Prediction of pain duration and intensity from patellofemoral pain maps using deep learning

Birgithe Kleemann Rasmussen, Ignas Kupcikevičius,
Linette Helena Poulsen, Mads Kristensen

Aalborg University

December 20, 2017

Abstract

Introduction: *Patellofemoral pain (PFP) syndrome is a musculoskeletal condition that presents as pain behind or around the patella without known structural changes [1]. Partial correlations between perceived size of PFP from pain maps and pain duration along with intensity has been indicated in previous studies [2], however morphology and location of PFP remains unexplored in terms of correlation. Based on the objects detection capabilities of deep learning, convolution methods can be used to detect image-features related to morphology. The aim of this study was to determine the performance of deep learning classification according to pain duration and intensity, based on morphology and location of perceived PFP from pain maps.*
Methods and materials: *PFP drawings were collected on lower extremities body-schema and encoded into three different data representations in respect to morphology of pain and location and a combination of the two. The distribution of the outputs were analyzed and used for defining the classification intervals for pain duration, below 12 months and above 36 months, and pain intensity, below 4 and above 8 on VAS. Estimation of generalization performance of the models was calculated through 10-fold cross validation during the training.*
Results: *The results during training showed a higher accuracy for pain intensity classification than pain duration classification using morphology-representation. Pain intensity had an accuracy on 65.04% (SD: 10.83%) and pain duration had an accuracy on 59.51% (SD: 11.20%). Furthermore, the combined-representation performed with the highest accuracy on 65.14% (SD: 12.87%). The location and morphology-representation scored 63.33% (SD: 1.67) and 65.04% (SD: 10.83%), respectively, based on pain intensity.*
Discussion: *Despite pain intensity being defined as multidimensional and subjective, the performance accuracy were higher than that of pain duration. The results may indicate that a combination of the morphology and the location of the pain had a higher classification performance in relation to pain duration or intensity. Currently, it is unclear if deep learning methods may be a suitable approach for classifying PFPS to work as support in clinical settings, to which further investigation is necessary. Improvements could be found when more data become available to better reflect generalization patterns in PFP drawings.*

I. Introduction

Patellofemoral pain syndrome (PFPS) is a painful musculoskeletal condition that is presented as pain behind or around the patella [1, 2]. PFPS affects 6-7% of adolescents, of whom two thirds are highly physically active [3]. Additionally the prevalence is more than twice as high for females than males [3, 4]. PFPS may be present over a longer period of time where a high number of individuals experience a recurrent or chronic pain [5]. Chronic pain may be maintained by the phenomenon central sensitization, which may result widespread pain over longer periods of time. Furthermore, PFPS may lead to osteoarthritis [4, 6]. Patellofemoral pain (PFP) is often described as diffuse knee pain, that can be hard for individuals to explain and localize [5]. Despite the fact that individuals feel pain in the knee, there is no structural changes in the knee such as significant chondral damage. Because PFP is not caused by structural changes, no definitive clinical test may be used to diagnose PFPS and thereby often diagnosed based on exclusion criterias [4] to which PFPS is also described as an orthopaedic enigma, and is one of the most challenging pathologies to manage [7]. To assist diagnosis of PFPS, pain maps may be used as a helpful tool for the individuals to communicate their pain by drawing pain areas on a body outline [8].

