



Hochschule  
Bonn-Rhein-Sieg  
University of Applied Sciences



# Lifelong Action Learning for Socially Assistive Robots

November 25th, 2022

Hasnainali Walli

*Advisors*

Prof. Dr. Paul G. Plöger, Prof. Dr. Sebastian Houben, Alex Mitrevski

## 1. Introduction

## 2. Comparative Analysis: Action Recognition

## 3. Comparative Analysis: Class-Incremental Learning



# Action Recognition



# Lifelong Action Learning



# Lifelong Action Learning

*Application in CRI*

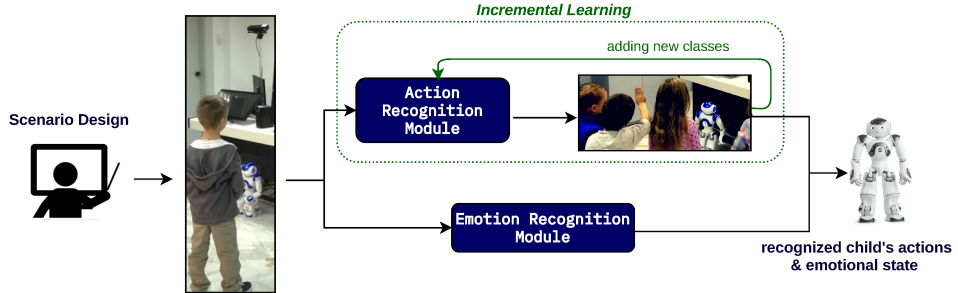


Hochschule  
Bonn-Rhein-Sieg  
University of Applied Sciences



# Lifelong Action Learning

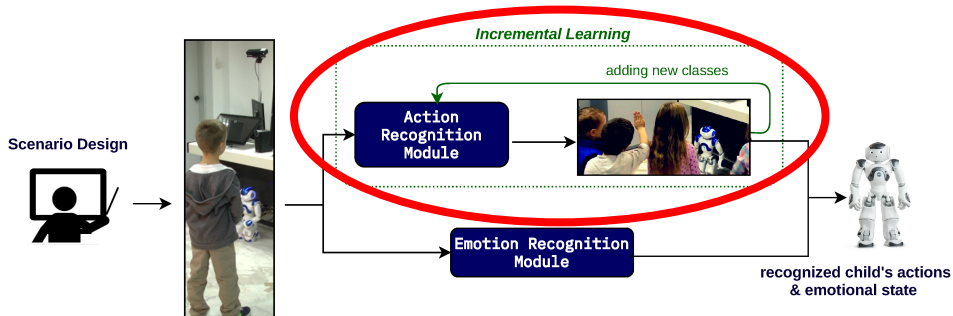
## Application in CRI



**Figure 1:** Incremental learning pipeline for action and emotion recognition. Source: Visual Robotic Perception System with Incremental Learning for Child–Robot Interaction Scenarios

# Lifelong Action Learning

## Application in CRI



**Figure 2:** Incremental learning pipeline for action and emotion recognition Source: Visual Robotic Perception System with Incremental Learning for Child–Robot Interaction Scenarios

## Their Approach

- RGB+D and Optical Flow data
- TSN Network
- iCaRL Algorithm
- BabyRobot Dataset

## Our Approach

- 3D Skeleton data
- CTR-GCN Network
- BiC Algorithm
- NTU RGB+D Dataset



# Our Approach

## *Methodology*

1. Perform comparative analysis on skeleton-based action recognition networks
  - CTR-GCN, MS-G3D, EfficientGCN, ViewAdaptive NN
  - Joint, Bone and Joint Motions
  - **Metrics**: Cross-Subject Accuracy, Training Time
2. Perform comparative analysis on class-incremental learning algorithms
  - LwF, iCaRL, LUCIR, BiC
  - 
  - **Metrics**: Task-Aware & Task-Agnostic Accuracy
3. Integrate final model on QTRobot
  - 
  -

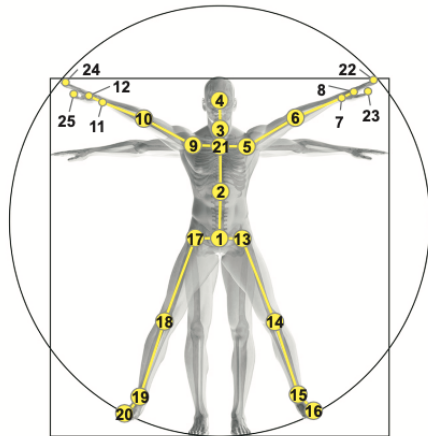
1. Introduction

2. Comparative Analysis: Action Recognition

3. Comparative Analysis: Class-Incremental Learning



# NTU Dataset



**Figure 3:** Human body joint configurations found in the NTU RGB-D dataset. Source: NTU RGB+D: A Large Scale Dataset for 3D Human Activity Analysis

# NTU Dataset

## *Subset of Actions*

Drink Water	Eat Meal	Brush Teeth	Drop
Pick Up	Throw	Sit Down	Stand Up
Clapping	Hand Waving	Kick Something	Hopping
Jump Up	Play with Phone/Tablet	Point to Something	Rub two Hands
Nod Head/Bow	Shake Head	Wipe Face	Cross Hands in Front

Table 1: Subset of action classes from the NTU RGB-D dataset

# Action Recognition Analysis Results

Network	Cross Subject	Cross View
CTR-GCN (Joint)	92.63%	96.37%
CTR-GCN (Bone)	92.78%	96.02%
CTR-GCN (Motion)	92.51%	96.40%
MS-G3D (Joint)	91.27%	96.85%
MS-G3D (Bone)	90.90%	95.44%
EfficientGCN-B4 (SG Layer)	94.05%	97.47%
EfficientGCN-B4 (EpSep Layer)	94.43%	97.56%
VA-NN (CNN)	92.97%	92.20%

Table 2: Model Accuracy

Network	Training Time
CTR-GCN	4 hrs
MS-G3D	8 hrs
EfficientGCN-B4	5 hrs
VA-NN (CNN)	0.5 hrs

Table 3: Model Training Time

# Action Recognition Analysis Results

Action	CTR-GCN	MS-G3D		Action	CTR-GCN	MS-G3D
Drink Water	82.48%	83.94%		Kick Something	97.83%	94.93%
Eat Meal	78.91%	73.82%		Hopping	98.91%	95.27%
Brush Teeth	90.84%	91.21%		Jump Up	98.91%	98.55%
Drop	90.18%	91.64%		Play with Phone/Tablet	86.91%	90.91%
Pick Up	98.91%	94.55%		Point to Something	92.39%	92.03%
Throw	96.36%	90.91%		Rub two Hands	90.58%	89.49%
Sit Down	98.90%	97.80%		Nod Head/Bow	96.01%	95.65%
Stand Up	98.17%	98.90%		Shake Head	96.00%	95.64%
Clapping	82.42%	72.89%		Wipe Face	92.39%	94.20%
Hand Waving	94.16%	94.89%		Cross Hands in Front	93.84%	94.57%

Table 4: Cross-Subject accuracy results per class for CTR-GCN and MS-G3D models

1. Introduction

2. Comparative Analysis: Action Recognition

3. Comparative Analysis: Class-Incremental Learning



# Incremental Learning

## *Concepts*





# Incremental Learning

## *Performance Metrics*



# Incremental Learning Analysis

## Task Sequence

Task #	Action		Task #	Action		Task #	Action
Task 1	Wipe Face		Task 5	Pick Up		Task 8	Nod Head/Bow
	Eat Meal			Brush Teeth			Hopping
Task 2	Cross Hands in Front		Task 6	Throw		Task 9	Drop
	Clapping			Point to Something			Drink Water
Task 3	Kick Something		Task 7	Hand Waving		Task 10	Rub two Hands
	Shake Head			Stand Up			Jump Up
Task 4	Sit Down						
	Play with Phone/Tablet						

Table 5: Task sequence for class-IL comparative analysis














# Bibliography

*T<sub>E</sub>X*, L<sup>A</sup>T<sub>E</sub>X, and Beamer

-  Donald E. Knuth. **The T<sub>E</sub>Xbook**. Addison-Wesley, 1984.
-  Leslie Lamport. **L<sup>A</sup>T<sub>E</sub>X: A Document Preparation System**. Addison-Wesley, 1986.
-  M. Goossens, F. Mittelbach, and A. Samarin. **The L<sup>A</sup>T<sub>E</sub>X Companion**. Addison-Wesley, 1994.
-  Till Tantau. **User's Guide to the Beamer Class Version 3.01**. Available at <http://latex-beamer.sourceforge.net>.
-  A. Mertz and W. Slough. Edited by B. Beeton and K. Berry. **Beamer by example** In TUGboat, Vol. 26, No. 1., pp. 68-73.