In The Name of Allah, The Most Beneficent, The Most Merciful

Deep Multimodal Pain Recognition: A Database and Comparison of Spatio-Temporal Visual Modalities

Mohammad A. Haque, Ruben B. Bautista, Fatemeh Noroozi
Kaustubh Kulkarni, Christian B. Laursen, Ramin Irani, Marco Bellantonio,
Sergio Escalera, Golamreza Anbarjafari, Kamal Nasrollahi, Ole K. Andersen,
Erika G. Spaich, Thomas B. Moeslund

VAP, CVC, SMI, CNAP, UT, AAU
Corresponding author’s email: mah@create.aau.dk

Contents

- Vision-based pain management challenges
- A novel database
- Exploitation of spatio-temporal information
- Exploitation of multimodality
- Remarks to our contributions

Pain Expression

- "Pain is associated with actual or potential tissue damage, or described in terms of such damage"- International Association for the Study of Pain (IASP)

- Needs to be managed as a-
  - Moral imperative
  - Professional responsibility
  - Duty of medical practitioners

- **Pain Expression**
  - Visually revealed in the face
  - A subset of facial expressions
  - Express severity of pain

Pain is an unpleasant sensory and emotional experience

Pain expression example from UNBC database

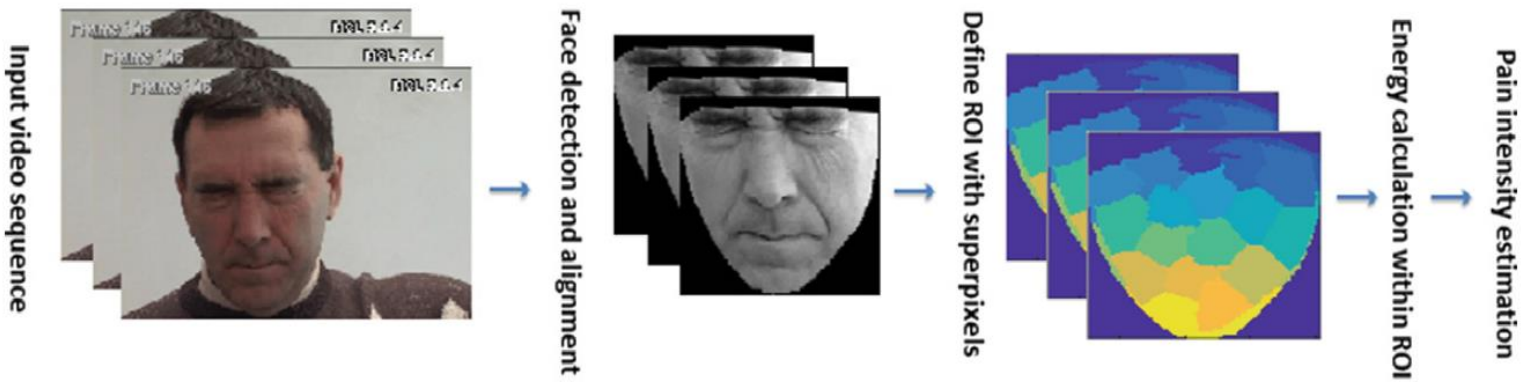
Visual Analysis of Pain

- The most primitive state of pain management is the assessment of pain.
  - Self-report
  - Visual inspection by experts
  - Automatic pain assessment from visual data
- Visual analysis of pain becomes more difficult to be correlated with self-report due to:
  - ‘Smiling in pain’
  - Social motives
  - Gender difference

Pain No pain

Researches going ON!

Two Challenges in Video-based Pain Detection



Ref: Dennis et al., Spatiotemporal Facial Super-Pixels for Pain Detection, 2016

- Exploiting both spatial and temporal information of the face to assess pain level
- Incorporating multiple visual modalities to capture complementary face information related to pain

But, we need DATABASE too!