A study by Boudreau et al. [9] indicates, through the

use of pain maps, that there is a correlation between the size of the pain (number of pain pixels) and the pain duration as well as intensity for individuals with PFP longer than five years.[9] However, it is unknown whether pain duration has an influence on morphology of the pain and location, as well as whether morphology of pain and location have an influence on pain intensity. The relation between pain maps and pain duration or pain intensity may be complex, because the perceived PFP is subjective, and considered as multifactorial [**?**]. Additionally the study by Boudreau et al. did not find a fully correlation between 35 pain maps and pain duration or pain intensity for individuals with a pain duration below 5 years. To investigate the potential nonlinear correlation, a deep learning method was used, which is a method that previously has not been applied on this type of data.
The goals of this study is to explore how accurate a deep learning model can classify pain maps according to pain duration or pain intensity. It is assumed that pain duration is a better predictor than pain intensity, because the perceived pain is subjective, and may be affected by multidimensional factors. The pain maps are encoded into multiple data representations to investigate whether morphology and location are correlated to pain duration or intensity.
The data representations are encoded into three representations, which reflect either morphology of pain or location. It is assumed that a deep learning model will perform better with more information, thus a combination of morphology and location of the pain constitute a data representation. The data representations are refereed to as morphology-, location- and combined-representation. There may be a difference in how gender reports pain intensity, where females reports more intense and frequent pain. Furthermore, there is an imbalance in prevalence between females and males, thus gender is included as a feature in the deep learning model.
The aim of this study was to explore classification performance of a deep learning model, using PFP maps and gender as input to classify according either pain duration or intensity. Furthermore, a secondary aim was to compare the performance accuracy of deep learning models with different pain map representations (morphology-, location- and combined-representation), when predicting pain duration or pain intensity.

II. Methods

Pain maps

Data used in this study were collected from an ongoing clinical trial (FOXH) which is conducted in collaboration with Danish and Australian universities. The pain maps were drawn by individuals with PFPS through the use of an application, Navigate Pain, in a clinical setting.
Navigate Pain is a software application that is used to visualise the location, morphology and spatial distribution of pain from individuals to healthcare personnel. The application permits individuals to draw their pain with different colors and line thickness onto a body outline, an example is shown in fig. 1. Navigate Pain android was developed at Aalborg University.[10]

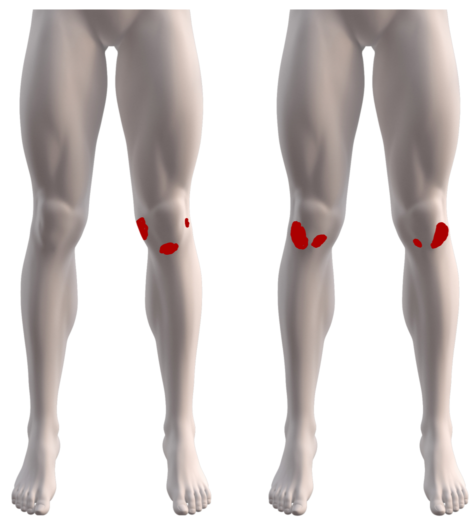

Fig. 1: *Pain maps from individuals with uni- and bilateral PFP. The red markings indicate the area of pain perceived by the individuals.*

The total number of pain maps available was 217, but only 205 pain maps with associated gender and pain duration, and 197 pain maps with associated gender and pain intensity was available.

Preprocessing

The pain maps were processed in MatLab version R2017b, where the images were resized, since they were collected at different resolutions (screen sizes) and cropped to only include the knees. To create more pain maps is a split body approach used, where pain maps are divided into two knees. Furthermore, pain was mirrored to represent pain on right knees to minimize the variance in the images. By using split body approach it was assumable that the pain duration and pain intensity were identical for both knees if PFP was bilateral. The total number of pain maps with gender and pain duration was 333, and pain maps with gender and pain intensity was 319, of which 15% was used as test data, and therefore not used to optimize and train the models.

Morphology-representation

The original pain maps reflect the morphology of the pain, and do not require further processing than converting the pain maps to a matrix including gender and the output, pain duration or pain intensity.

Pain location

The knees are divided into regions based on the underlying anatomical structures, which may have a correlation to pain duration or pain intensity. The locations are divided into 20 regions, which are inspired by Photographic Knee Pain Map (PKPM). The divisions are designed to categorise location of knee pain for diagnostic and research purposes. PKPM represent both knees that makes it possible to identify unilateral and bilateral pain.[11] The knee regions are illustrated in fig. 2.

