Spot the Pain: Exploring the Application of Skeleton Pose Estimation for Automated Pain Assessment

Author: Angelica Hjelm Gardner
Supervisor: Welf Löwe
Course code: 5DV50E
Date: June 3, 2022

INTRODUCTION

Foundation of pain management

Focus on describing the pain

Automated pain assessment

Reliable, objective, continuous monitoring

At least one input modality

PAIN INDICATORS

Behavioural

Facial expressions, Body gestures, Paralinguistic vocalisation

Physiological

Brain activity, Cardiovascular activity, Skin conductance response

Most focus on facial expressions

Uni-, Bi- and Multimodality

No research focus on body movements

BODY MOVEMENT REPRESENTATION

No formal standard

Facial Action Coding System (FACS)

Pain-related facial muscle movements

BODY MOVEMENT REPRESENTATION

No formal standard

Facial Action Coding System (FACS)

Pain-related facial muscle movements

BODY MOVEMENT REPRESENTATION

No formal standard

Facial Action Coding System (FACS)

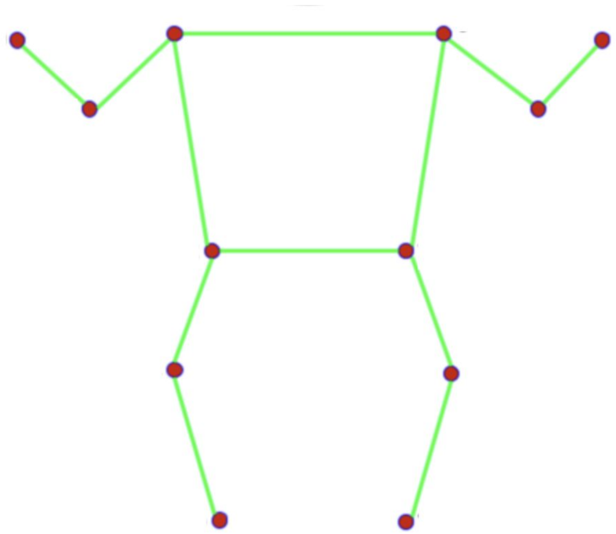
Pain-related facial muscle movements

Abrupt actions, limping, hesitation, stiffness

Skeleton avatar model

SKELETON POSE ESTIMATION

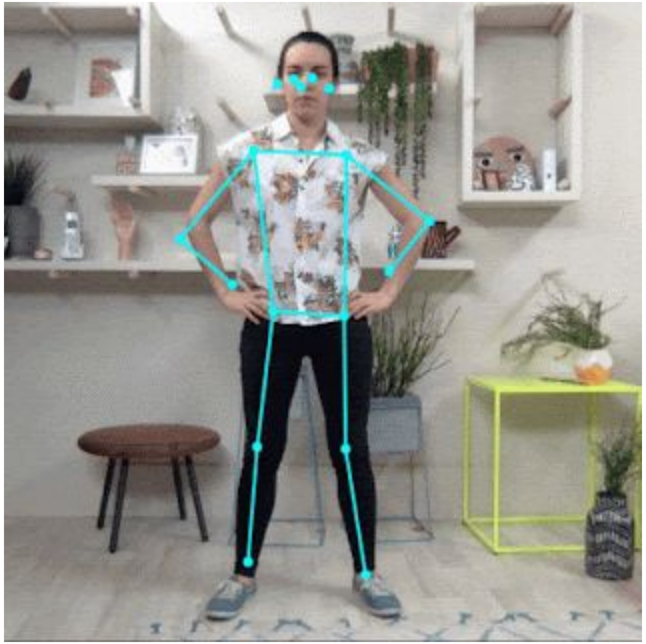
Kinematic method



SKELETON POSE ESTIMATION

Kinematic method

Sequence of movements



SKELETON POSE ESTIMATION

Kinematic method

Sequence of movements

Machine learning model

Learn pain-related patterns

Skeleton pose representation

RESEARCH QUESTIONS

RQ1

What pain assessment performance do we achieve when using skeleton pose estimation to represent body movements as the only pain indicator in a system?

RQ2

Can skeleton pose estimation identify areas of pain in the human body?

RQ3

Does including body movement data improve pain assessment performance in a bimodal approach?

METHOD DESCRIPTION

Machine learning experiments

Independent variables:

- Input modality
- Model architecture
- Model hyperparameters
- Bimodal approaches
- Experimental objectives

Model performance on unseen data

METHOD DESCRIPTION

Machine learning experiments

Independent variables:

- Input modality
- Model architecture
- Model hyperparameters
- Bimodal approaches
- Experimental objectives

1. Body
2. Face

Model performance on unseen data

METHOD DESCRIPTION

Machine learning experiments

Independent variables:

- Input modality
- Model architecture
- Model hyperparameters
- Bimodal approaches
- Experimental objectives

1. CNN-LSTM
2. RCNN

Model performance on unseen data

METHOD DESCRIPTION

Machine learning experiments

Independent variables:

- Input modality
- Model architecture
- Model hyperparameters
- Bimodal approaches
- Experimental objectives

Different model configurations

Model performance on unseen data

METHOD DESCRIPTION

Machine learning experiments

Independent variables:

- Input modality
- Model architecture
- Model hyperparameters
- Bimodal approaches
- Experimental objectives

1. Early Fusion
2. Late Fusion
3. Ensemble

Model performance on unseen data

METHOD DESCRIPTION

Machine learning experiments

Independent variables:

- Input modality
- Model architecture
- Model hyperparameters
- Bimodal approaches
- Experimental objectives

1. Pain recognition
2. Pain intensity estimation
3. Pain area classification

Model performance on unseen data

METHOD DESCRIPTION

Machine learning experiments

Independent variables:

- Input modality
- Model architecture
- Model hyperparameters
- Bimodal approaches
- Experimental objectives

Model performance on unseen data

METHOD DESCRIPTION

Machine learning experiments

Independent variables:

- Input modality
- Model architecture
- Model hyperparameters
- Bimodal approaches
- Experimental objectives

Model performance on unseen data

Performance metrics

METHOD DESCRIPTION

Machine learning experiments

Independent variables:

- Input modality
- Model architecture
- Model hyperparameters
- Bimodal approaches
- Experimental objectives