Contribution of the paper

- We present
  - The first state-of-the-art publicly available database, ’Multimodal Intensity Pain (MIntPAIN)’ database, for RGBDT pain level recognition in sequences.
  - Baseline results including 5 pain levels recognition by analyzing independent visual modalities and their fusion
  - Employed state-of-the-art deep learning CNN+LSTM model to exploit spatio-temporal information

The Database

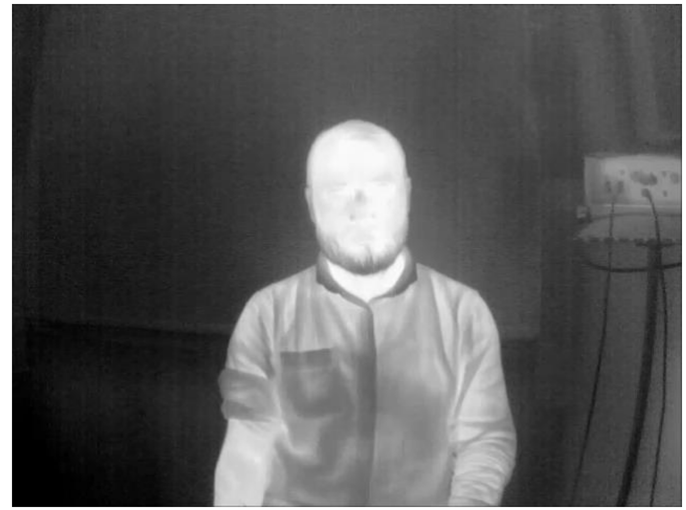
Comparison of the Publicly Available Video Databases on PAIN

Attribute	UNBC-McMaster database (2011)	BioVid database (2013)
No. of subjects	129 (16 are available)	90 (87 are available)
Subject's type	Self-identified pain patient	Healthy voluteers
Pain type	Natural sholder pain	Stimulated heat pain
Pain levels	0-16 (PSPI) and 0-10 (VAS)	1-4 (Stimuli)
Modalities	RGB	RGB
Size of the database	200 variable length videos with 31,571 frames	17,300 5s videos with 25 fps

Comparison of the Publicly Available Video Databases on PAIN

Attribute	UNBC-McMaster database (2011)	BioVid database (2013)	MIntPain database (2018)
No. of subjects	129 (16 are available)	90 (87 are available)	20
Subject's type	Self-identified pain patient	Healthy voluteers	Healthy volunteers
Pain type	Natural sholder pain	Stimulated heat pain	Stimulated electrical pain
Pain levels	0-16 (PSPI) and 0-10 (VAS)	1-4 (Stimuli)	0-4 (Stimuli)
Modalities	RGB	RGB	RGB, Depth, Thermal
Size of the database	200 variable length videos with 31,571 frames	17,300 5s videos with 25 fps	9366 variable length videos with 1,87,939 frames

