LeRobot Med

Hackathon Imitation Learning

1. General Description

Robotics is transforming modern medicine by providing tools that enhance precision, efficiency, and safety in healthcare delivery. Robotic arms, in particular, are beginning to play a critical role in clinical tasks that demand repetitive motion and delicate control—such as patient auscultation.

Imitation learning, a technique within machine learning, enables these robots to acquire skills by observing and replicating human actions. In the medical context, this means a robotic arm can learn how to use instruments like a stethoscope from data collected by a human expert. This dataset typically consists of video recordings from two cameras and the joint angles of the servomotors during the manipulation.

Once trained, the system can accurately reproduce the necessary movements to position the device correctly on the patient's body. This ability opens new possibilities for telemedicine, enables assistance in settings with limited medical personnel, and fosters innovative forms of human-robot interaction in clinical diagnostics.

The robotic arm used in this project is a modified version of the SO-100 model. Specific adjustments were made to the upper arm and forearm components to significantly reduce 3D printing time, without compromising the system's overall functionality. The remaining mechanical and electronic parts were left unchanged, using the original SO-100 components to ensure full compatibility with the existing simulation and control environment.

The integration of intelligent robotic arms into home settings represents a significant advancement in improving the quality of life for elderly individuals or people with disabilities—especially those living alone who may face challenges accessing regular medical care. These systems can assist with basic health monitoring tasks such as periodic auscultation, vital signs measurement, or the administration of simple treatments, thereby reducing reliance on human caregivers. Thanks to their imitation learning capabilities, these robots can adapt to personalized routines and perform procedures autonomously and safely—enhancing patient independence while easing the burden on healthcare systems, particularly in regions with a shortage of medical staff.

2. Hardware / Simulation

A modified version of the SO-100 robotic arm was used in this project, specifically adapted to optimize the manufacturing process. The modifications focused on the upper arm and forearm components, significantly reducing 3D printing time without compromising the overall functionality of the system. The remaining parts—such as the base, wrist, and gripper—were preserved from the original SO-100 design, ensuring full compatibility with the simulation and control environment.

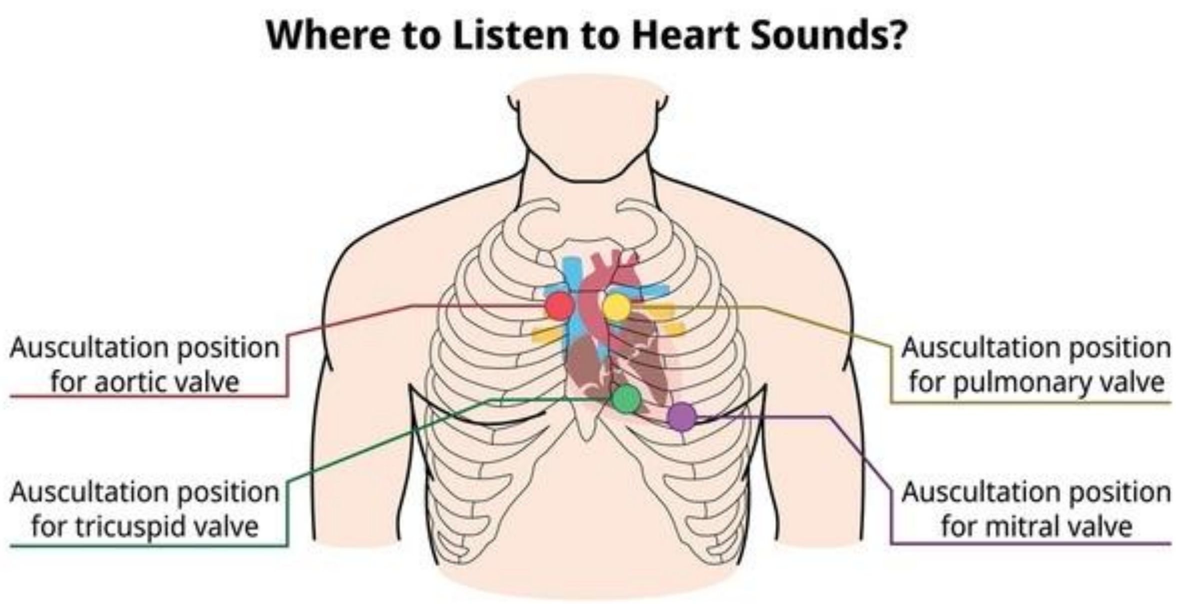
Python was used as the primary programming language for the development of the control system. It was employed to implement data collection routines, as well as the training and validation of the imitation learning model. In particular, a policy based on the **Action Chunking Transformer (ACT)** was trained, allowing the robot to accurately reproduce the expert-demonstrated movements.

For the acquisition of biosignals, a digital stethoscope was used to collect real patient data during auscultation sessions. These signals were later plotted and analyzed to validate both the quality of the recordings and the correct positioning of the device by the robotic arm. This integration of real medical hardware reinforces the clinical relevance of the system, especially in applications involving remote diagnostics and automated healthcare assistance.

3. Data Collection

The data collection process was one of the main challenges of the project, particularly due to logistical constraints related to working alone. Without external assistance, it was necessary to simultaneously interact with the master arm while maintaining a stable posture to allow for the precise placement of the stethoscope on the body. This dual role—controlling the system while also acting as the patient—made the data capture process quite a challenge, requiring multiple attempts to achieve consistent results.

Despite these difficulties, a total of **20 complete samples** were successfully recorded. Each sample included synchronized video from two cameras along with the joint angles of the robotic arm. Each demonstration lasted approximately **30 seconds**, which was sufficient to cover the full trajectory required to position the stethoscope correctly. These data were later used to train the imitation learning model, providing a realistic and representative dataset of the desired behavior.



https://www.shutterstock.com/es/search/auscultating-heart

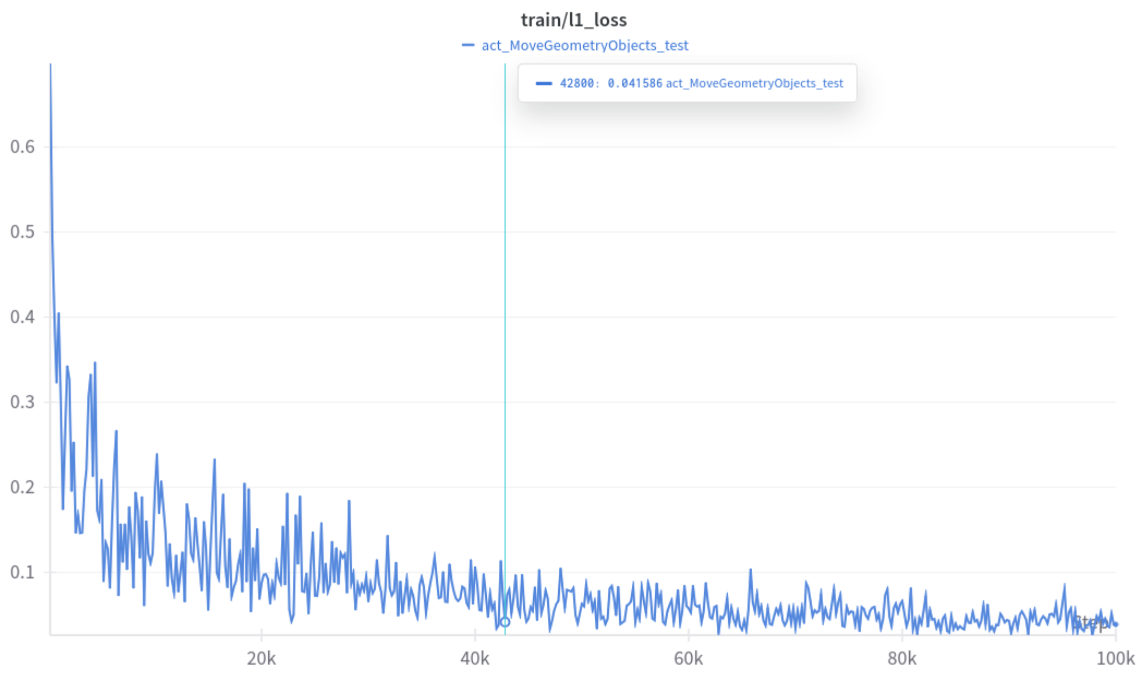
Faltan los ejemplo de capturas

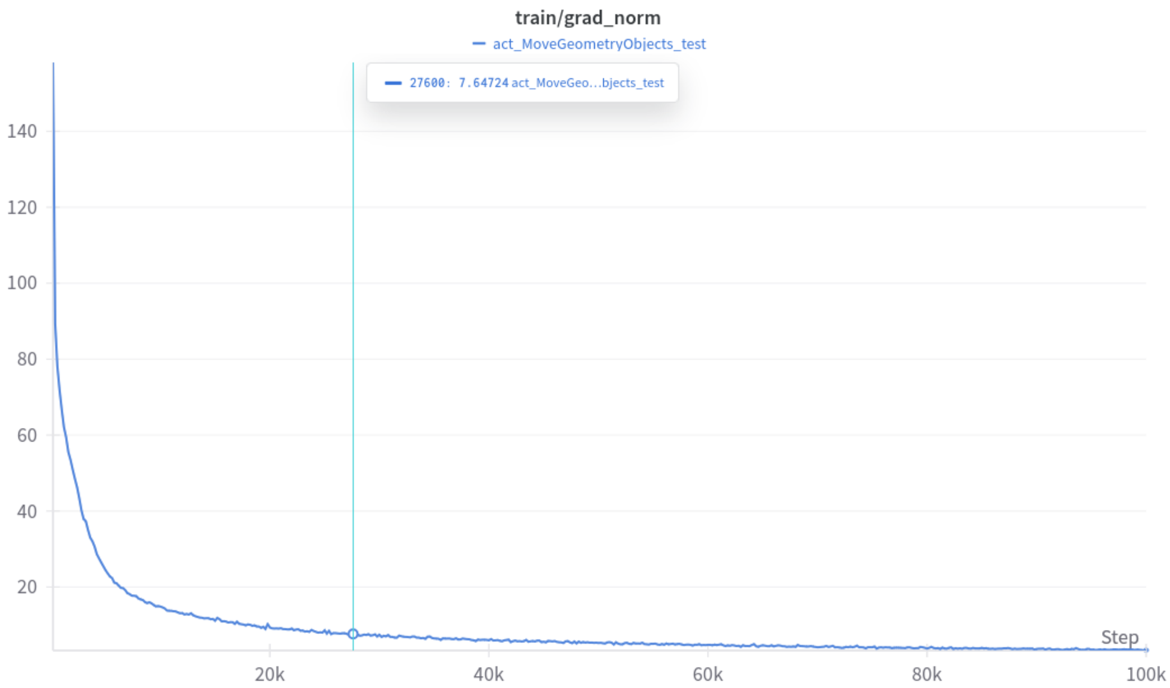
4. Model Training

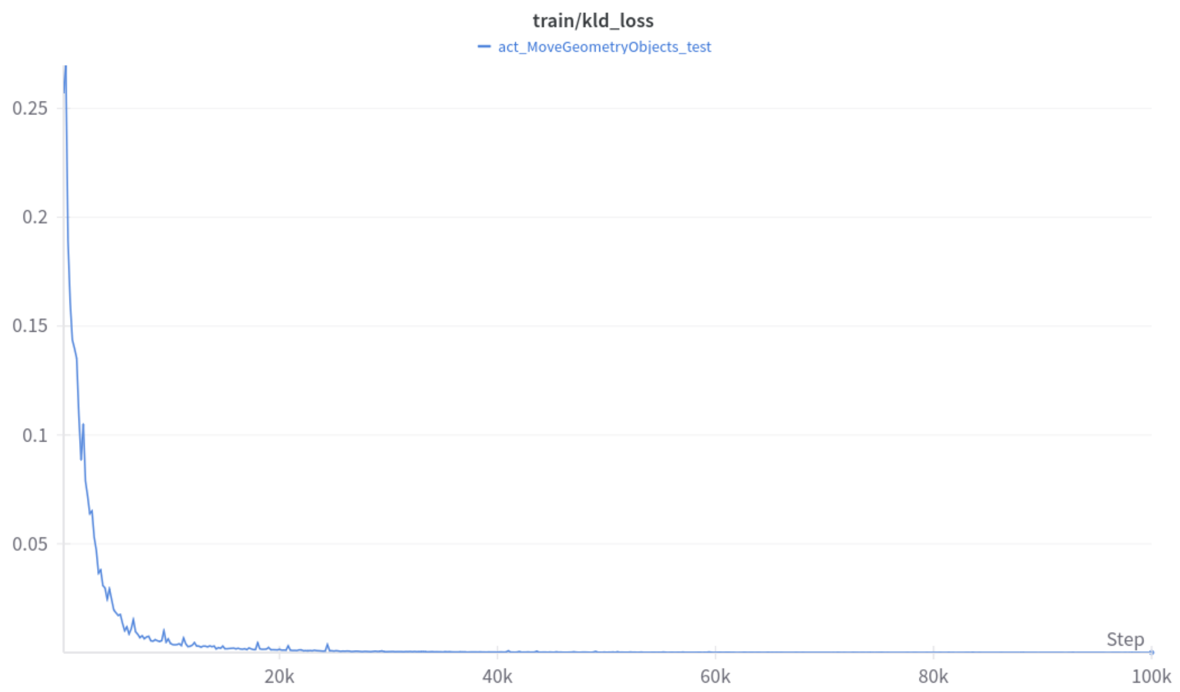
The model used in this project was based on the **Action Chunking Transformer (ACT)** architecture. The training process was carried out using **Google Colab**, which provided access to a **T4 GPU**, allowing for accelerated computation and reduced training times.

