Towards Predicting Fine Finger Motions from Ultrasound Images via Kinematic Representation

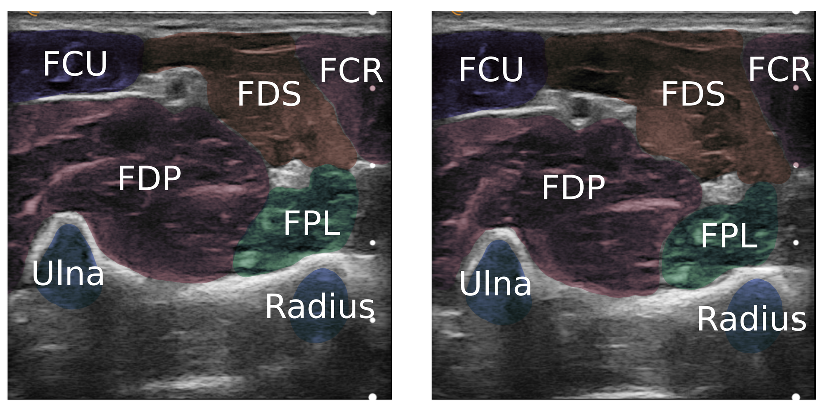
Paper
Code
Data

TECHNION Israel Institute of Technology

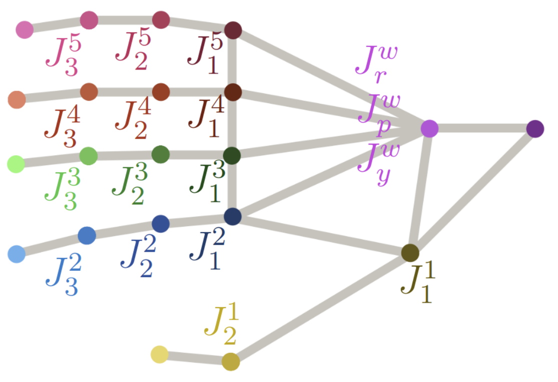
Dean Zadok, Oren Salzman, Alon Wolf and Alex Bronstein

- Robotic prostheses are mostly tailored to perform grasping tasks.
- Lack of broader functionality such as keyboard typing results in high rejection rates of available prostheses.
- We replace electromyography with an ultrasound as the input sensor to our system.
- This changes the input signal from electric potentials to lower-arm muscle images which allows to exploit SOTA learning techniques.
- We propose a model-based method to detect fine finger motion given a sequence of ultrasound images and predict pressed keys for piano playing and keyboard typing, by using robotic hand configurations as intermediate representations.

Anatomical Background

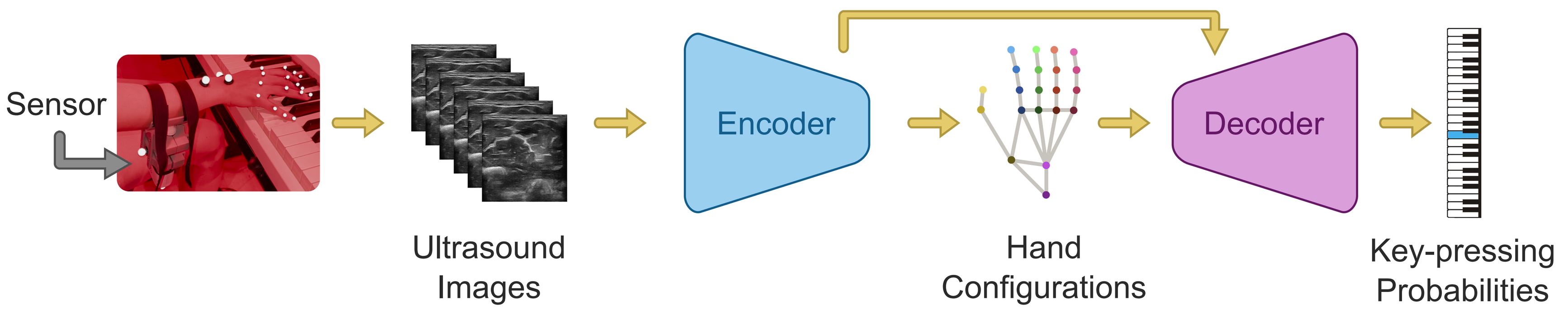


Visual segmentation of two ultrasound images. All fingers are at rest in the left image while in the right image the Ring finger is fully flexed.



Visual illustration of a configuration of the hand. Two joints for the thumb, and three joints for each finger.

Learning Framework

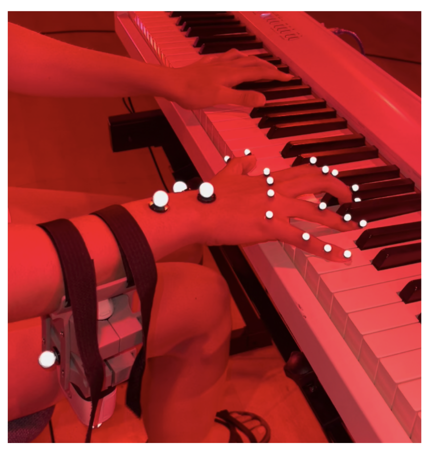


A schematic flow of our Configuration-based Multi Frame (CBMF) model. The sensor is placed on the lower arm while the subject is playing the piano. A continuous stream of ultrasound images is fed into the neural-network encoder which creates a latent representation of hand configurations. The decoder receives the representation and outputs probabilities indicating pressed keys.

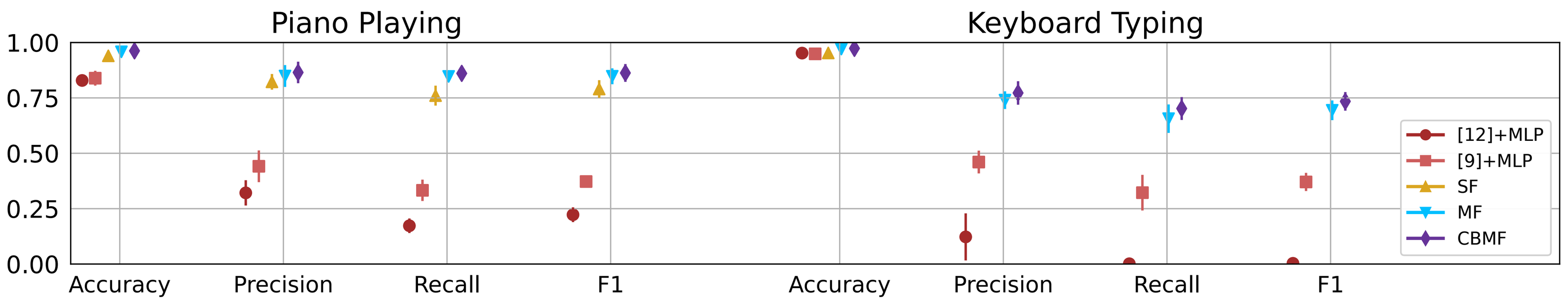
- In the first phase, we train only the encoder to produce hand configurations alone.
- In the second phase, we simultaneously train the pre-trained encoder together with the decoder to predict both hand configurations and key-pressing probabilities.

Data Collection

- The ultrasound sensor is fixed using a 3D-printed housing during recordings.
- The dataset was well obtained from 12 subjects and includes 44K samples for piano playing and 42K samples for keyboard typing.



Tasks Evaluation

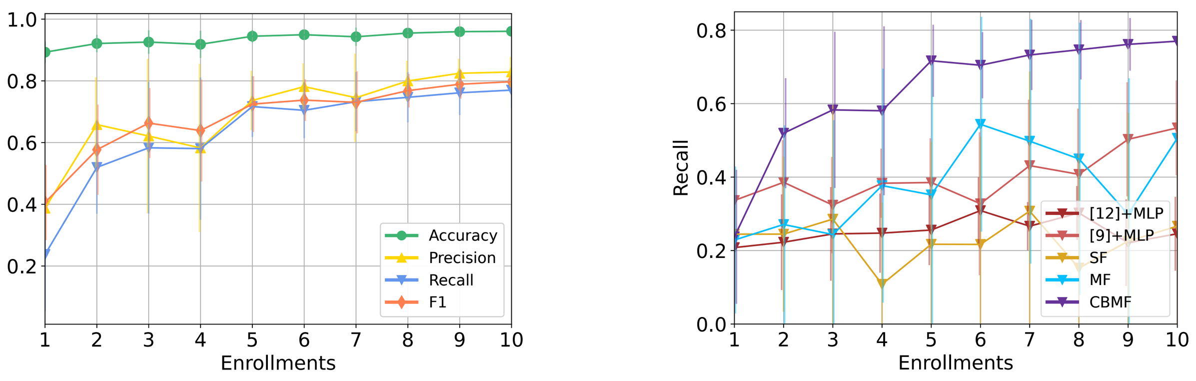


Evaluation of our methods (MF and CBMF) as well as three more baselines.

- Piano playing is significantly easier to detect, unlike keyboard typing, where some baselines were unable to learn.
- Using hand configurations (with our CBMF model) resulted in less false-positive and false-negative predictions.