Model performance on unseen data

Performance metrics

Approach	Objective	Model architecture
Unimodal (Body)	Pain recognition	CNN-LSTM
Unimodal (Body)	Pain recognition	RCNN
Unimodal (Body)	Pain intensity	CNN-LSTM
Unimodal (Body)	Pain intensity	RCNN
Unimodal (Body)	Pain area	CNN-LSTM
Unimodal (Body)	Pain area	RCNN
Unimodal (Face)	Pain recognition	CNN-LSTM
Unimodal (Face)	Pain recognition	RCNN
Unimodal (Face)	Pain intensity	CNN-LSTM
Unimodal (Face)	Pain intensity	RCNN
Unimodal (Face)	Pain area	CNN-LSTM
Unimodal (Face)	Pain area	RCNN
Bimodal (Early Fusion)	Pain recognition	CNN-LSTM
Bimodal (Early Fusion)	Pain recognition	RCNN
Bimodal (Early Fusion)	Pain intensity	CNN-LSTM
Bimodal (Early Fusion)	Pain intensity	RCNN
Bimodal (Early Fusion)	Pain area	CNN-LSTM
Bimodal (Early Fusion)	Pain area	RCNN
Bimodal (Late Fusion)	Pain recognition	CNN-LSTM
Bimodal (Late Fusion)	Pain recognition	RCNN
Bimodal (Late Fusion)	Pain intensity	CNN-LSTM
Bimodal (Late Fusion)	Pain intensity	RCNN
Bimodal (Late Fusion)	Pain area	CNN-LSTM
Bimodal (Late Fusion)	Pain area	RCNN
Ensemble	Pain recognition	all unimodal approaches
Ensemble	Pain intensity	all unimodal approaches
Ensemble	Pain area	all unimodal approaches

DATASET

Real-world dataset provided by AIMO[1]

1059 videos from 807 participants

Overhead deep squat

Self-assessment questionnaire about pain

[1] https://www.aimo-fit.com/

OBJECTIVES

Pain recognition:

Pain vs. no pain

Pain intensity estimation:

1-3 (mild pain), 4-7 (moderate pain), 8-10 (severe pain)

Pain area classification:

Head and neck, Upper body, Lower body, Back region

MODEL ARCHITECTURES

Spatial and temporal information

Deep learning algorithms: CNN & RNN

Hybrid CNN-LSTM architecture

- Convolutional layers for representing spatial information
- Recurrent layers for temporal information

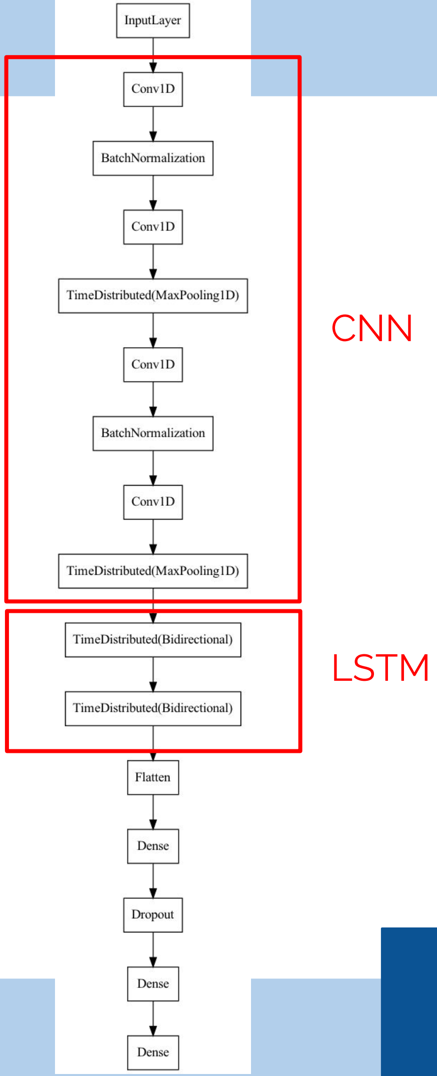
MODEL ARCHITECTURES

Spatial and temporal information

Deep learning algorithms: CNN & RNN

Hybrid CNN-LSTM architecture

- Convolutional layers for representing spatial information
- Recurrent layers for temporal information



MODEL ARCHITECTURES

- Spatial and temporal information
- Deep learning algorithms: CNN & RNN
- Hybrid CNN-LSTM architecture
  - Convolutional layers for representing spatial information
  - Recurrent layers for temporal information
- Recurrent CNN (RCNN) architecture
  - Recurrent convolutional layers

