Master's Thesis Proposal Nr.

Topic: Classification of hand motion patterns in inflammatory arthritis patients using machine learning

Inflammatory arthritis describes a group of autoimmune diseases characterized by inflammation of the articular structures. [1] There are several distinct kinds of arthritis, including psoriatic arthritis (PsA) and rheumatoid arthritis (RA). Especially at the disease onset the small joints of the hands and fingers show a high involvement for both types. However, RA often shows a common trait of symmetric involvement, whereas in PsA it is more likely to appear asymmetrical[2].In addition, in RA patients the disease manifests in more than 90 percent with synovitis, and the metacarpophalangeal joints are commonly affected, whereas PsA patients tend to develop tenosynovitis or dactylitis and the distal interphalangeal joints are affected [3]. It is thus likely that the different clinical manifestations will lead to observable differences in hand function impairment.

A good and timely medical treatment is necessary to keep the patients in clinical remission with no active inflammation to prevent destructive joint damage. Frequent testing would therefore be desirable. To evaluate specific damage in arthritis diagnosis, doctors commonly use blood tests for inflammation and imaging techniques, which can, however, be costly and time-consuming. Given that RA and PsA primarily affect the motion range of the patients, the analysis of movement restrictions in inflammatory arthritis could offer an additional, objective approach to finding functional biomarkers for the disease.

Marker-based motion capturing is the current gold standard for movement analysis. This technology provides the opportunity to objectively analyze movement performance. In a current study of the subproject D01 and C04 (GAPX) of the CRC EmpkinS this technology was used to assess hand movements during a wide range of different elementary and functional tasks from patients with RA and PsA as well as a healthy non-arthritic control group with the overall aim of identifying disease-specific functional biomarkers of physical impairment.

An issue with motion capture data is that it is typically complex and high-dimensional. Machine learning (ML) can help to identify patterns and underlying relationships in the data that cannot be captured with traditional statistical tools. Thus, applying ML methods to the recorded hand motion capture data of RA and PsA patients and controls could offer a more objective approach to identifying group characteristics and functional biomarkers. In a previous pilot study, we first investigated the feasibility of using ML to distinguish RA patients and controls from trajectories of different hand motion tasks using only automatically extracted features. However, the classification performance was below the anticipated level of accuracy (>70%), which could potentially be explained by the small dataset (24 patients, 23 controls) with very limited data quality.

This thesis aims to expand the previous work using the larger dataset from GAPX. The repetition of the former classification experiment serves as a crucial starting point to determine if the poor results were due to data quality issues or methodological shortcomings. However, this thesis goes beyond the prior work in multiple aspects. First, we take into account two subgroups of inflammatory arthritis with different clinical manifestations (RA and PsA). The goal is to investigate if a machine-learning model based on hand motion data can not only distinguish between patients and controls but also between RA and PsA patients. If successful, this would support our assertion that different disease types result in different motion patterns, which is a big step toward identifying disease-specific functional biomarkers. Second, for the first time, an approach will be applied to investigate whether we can predict the stage of the disease as defined by clinical disease activity scores.

The classification tasks will be achieved by training different machine learning pipelines consisting of a standard workflow that includes data preprocessing, feature extraction, feature selection, classification, and evaluation. However, other than before, we will not limit the feature space to automatic features, but extract and incorporate additional expert features (e.g. range of motion, peak distances,..) using input from clinical experts. From the classification results, we aim to identify the movement best suited to distinguish spe-

cific groups. We will further determine which joints or features are most discriminative by looking at the feature importance. The thesis will help us to identify disease-characteristic patterns of hand movement impairment and gain information on biomarkers of hand function that can be used for fine-tuned monitoring of the two different diseases, RA and PsA.

The proposed work consists of the following parts:

- Literature review focusing on clinical manifestations and current monitoring and treatment options in different types of inflammatory arthritis, as well as, the use of machine learning, with a focus on (hand) movement analysis and classification of inflammatory arthritis.
- Validation of the existing classification pipeline based on automatic feature extraction for a broader dataset with clearly defined events for different subcomponents during a movement task.
- Extraction and identification of informative expert features such as the range of motion of specific joints using clinical findings and expert knowledge.
- Classification of RA and PsA patients and controls.
- Definition of sub-groups within the patients based on common clinical disease activity scores to classify different stages of the disease.
- Evaluation of the classification performance based on common metrics (accuracy, F1 score..).
- Identification of the movement task that is most suited for the classification of groups especially between RA and PsA patients and disease stages.
- Identification of the most discriminative features and joints for each patient group.

The thesis must contain a detailed description of all developed and used algorithms as well as a profound result evaluation and discussion. The implemented code has to be documented and provided. Extended research on literature, existing patents, and related work in the corresponding areas has to be performed.

Advisors:	Sophie Fleischmann, Robert Richer, Prof. Dr. Bjoern Eskofier (Machine Learning and Data Analytics Lab, FAU Erlangen-Nürnberg) Birte Coppers, PD Dr. Anna-Maria Liphardt (Department of Internal Medicine 3 - Rheumatology and Immunology, University Hospital Erlangen) Simon Heinrich, Prof. Dr. Sigrid Leyendecker (Lehrstuhl für technische Dynamik, FAU Erlangen-Nürnberg
Student:	Saloni Jethwa
Start – End:	15.02.2024 - 15.08.2024

Literatur

[1] Folle L, Simon D, Tascilar K, Kronke G, Liphardt AM, Maier A, Schett G, Kleyer A. *"Deep Learning-Based Classification of Inflammatory Arthritis by Identification of Joint Shape Patterns-How Neural Networks Can Tell Us Where to "Deep Dive"Clinically."* Front Med (Lausanne). 2022 Mar 10;9:850552. doi: 10.3389/fmed.2022.850552 PMID: 35360728 ; PMCID: PMC8960274.

[2] Connolly J, Condell J, Curran K, Gardiner P. *Ä new method to determine joint range of movement and stiffness in rheumatoid arthritic patients."*Annu Int Conf IEEE Eng Med Biol Soc. 2012;2012:6386-9. doi: 10.1109/EMBC.2012.6347455 . PMID: 23367390 .

[3] Ogdie A, Coates LC, Mease P *"Measuring Outcomes in Psoriatic Arthritis."*Arthritis Care Res (Hoboken). 2020 Oct;72 Suppl 10(Suppl 10):82-109. doi: 10.1002/acr.24242 . PMID: 33091263 ; PMCID: PMC8528225. 23367390.

[4] Smolen JS, Schoels M, Aletaha D. *"Disease activity and response assessment in psoriatic arthritis using the Disease Activity index for PSoriatic Arthritis (DAPSA). A brief review. Clin Exp Rheumatol."*2015 Sep-Oct;33(5 Suppl 93):S48-50. Epub 2015 Oct 15. PMID: 26471734 .

[5] Adamczyk, J., and Malawski, F. *"Comparison of Manual and Automated Feature Engineering for Daily Activity Classification in Mental Disorder Diagnosis."* COMPUTING AND INFORMATICS, 40(4), 850879 .here

[6] Christ, M., Braun, N., Neuffer, J., and Kempa-Liehr, A. W. *"Time Series Feature Extraction on basis of Scalable Hypothesis tests (tsfresh A Python package)"* Neurocomputing, 307, 72-77. .