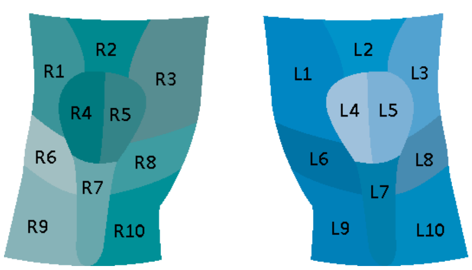

Fig. 2: *The regions of the left (L1-L10) and right (R1-R10) knees, where each knee is split into ten regions.*

There are ten regions on each knee, where region 1 and 3 represent the superior lateral and superior medial areas for patella. Region 2 refers to quadriceps tendon. The patella is divided into lateral and medial regions, which are region 4 and 5. Region 6 and 8 are lateral and medial joint line areas. Patella tendor is region 7 and the two last regions, 9 and 10, are tibia lateral and medial.[11]

Location-representation

To investigate whether the location alone have a correlation to the outputs, a simplified representation of the pain maps was created. The location of the pain was reflected by the use of the defined knee regions (fig. 2), where each region represented a value of 0 (not active) or 1 (active) in a vector. The values were defined by using a threshold to determine whether a region was considered active in relation the amount of pain. A threshold was required to increase the confidence of an active pain region by avoiding minimal contributions e.g. small pain areas in the associated regions. Simultaneously the threshold should not be too large so that pain areas was excluded. The threshold was decided based on an analysis on five random pain maps, where threshold values of 0, 5, 10 and 15% was compared. The threshold represent which minimal percentage of pain should be present in a specific region before it is considered active. Based on the analysis a 5% threshold was chosen. As a result for adding the threshold value the number of pain maps with pain duration decrease to 316, and number of pain maps with pain intensity decrease to 302.

Combined-representation

A combination of morphology and location of the pain is created based on components from morphology- and location-representations. The original pain maps are superimposed on the regions, which result in pain pixels reflecting the location with a number from 1 to 10. Before using the representation as input, one-hot encoding approach was used, which made it possible to separate categorical data into binary data [46]. This means that the 20 values for each knee region do not have a correlation when analysed in the deep learning model. The number of pain maps with pain duration was 331, and number of pain maps with pain intensity was 317. The number of pain maps increased according to the location-representation, because no threshold was applied in this data-representation.

Nonlinearity in pain maps

Given that PFP is subjective and multifactorial it is unlikely that the pain maps and pain duration or pain intensity are linearly correlated. In order to determine if there was a linear relationship, linear regressions were done on simple features reflecting the size of the pain and number of active pain regions. The linear regressions were made in MatLab, and composed a correlation between number of pain pixels and pain duration, number of pain pixels and pain intensity, number of active pain regions and pain duration, and number of active pain regions and pain intensity.

Deep learning models

Deep learning models were developed on a computer with 4x "Intel® Core^{TM} i7" CPUs and one single GPU of type "Geforce GTX 970M", using the programming language Python v3.6.3. Libraries used was Keras with a TensorFlow backend.

Multiple deep learning models suitable to the three data representation were created. The models used supervised learning, which is defined as a network learning to classify a given input corresponding to a specific output [12]. The models classify the input, pain maps and gender, in relation to the determined outputs, pain duration or pain intensity.

Two of the models, which managed the morphology-, and the combined morphology and location-representations, were developed using the same model architecture consisting of convolutional-followed by pooling layers and fully connected layers. Convolutional were used because it's highly classification in images that automatically learn a complex pattern by extracting visual features from the pixel-level content [13, 14]. The combination of convolutional and pooling layers performed feature extraction while the classification was made by fully connected layers.

The model that classified the location should not process morphology, thus a convolutional layer was not necessary, and thereby only contained fully connected layers.

Optimization of models

Optimization of the three model were done using a validation subset, whereto graphs plotting validation accuracy and loss were compared with training accuracy and loss. This was used to estimate the optimal number of epochs to reduce overfitting the model to the training subset. Further optimization were done using manual search on hyperparameters, from which improvements were based on an average accuracy, sensitivity and specificity gained from 10-fold cross validation.

III. Results

Linear correlations

The linear regression between simple features, number of pain pixels or active pain regions, and outputs, pain duration or pain intensity, resulted in the plots shown in fig. 3. The R^2-values support the nonlinearity, shown in the plots, where correlation fig. 3a resulted in a $R^2 = 0.018$, fig. 3b resulted in $R^2 = 0.008$, fig. 3c resulted in $R^2 = 0.011$ and fig. 3d resulted in $R^2 = 0.011$.