MODEL ARCHITECTURES

Spatial and temporal information

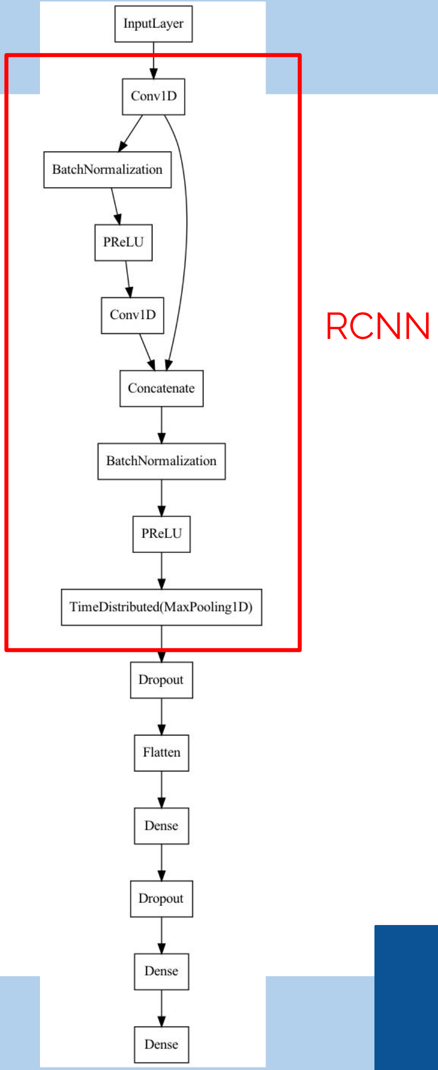
Deep learning algorithms: CNN & RNN

Hybrid CNN-LSTM architecture

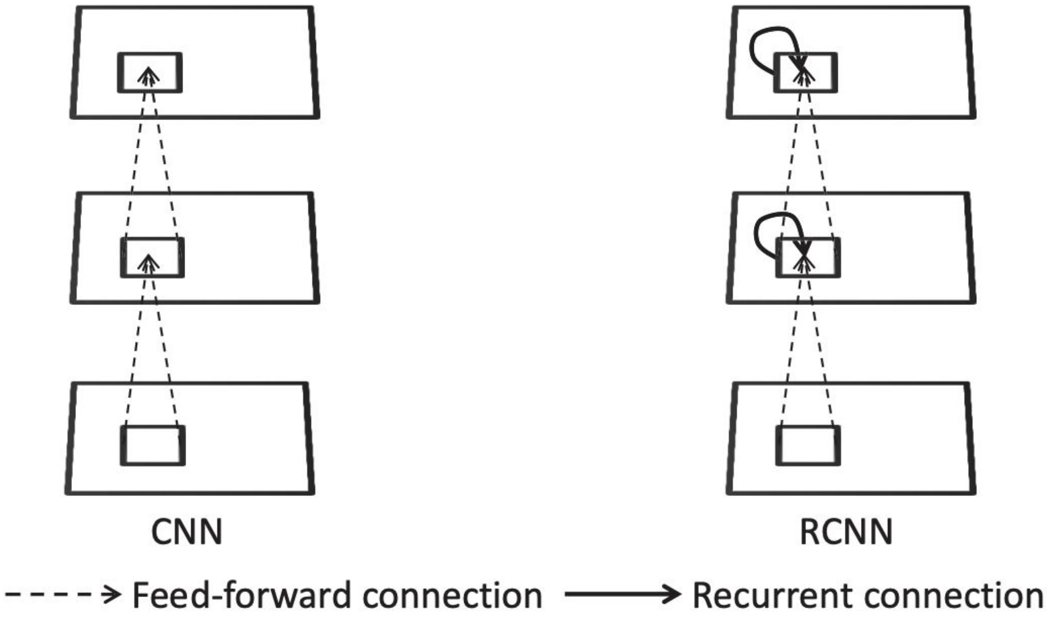
- Convolutional layers for representing spatial information
- Recurrent layers for temporal information