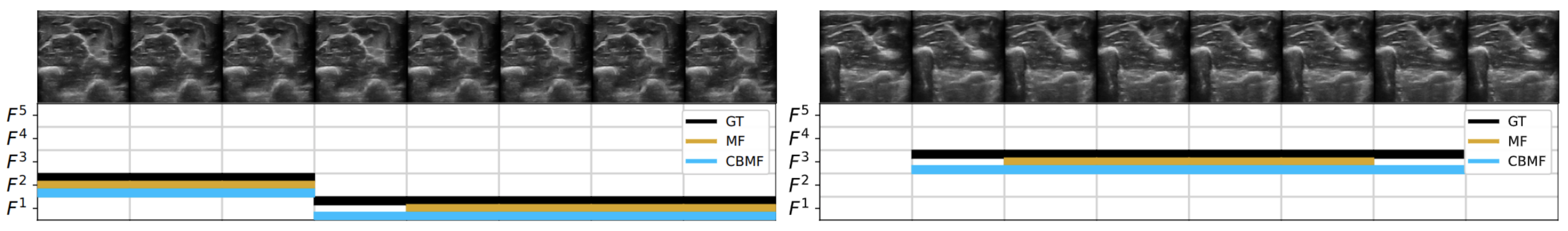
An artistic painting illustrating our desires in creating a computational robotic system that can imitate piano playing perfectly by learning lower arm muscle behavior in real time. (Credit: Daniel Philosoph)

Real-time Applicability

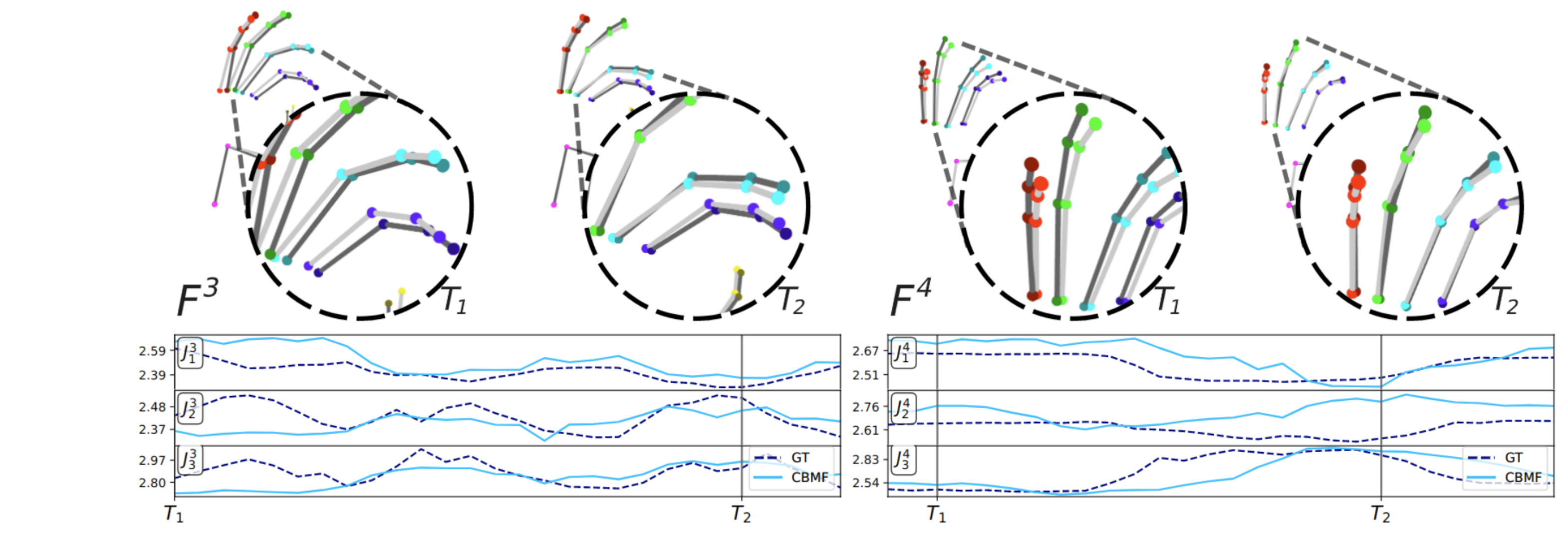


Piano-playing evaluation given a certain number of times the sensor was removed and wore back on (enrollment). At each step, we randomly add a new enrollment to the training set and evaluate an unseen enrollment. (Left) Behavior of the CBMF model for the four metrics and (Right) Recall values for the evaluated methods.

Examples



Executing the MF and CBMF models for piano playing and keyboard typing. The differences are expressed in consistency and correctly deciding when to press or leave a key.



Hand configurations taken from the intermediate representation of the CBMF model (dark grey) as well as the ground truth (light grey). In each rollout, the three graphs represent the joint angles in radians of the relevant finger that is enlarged in the top two illustrations.