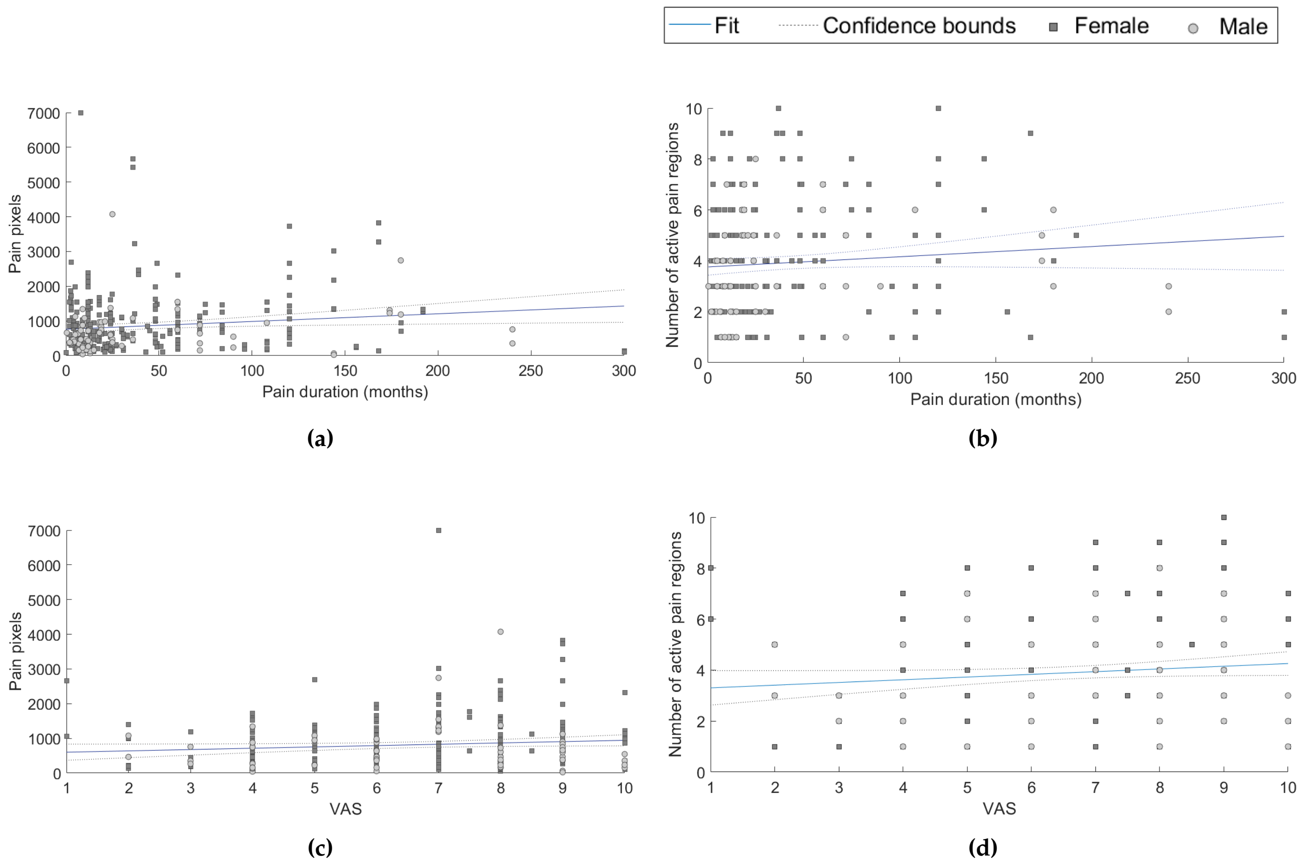

Fig. 3: *Linear correlations of pain pixels and pain duration (a), active pain regions and pain duration (b), pain pixels and pain intensity indicated in VAS (c), and active pain regions and pain intensity indicated in VAS (d).*

Classification according to outputs
tyfuihoj

Classification of pain maps
yubi

Results on generalization performance when using 10-fold cross validation, after final optimization for morphology-, region- and combined representation are shown in tab. 1. This is done for pain duration and pain intensity classifications to which an average accuracy, sensitivity, and specificity are calculated along with it's corresponding standard deviations.