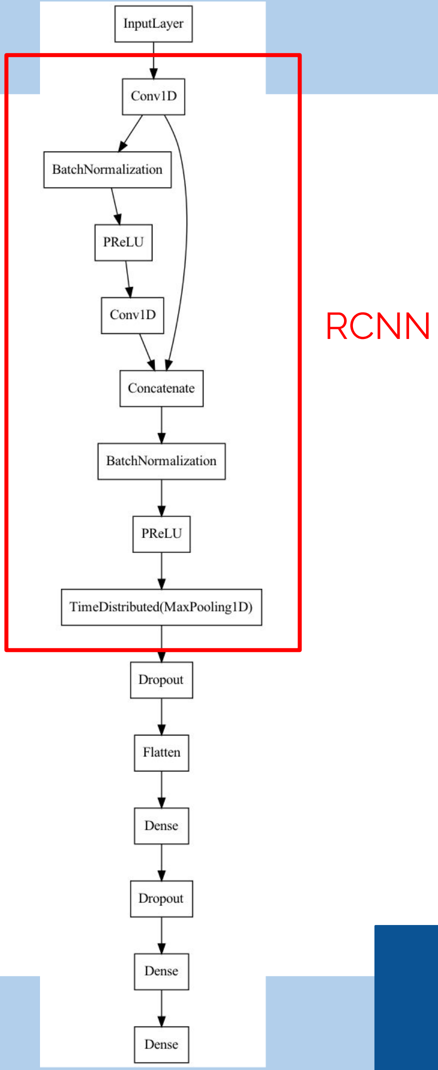
Recurrent CNN (RCNN) architecture

- Recurrent convolutional layers



MODEL ARCHITECTURES





BIMODAL FUSION STRATEGIES

Early Fusion, feature-level

Late Fusion, decision-level

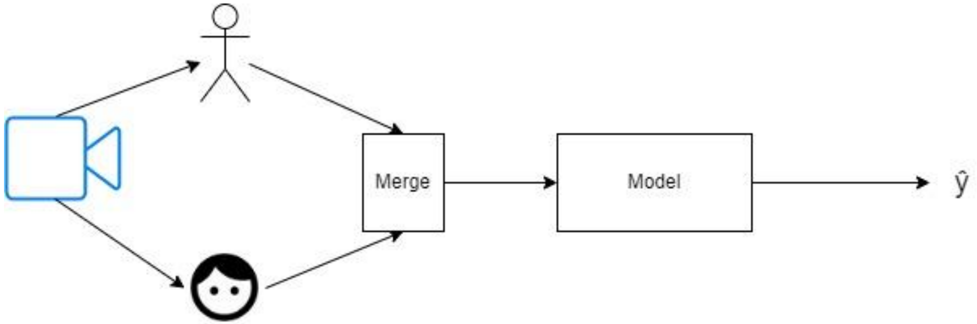
Ensemble learning, weighted average

BIMODAL FUSION STRATEGIES

Early Fusion, feature-level

Late Fusion, decision-level

Ensemble learning, weighted average

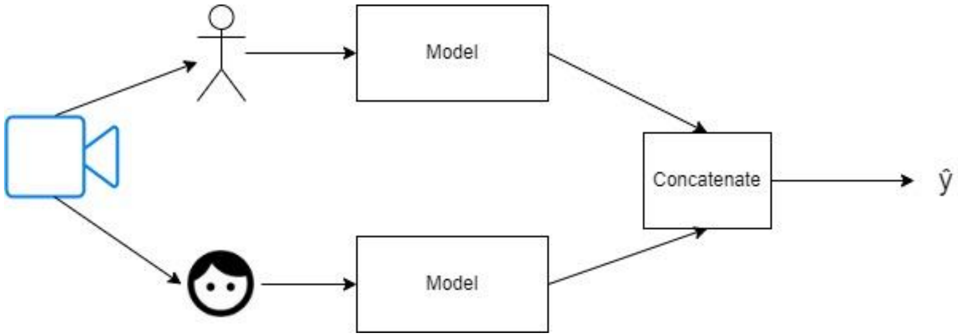


BIMODAL FUSION STRATEGIES

Early Fusion, feature-level

Late Fusion, decision-level

Ensemble learning, weighted average

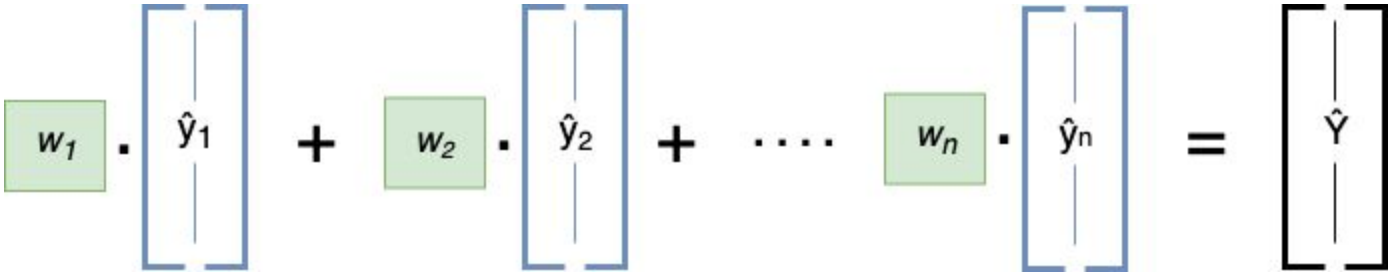


BIMODAL FUSION STRATEGIES

Early Fusion, feature-level

Late Fusion, decision-level

Ensemble learning, weighted average



AREA UNDER THE CURVE (AUC)

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

Used for comparison

0.0 - 1.0

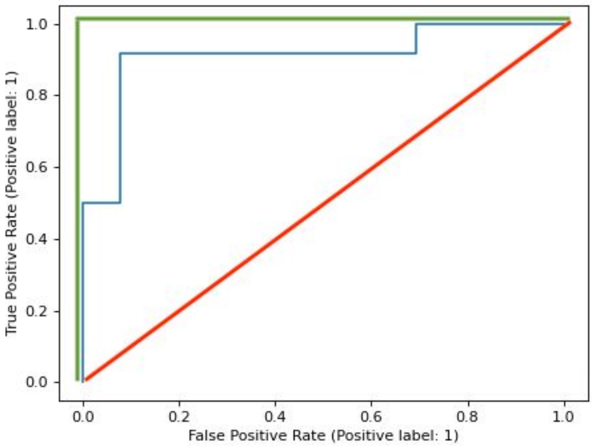
Models predict probabilities

0.95, 0.03, 0.6

Default threshold = 0.5

Performance across all thresholds

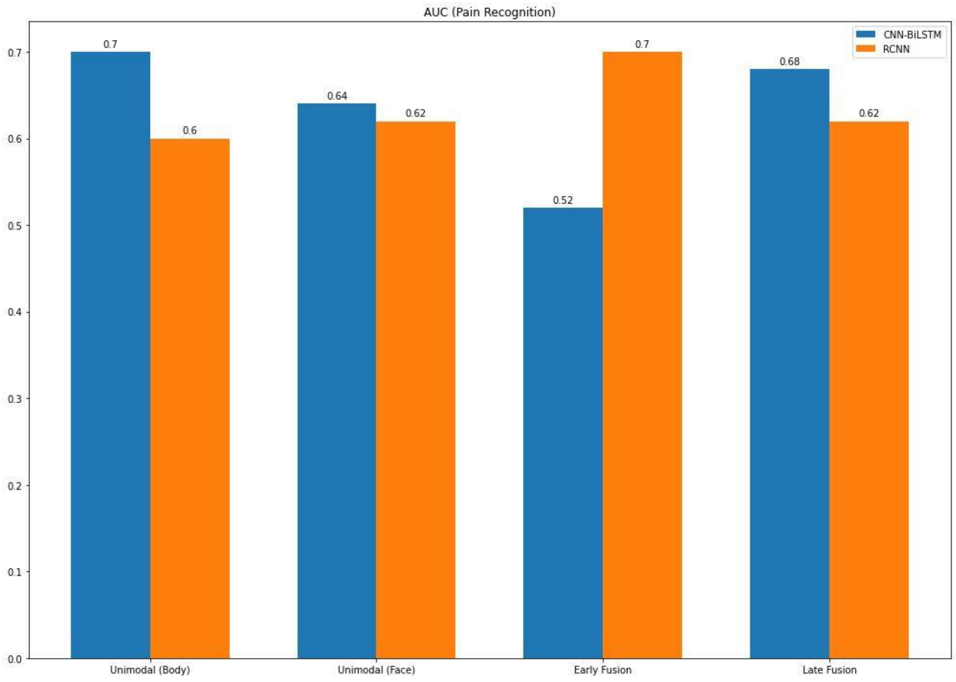
ROC curve



PAIN RECOGNITION

Binary classification:

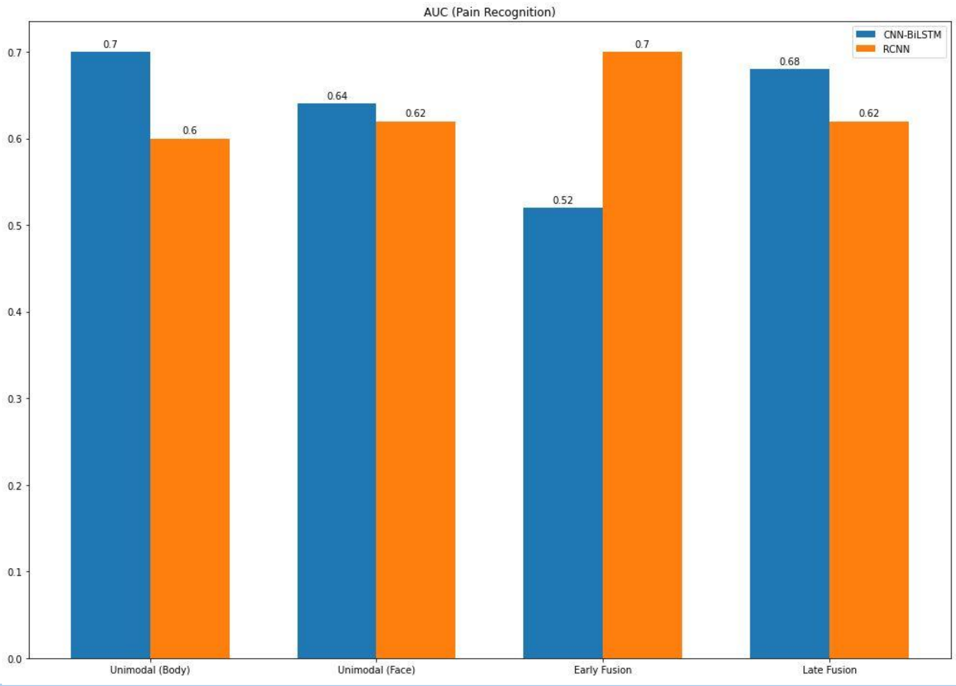
Pain vs. no pain



PAIN RECOGNITION

Binary classification:

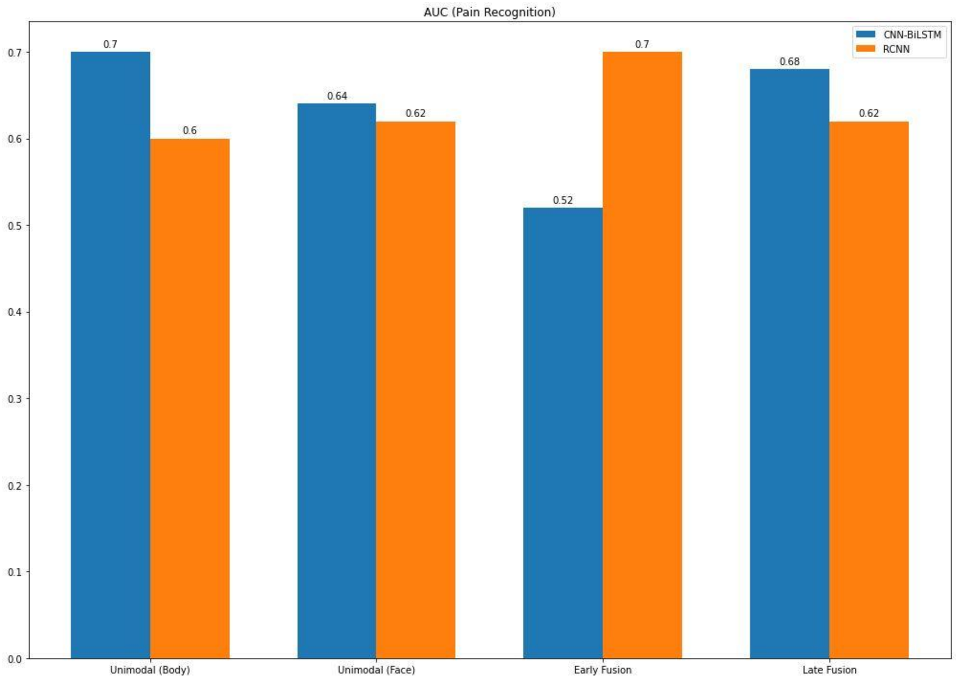
Pain vs. no pain



PAIN RECOGNITION

Binary classification:

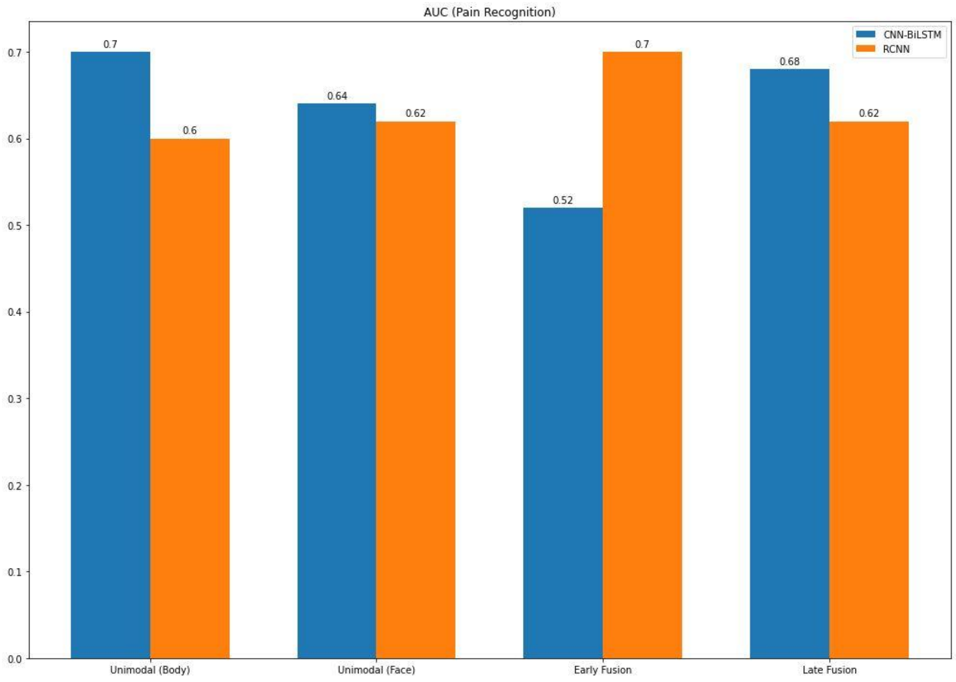
Pain vs. no pain



PAIN RECOGNITION

Binary classification:

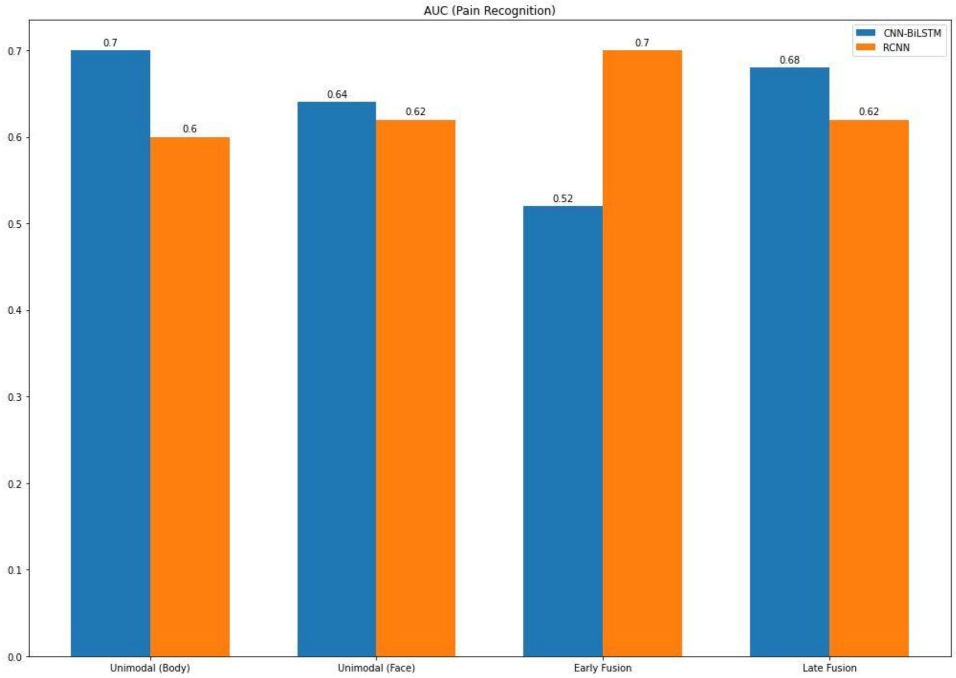
Pain vs. no pain



PAIN RECOGNITION

Binary classification:

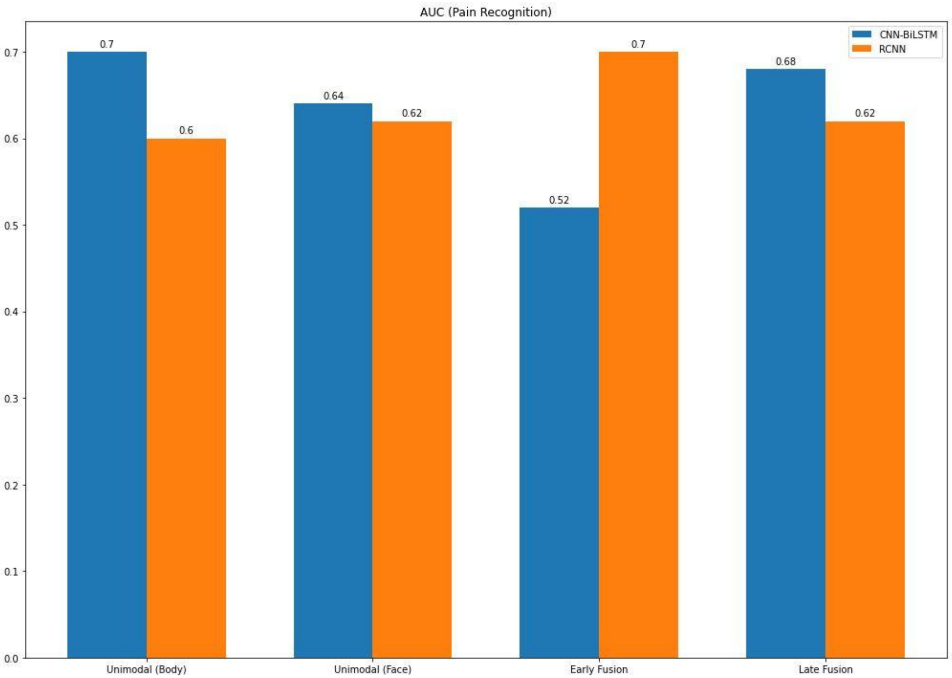
Pain vs. no pain



PAIN RECOGNITION

Binary classification:

Pain vs. no pain

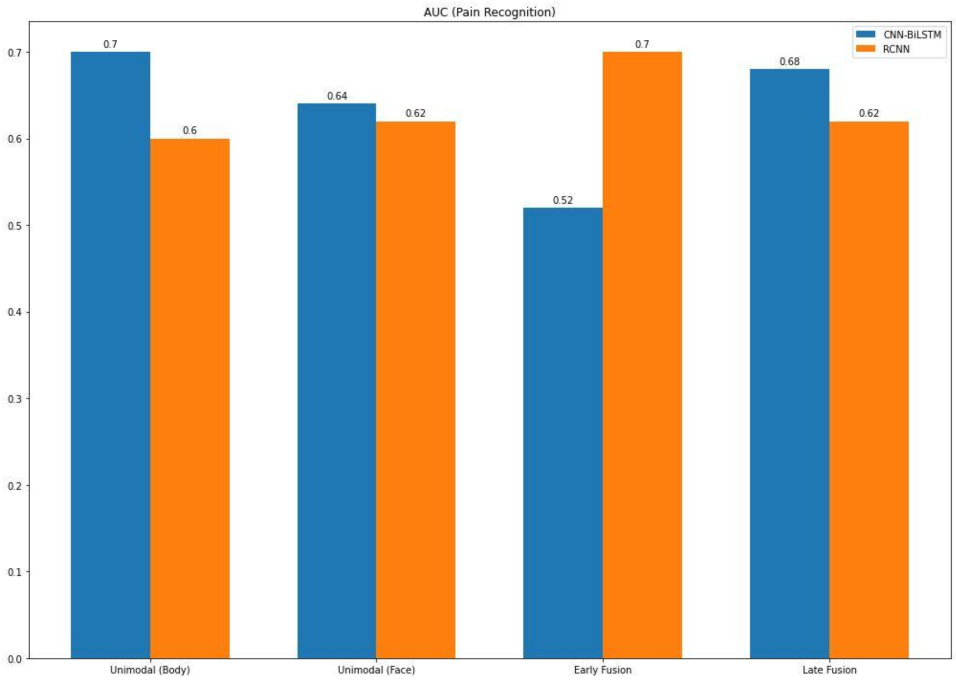


PAIN RECOGNITION

Binary classification:

Pain vs. no pain

Small improvements



PAIN RECOGNITION

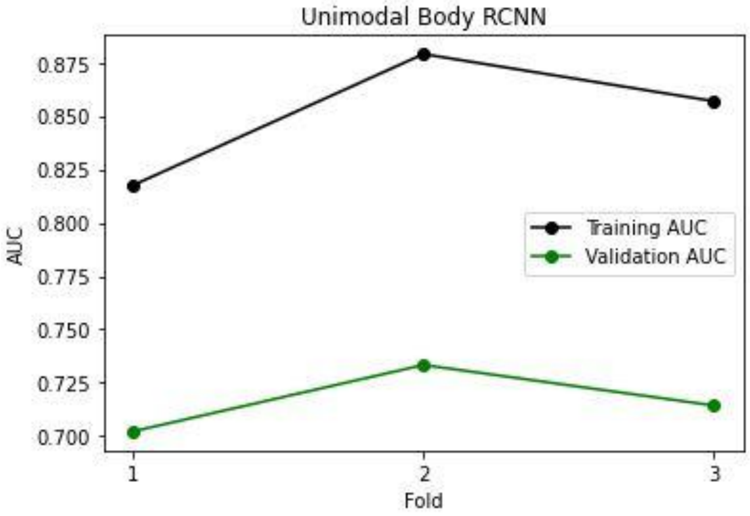
Binary classification:

Pain vs. no pain

Small improvements

Training performance

Individual responses



PAIN RECOGNITION

Binary classification:

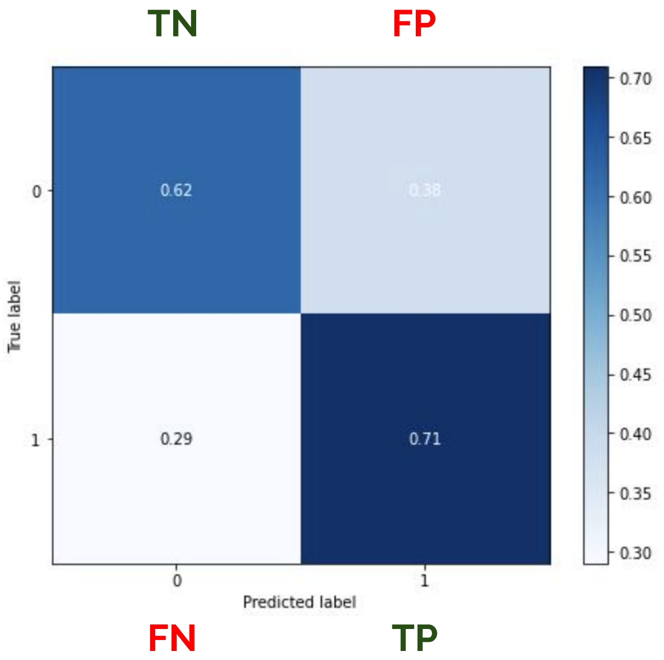
Pain vs. no pain

Small improvements

Training performance

Individual responses

Distinguishable patterns

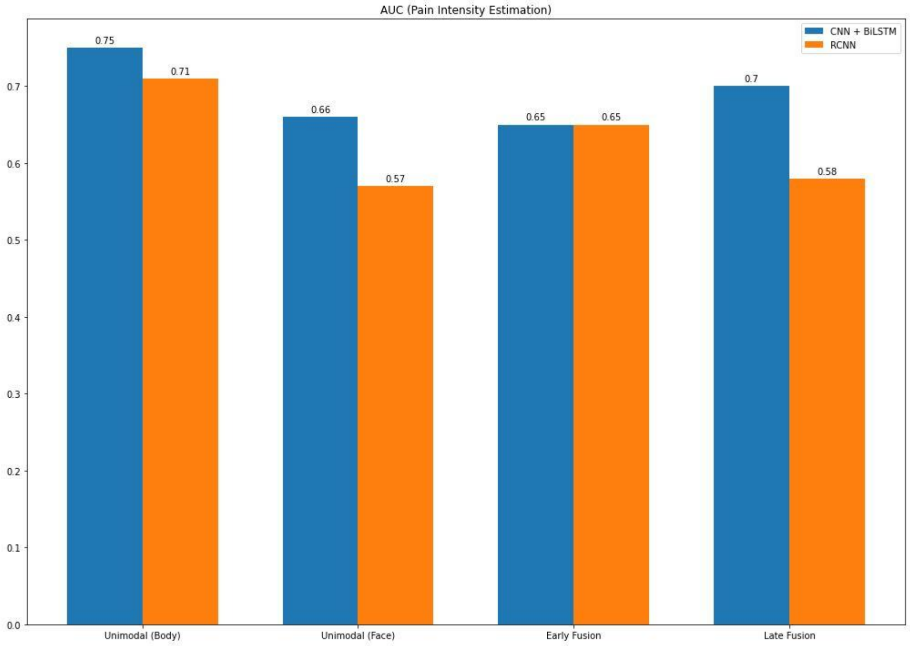


Confusion matrix

PAIN INTENSITY ESTIMATION

Multiclass classification:

Mild, Moderate, Severe



PAIN INTENSITY ESTIMATION

Multiclass classification:

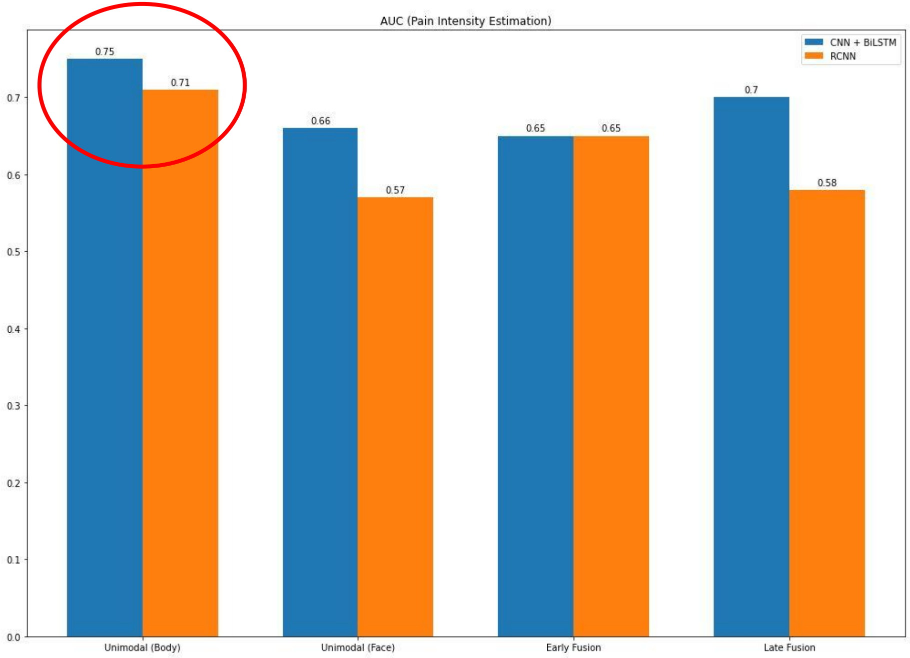
Mild, Moderate, Severe

Best-performing

Too small face in videos

Greater availability

Real-world data

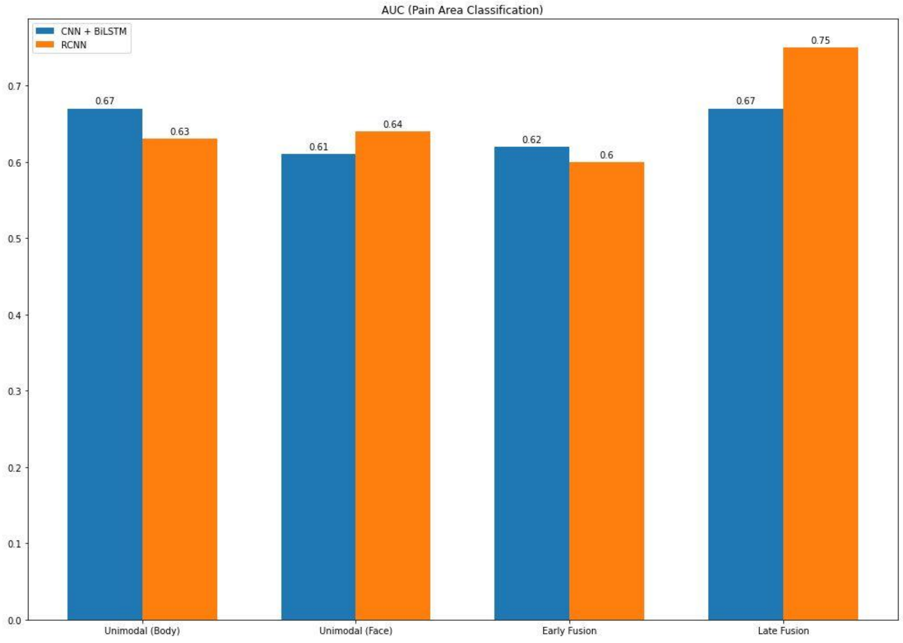


PAIN AREA CLASSIFICATION

Multiclass classification:

Back Region, Head and Neck

Lower Body, Upper Body



PAIN AREA CLASSIFICATION

Multiclass classification:

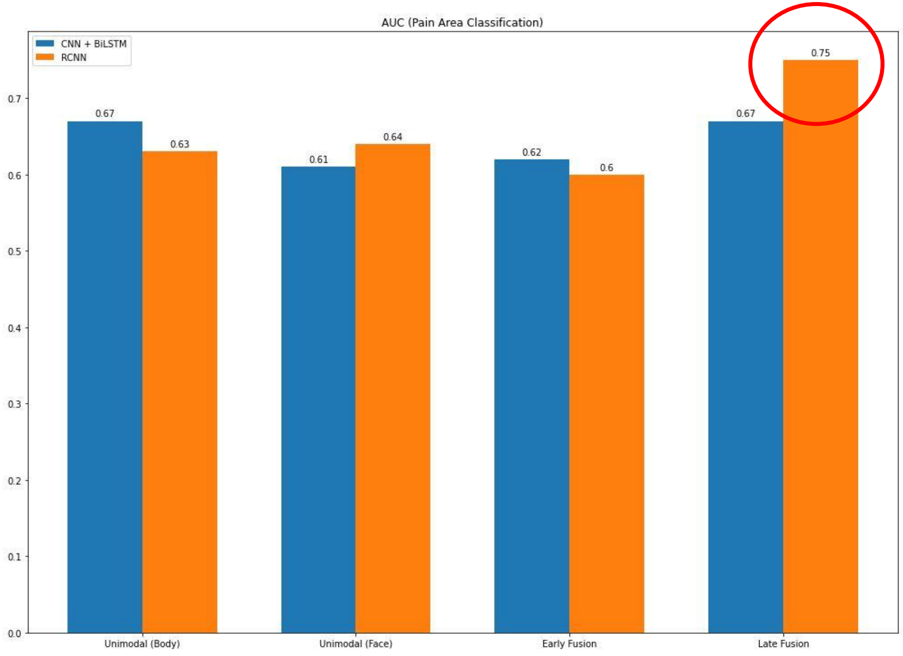
Back Region, Head and Neck

Lower Body, Upper Body

Best-performing model

Never used before

Multioutput classification



BODY CONTRIBUTION TO BIMODAL APPROACH

Body weight by ensembles

Low performance for two objectives

Metric	Pain Recognition	Pain Intensity	Pain Area
Accuracy	100.00%	53.68%	39.39%
AUC	0.71	0.61	0.54
Precision	1.00	0.40	0.31
Recall	1.00	0.38	0.30
F-1 Score	1.00	0.30	0.26

BODY CONTRIBUTION TO BIMODAL APPROACH

Body weight by ensembles

Low performance for two objectives

```
Body CNN-BiLSTM (weight: 0.0928)
Face CNN-BiLSTM (weight: 0.8981)
Body RCNN (weight: 0.0073)
Face RCNN (weight: 0.0019)
```

Area

```
Body CNN-BiLSTM (weight: 0.0037)
Face CNN-BiLSTM (weight: 0.8544)
Body RCNN (weight: 0.1301)
Face RCNN (weight: 0.0117)
```

Intensity

Metric	Pain Recognition	Pain Intensity	Pain Area
Accuracy	100.00%	53.68%	39.39%
AUC	0.71	0.61	0.54
Precision	1.00	0.40	0.31
Recall	1.00	0.38	0.30
F-1 Score	1.00	0.30	0.26

BODY CONTRIBUTION TO BIMODAL APPROACH

Body weight by ensembles

Low performance for two objectives

Best-performing pain recognition

Performance across all thresholds

```
Body CNN-BiLSTM (weight: 0.6587)
Face CNN-BiLSTM (weight: 0.2500)
Body RCNN (weight: 0.0899)
Face RCNN (weight: 0.0014)
```

Recognition

Metric	Pain Recognition	Pain Intensity	Pain Area
Accuracy	100.00%	53.68%	39.39%
AUC	0.71	0.61	0.54
Precision	1.00	0.40	0.31
Recall	1.00	0.38	0.30
F-1 Score	1.00	0.30	0.26

CONCLUSIONS

Basis for future studies

Promising unimodal body results (**RQ1**)

Pain area classification opportunity (**RQ2**)

Results show an affirmative indication (**RQ3**)

FUTURE WORK

Multioutput classification

Explainable AI (XAI) and feature importance

Combinations with other modalities

- Audio, Physiological signals (e.g. muscle activity)

Different architectures and fusion strategies

Thank you!