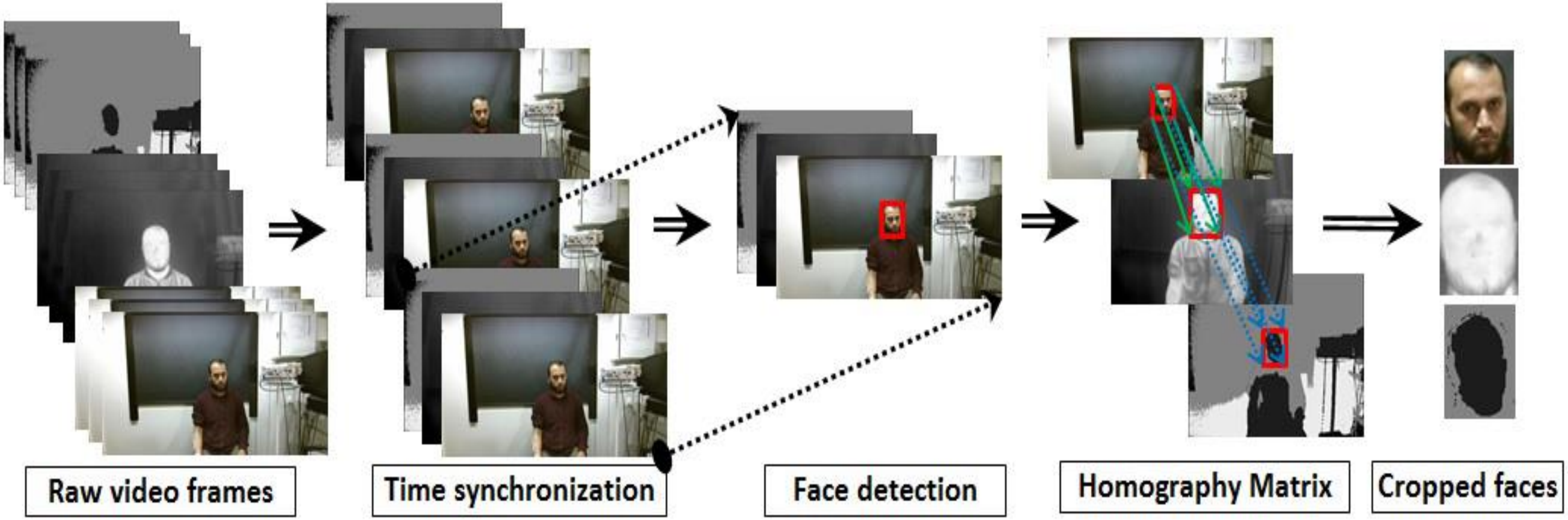
Raw Frames from the Database



Preprocessing of the Database

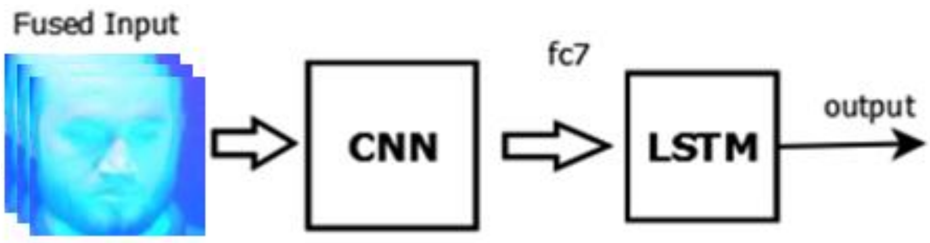
Raw Frame to Cropped Face

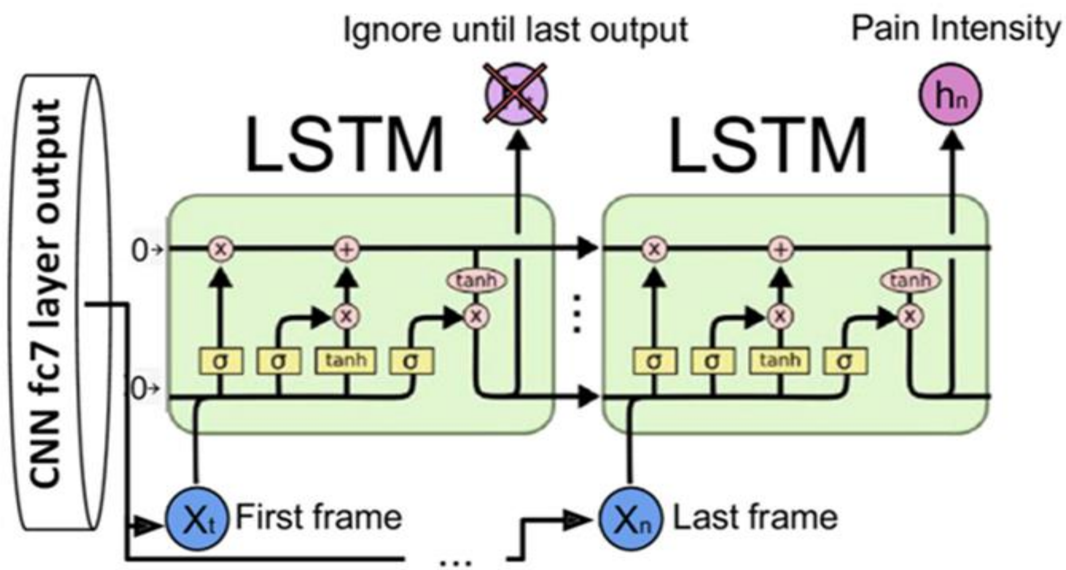
- Frame by frame synchronization using timestapms
- Markus's Face Detection
- 8-point homography



Exploiting Spatio-Temporal Information

The CNN+LSTM Architecture





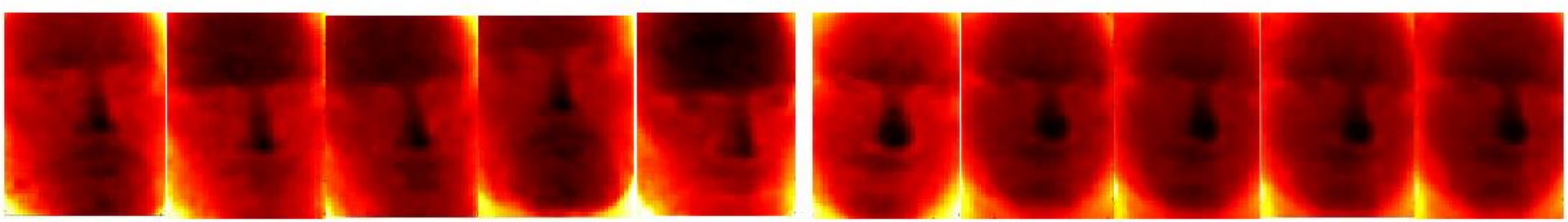
Architecture of the hybrid CNN+LSTM deep learning framework

Exploiting Multimodality

Faces with 5-Pain Levels

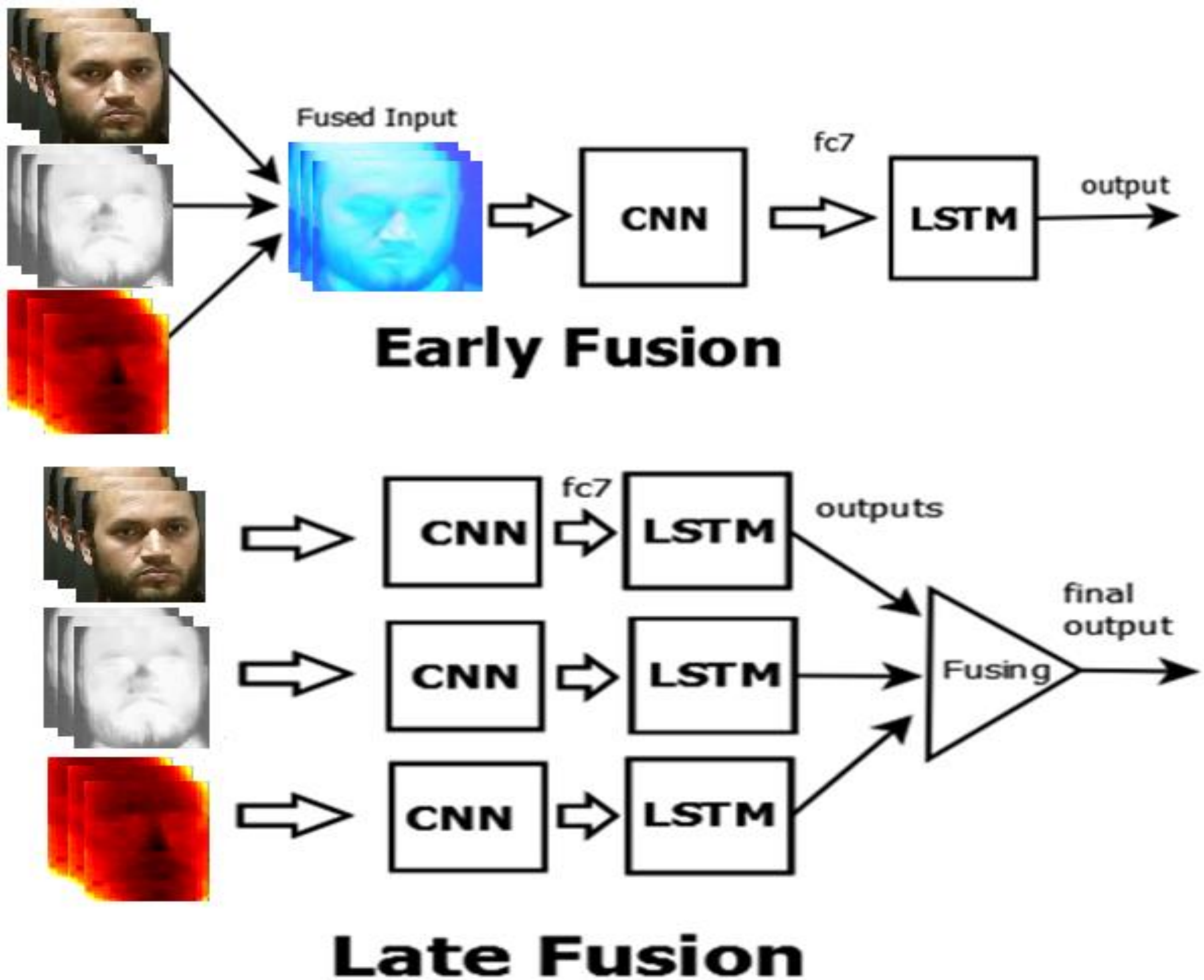
(a) RGB faces

(b) Thermal faces



(c) Depth faces

Early and Late Fusions



The Baseline!

Independent Modalities with VGG-Face CNN and LSTM

Modalities	CNN-RGB	CNN-T	CNN-D
Mean Frame(%)	**18.17**	18.08	16.71
Mean Sequence (%)	**18.55**	18.33	17.41

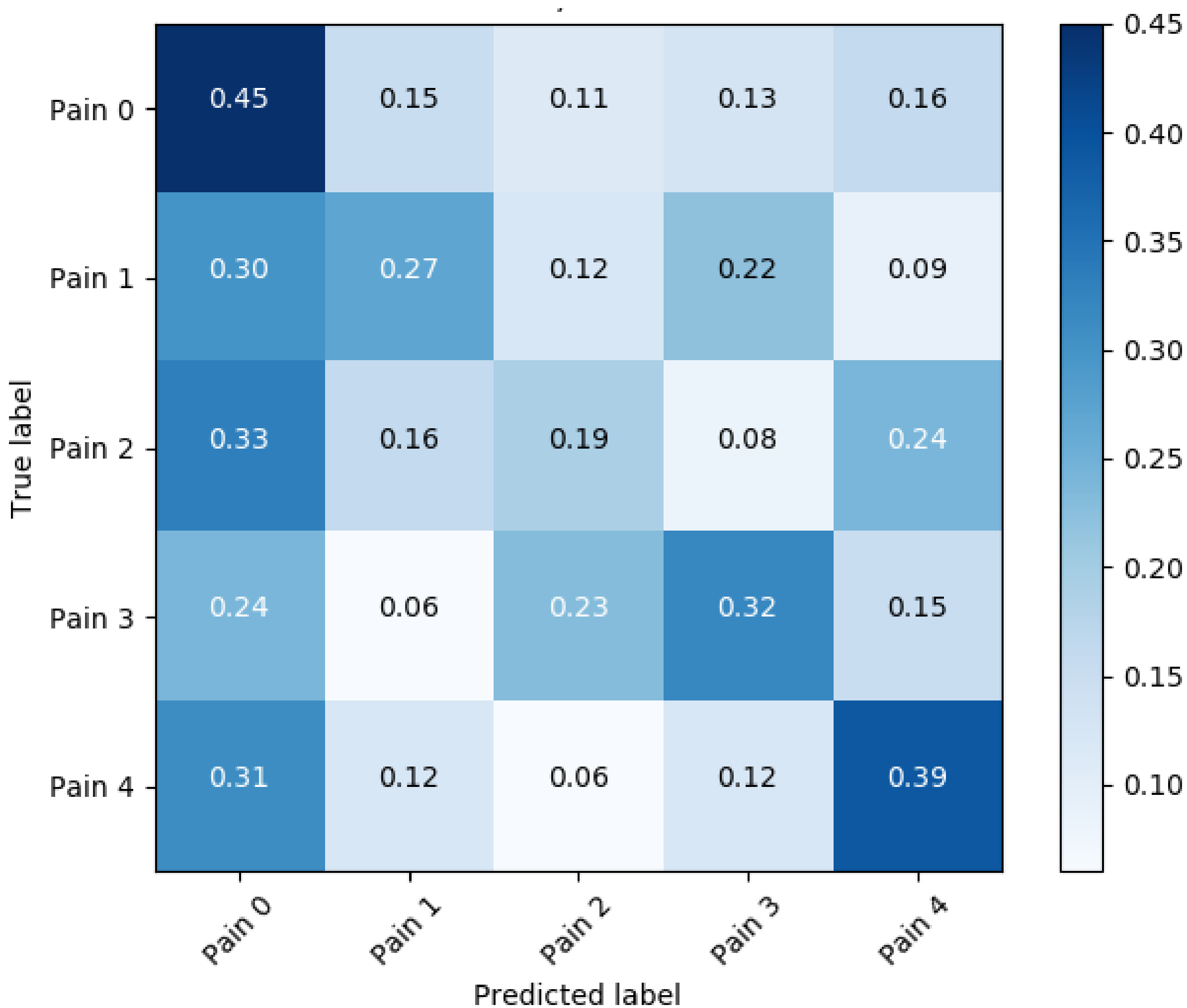
LSTM-RGB	LSTM-D	LSTM-T
15.36	14.72	13.13
15.36	14.72	13.13

Early and Late Fusion Results

Fusion	EF-RGB-T	EF-RGB-D	EF-D-T	EF-RGB-DT
Mean Frame (%)	23.85	24.62	23.12	**32.40**
Mean Sequence(%)	30.77	27.92	25.30	**36.55**

LF-RGB-T	LF-RGB-D	LF-D-T	LF RGB-D-T
21.80	23.20	22.50	25.20
22.10	22.30	22.70	25.40

Confusion Matrix for EF of All Modalities



MISC

- The database is available here:
  - http://www.vap.aau.dk/mintpain-database/

- An implementation of CNN+LSTM available here:
  - https://github.com/prlz77/LSTM-on-CNN


Q
&
A