IV. Discussion

Bullet points:

- results, compare them to regression
- accuracy, sensitivity and specificity
- Pain duration vs pain intensity
- Split body approach
- data amount (not large enough)

number of pain maps and split body and Linear correlation, threshold –> splitbody assumption according duration, intensity, independent) –> threshold, losing pain maps Classification according to output Classification of pain maps optimization of the model –> gridsearch, computerpower, input parameters (subjects activity level, age, other pain (hip pain ect.))

V. Conclusion

conclusionnnnn

References

[1] Liam R. Maclachlan, Natalie J. Collins, and Et.al. The psychological features of patellofemoral pain: a systematic review. 2017. doi: 10.1136-bjsports-2016-096705.

[2] T.O. Smith, B.T. Drew, and Et.al. Knee orthoses for treating patellofemoral pain syndrome (review). 2015. doi: 10.1002/14651858.CD010513.pub2.

[3] M. S. Rathleff, B. Vicenzino, and Et.al. Patellofemoral Pain in Adolescents and adulthood: same same, but different? 2015. doi: 10.1007/s40279-015-0364-1.

[4] Wolf Petersen, Andree Ellermann, and Et.al. Patellofemoral Pain Syndrome. *Clinical Orthopaedics and Related Research*, 2013. doi: 10.1097/01.blo.0000229284.45485.6c.

[5] Erik Witvrouw, Michael J. Callaghan, and Et.at. Patellofemoral Pain: consensus statement from the 3rd International Patellofemoral Pain Research Retreat held in Vancouver, September 2013. 2014. doi: 10.1136/bjsports-2014-093450.

[6] Kay M. Crossley, Michael J. Callaghan, and Et.al. Patellofemoral pain. 2016. doi: 10.1136/bjsports-2015-h3939rep.

[7] Scott F Dye. Patellofemoral Pain Current Concepts: An Overview. *Sports Medicine and Arthroscopy Review*, 2001.

[8] Shellie A. Boudreau and Susanne et.al Badsberg. Digital pain drawings: Assessing Touch-Screen Technology and 3D Body Schemas. 2016. doi: 10.1097/AJP.0000000000000230.

[9] Shellie A. Boudreau, E. N. Kamavuako, and Et.al. Distribution and symmetrical patellofemoral pain patterns as revealed by high-resolution 3D body mapping: a cross-sectional study. 2017. doi: 10.1186/s12891-017-1521-5.

[10] Aglance Solutions. Visual insight for clinical

	Avg. accuracy (%)	Avg. sensitivity (%)	Avg. specificity (%)
		Morphology-representation	
Pain duration	59.51% (±11.20%)	62.16% (±21.59%)	62.20% (±18.13%)
Pain intensity	65.04% (±10.83%)	48.23% (±0.28%)	71.02% (±0.12%)
		Location-representation	
Pain duration	54.56% (±12.81%)	50.68% (±0.15%)	59.55% (±0.15%)
Pain intensity	63.33% (±1.67%)	0.00% (±0.00%)	63.33% (±0.02%)
		Combined-representation	
Pain duration	55.49% (±9.55%)	55.23% (±0.15%)	56.99% (±0.12%)
Pain intensity	65.14% (±12.87%)	37.50% (±0.35%)	67.34% (±0.15%)

Table 1: *Generalization performance of the three models, which use the morphology-, location- and combined-representation.*

reasoning – Navigate Pain, 2015. URL http://www.navigatepain.com/.

[11] D. W. Elson, S. Jones, and Et.al. The photographic knee pain map: Locating knee pain with an instrument developed for diagnostic, communcation and research purposes. 2010. doi: 10.1016/j.knee.2010.08.012.

[12] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. URL http://www.deeplearningbook.org.

[13] Jacopo Acquarelli, Twan van Laarhoven, Jan Gerretzen, Thanh N. Tran, Lutgarde M.C. Buydens, and Elena Marchiori. Convolutional neural networks for vibrational spectroscopic data analysis. *Analytica Chimica Acta*, 954:22–31, 2017. ISSN 18734324. doi: 10.1016/j.aca.2016.12.010. URL http://linkinghub.elsevier.com/retrieve/pii/S0003267016314842.

[14] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2323, 1998. ISSN 00189219. doi: 10.1109/5.726791. URL http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf.