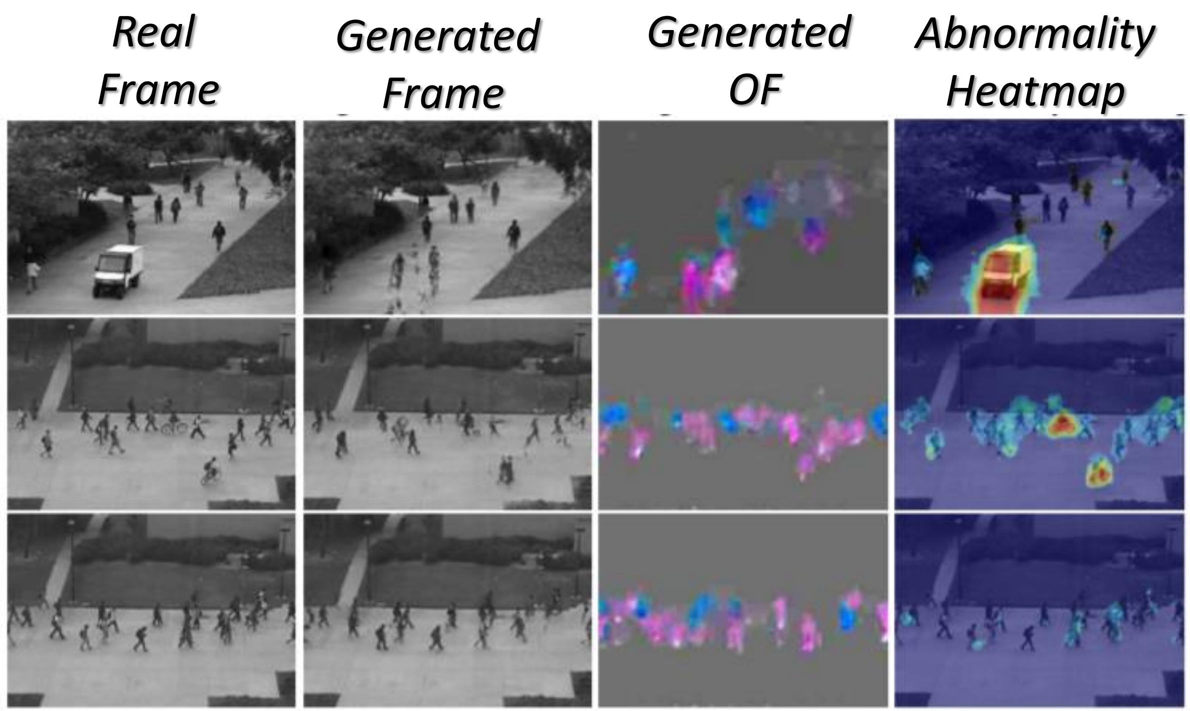
Methodology:

- Two Conditional GAN networks are trained: $N^{F \to O}$, which generates optical flow from frames and $N^{O \to F}$, which generates frames from optical-flow.
- At testing time, the real data are compared with both the appearance and the motion representations reconstructed by the two GANs and abnormal areas are detected by computing local differences.

Experiments & Accuracy:

- UCSD Ped1 (97.4%), Ped2 (93.5%) and UMN dataset (99%)

Abnormal Detection Example Work



The training pairs of frame-optical flow & images, $X = \{(F_t, O_t)\}$ are collected using only the frames of the *normal videos.*

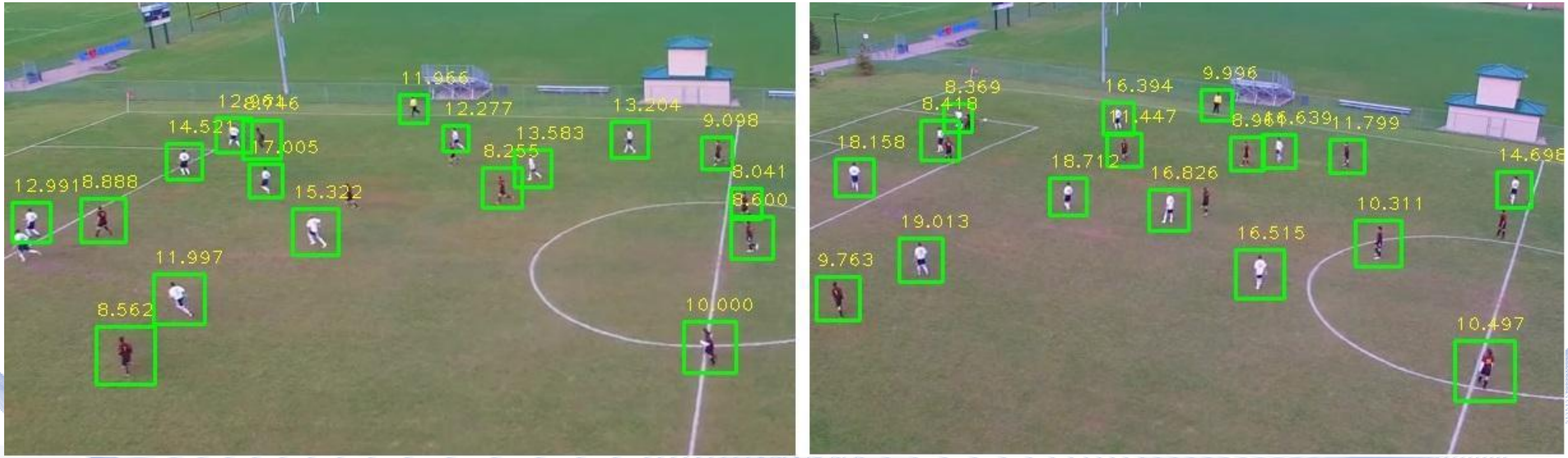
In testing time, the generators $G^{F\rightarrow O}$ and $G^{O\rightarrow F}$ fail to reconstruct abnormal events.

[RAV2017]

Sports: special case of human action

- Biomechanics
- Sports data analysis.

Sports: special case of human action



Player data analysis. Computer-assisted coaching.

Artificial Intelligence & Information Analysis Lab

Bibliography

[PIT2021] I. Pitas, “Computer vision”, Createspace/Amazon, in press.

[PIT2017] I. Pitas, “Digital video processing and analysis” , China Machine Press, 2017 (in Chinese).

[PIT2013] I. Pitas, “Digital Video and Television” , Createspace/Amazon, 2013.

[NIK2000] N. Nikolaidis and I. Pitas, “3D Image Processing Algorithms”, J. Wiley, 2000.

[PIT2000] I. Pitas, “Digital Image Processing Algorithms and Applications”, J. Wiley, 2000.

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr