Deep Spatio-Temporal Video Based Analysis for Shoulder Pain Intensity Measurement

Maimoon Siddiqui[1], Zakia Hammal[1], Diyala Erekat[2], Hamdi Dibeklioğlu[2]
[1]School of Computer Science, The Robotics Institute Carnegie Mellon University, [2]Department of Computer Engineering, Bilkent University

Motivation

- Pain assessment and management are important across a wide range of disorders and treatment interventions[1].
- The standard clinical assessment of pain is limited primarily to the subjective reports (e.g., Visual Analog Scale (VAS)).
- While convenient and useful, subjective reports have several limitations (e.g., inconsistent metrics, reactivity to suggestion).
- ***We propose an automatic and objective pain intensity measurement using spatio-temporal changes in facial expression.***

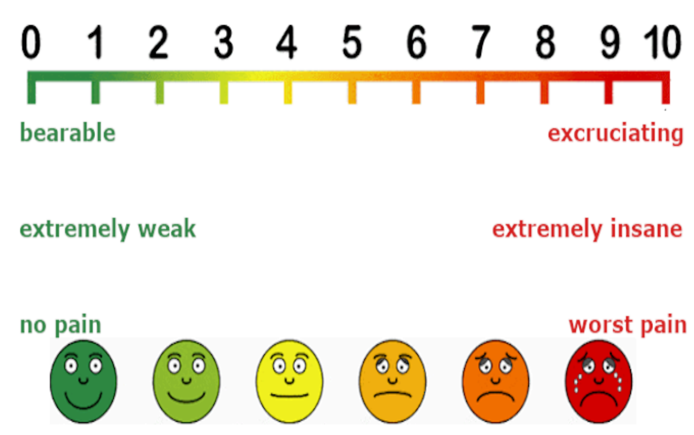
Dataset

UNBC McMaster Pain Archive[2]

- 25 Participants with shoulder pain
- 200 video sequences

For each video sequence:

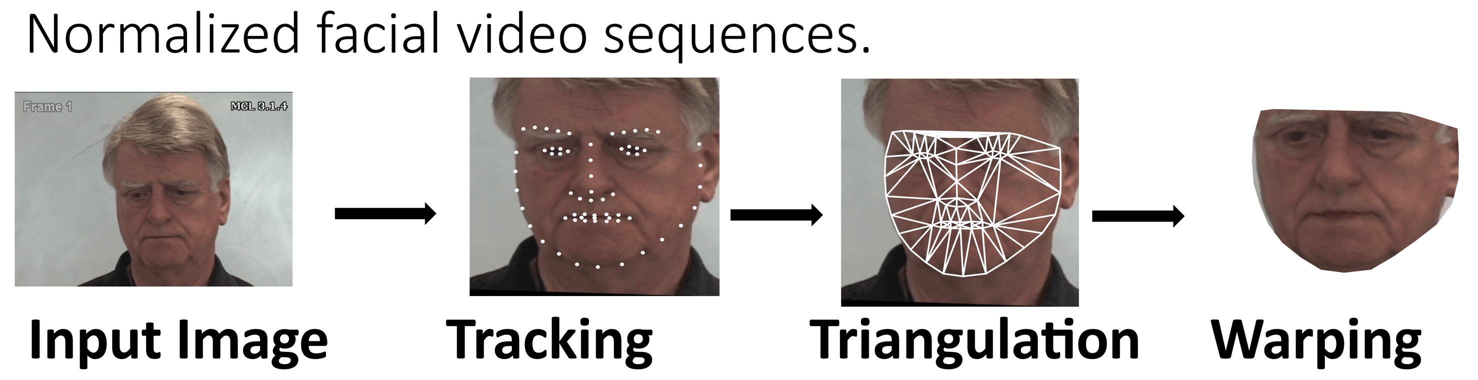
- Three self-reported pain scores
  - ✓ Affective Scale (AFF)
  - ✓ Sensory Scale (SEN)
  - ✓ Visual Analogue Scale (VAS)



- Offline Observer Pain Intensity Rating (OPI)

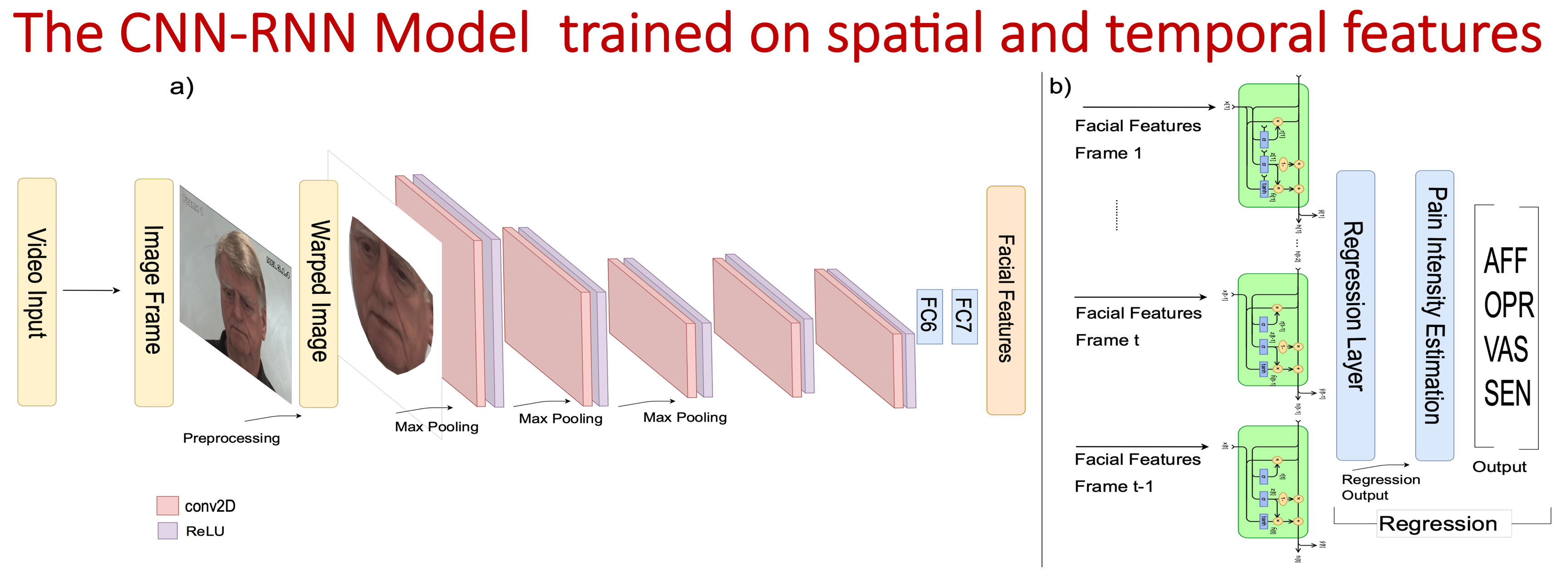
Face Registration And Warping

- 66 facial points tracked using Active Appearance Model[2].
- Face registration and warping using Delaney triangulation.
- Normalized facial video sequences.



End-to-End Spatio-Temporal Deep Model

The CNN-RNN Model trained on spatial and temporal features



CNN: AlexNet[3] trained to learn frame-by-frame spatial feature (4096D per frame).
RNN: 2-layer GRU[4] trained to learn per-video temporal dynamics of facial features.

Loss Function:

MSE   Inter-variance Loss.   L2 Regularisation

$$\frac{\alpha}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+(1-\alpha)Var(\hat{Y})+\lambda\sum_{p=1}^{P}(\beta_p)$$

Inter-variance loss penalizes the difference between self-reported and observer's reported pain scores.

Experimental Setup

Stratified distribution of data

- Dataset is split into five independent folds
- Well distributed pain intensities per fold.
- Ensures model training is not biased due to skewed training distribution.

Fold #	AFF	VAS	OPR	SEN
Fold 1	97.47	97.6	93.86	97.95
Fold 2	98.93	99.42	99.45	98.41
Fold 3	97.89	98.28	98.33	97.86
Fold 4	98.39	98.28	99.25	97.96
Fold 5	93.92	96.79	98.37	96.23

[a]Mean square similarity with the entire Pain Archive

Data distribution across 5 folds: Mean Square similarity to the entire dataset

Two level 5 fold Cross Validation

```
params = [num layers, dropouts, learning_rate, λ]
for outer_fold (5) do
    min_val_loss = infinity; test_data = data[outer_fold];
    train_val_data = data - test_data;
    for inner_folds do
        val_data = data[inner_fold];
        train_data = train_val_data - val_data;
        net.init(params[inner_fold]);
        net.fit(train_data);
        val_loss = net.history['valid_loss'];
        if val_loss < min_val_loss then
            net.save_params(params[inner_fold]);
            min_val_loss = val_loss;
            min_param_idx = params[inner_fold];
        end
        test_net = net.init(params[min_param_idx]);
        test_net.fit(train_val_dataset);
        test_loss = evaluate(test_net,test_data);
    end
end
```

Algorithm 1. Nested two level cross validation

Experimental Results

- Mean Absolute Error (MAE) in pain intensity measurement: VAS [0-10], OPI [0-5]

			Random Distribution		Stratified Distribution	
Method	Setting	Pain Scale	MAE (VAS)	MAE (OPI)	MAE (VAS)	MAE (OPI)
Our Model	One Label	VAS	2.98	-	2.57	-
	One Label	OPI	-	1.46	-	1.33
	Two Labels	VAS,OPI			2.25	1.72
	Four Labels	VAS,OPI,AFF,SEN	3.09		2.54	1.39
DeepFaceLift Liu et al	3rd Neural Net Layer	VAS, OPI	2.34	-	-	-
		VAS	2.24	-	-	-
	Neural Net Layer	VAS, OPI	2.41	-	-	-
		VAS	2.22	-	-	-

- Our Model: spatio-temporal modelling of temporal changes (OPI and VAS).
- DeepFaceLift[5]: summary statistics as proxy of temporal changes (VAS only).

Conclusions

- Automatic, objective, and reliable measurement of pain intensity from facial expression is feasible.
- The proposed loss function exploits the consistency between different pain intensity measures.
- Stratifying data on average improved VAS and OPI results by 13.6% and 8.9%, respectively .
- OPI offers a more objective assessment of pain intensity.

References

[1] Z. Hammal, J. F. Cohn, "Automatic, Objective, and Efficient Measurement of Pain Using Automated Face Analysis, Social and Interpersonal Dynamics in Pain", pp. 121–146, 2018.

[2] P. Lucey, J. Cohn, K. Prkachin, P. Solomon, I. Matthews. "Painful data: The UNBC-McMaster shoulder pain expression archive database", IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, FG 2011. pp. 57-64, 2011.

[3] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks", Proceedings of NIPS, IEEE, Neural Information Processing System Foundation, pp. 1097-1105, 2012.

[4] J. Chung, Ç. Gülçehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling", NIPS 2014 Workshop on Deep Learning, December 2014.

[5] D. L. Martinez, O. Rudovic, R. Picard, "Personalized Automatic Estimation of Self-Reported Pain Intensity from Facial Expressions," IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017.

Acknowledgement

Thank you Dr Zakia Hammal, Dr Hamdi Dibeklioglu , Diyala , Dr John Dolan, Rachel Burcin and all the members of RISS who made this project possible.
Research reported in this publication was supported in part by the National Institute of Nursing Research of the National Institutes of Health under Award Number R21NR016510. The content is solely the responsibility of the authors and does not necessarily represent the official views of the sponsors.

Carnegie Mellon University