The training pipeline included data preprocessing, normalization of the joint angles, synchronization of video frames, and preparation of the sinusoidal control signal used in the temporal embedding. Once the dataset was structured, the model was trained to map input observations—composed of camera frames and joint states—to the corresponding motor actions demonstrated during the data collection phase.

Although training time was limited due to hardware constraints and time restrictions during the hackathon, the resulting model demonstrated a strong ability to generalize the learned behavior in controlled test scenarios. Further training with additional data could enhance both precision and robustness.


train/l1_loss
act_MoveGeometryObjects_test
42800: 0.041586 act_MoveGeometryObjects_test
0.6
0.5
0.4
0.3
0.2
0.1
20k
40k
60k
80k
100k
Step


train/grad_norm
act_MoveGeometryObjects_test
27600: 7.64724 act_MoveGeo...bjects_test
140
120
100
80
60
40
20
20k
40k
60k
80k
100k
Step



5. Deployment and Execution

After completing the training phase, the learned policy was deployed on a desktop computer connected to the robotic arm, the data acquisition base, and a microphone input. The microphone was used to trigger the controlled acquisition of phonocardiographic signals during each execution cycle.

The system operated autonomously during the testing phase, performing all actions without direct human intervention. This allowed for a realistic evaluation of the trained model under conditions similar to its intended clinical application.

Overall, the system achieved an **80% success rate** in positioning the stethoscope correctly on the target area, demonstrating a robust and consistent performance of the learned policy.

6. Evaluation and Results

The system's performance was evaluated by measuring how consistently the robotic arm was able to position the stethoscope accurately over the target area. The primary metric used was **success rate**, defined as the percentage of trials in which the target position was reached precisely relative to the total number of executions. Additional indicators included **execution time per attempt**, as well as qualitative observations such as alignment errors or failure to maintain contact with the surface.

One of the main limitations observed was the **complexity of the task**—accurate placement of the stethoscope for cardiac auscultation requires fine motor control and high spatial precision. This significantly increases the difficulty of the problem and highlights the need for a larger number of training episodes to further improve accuracy and reliability.

Furthermore, due to time constraints imposed by the hackathon environment, the policy was trained for **only one auscultation site**—specifically the aortic valve location, one of the four typical cardiac listening points. As a result, the trained model currently lacks generalization to other areas of the chest, which presents an opportunity for future work.

7. Conclusions and Future Improvements

Since I began working with LeRobot, I have focused my research on imitation learning and its various implementations—an area I am actively developing and plan to consolidate in the near future. In parallel, I have designed and built a custom gripper for the robotic arm, tailored to perform specific tasks and enhance its versatility.

During the hackathon, I gained valuable hands-on experience with **SmolVLA**, a powerful tool for training robotic policies. However, due to time constraints, I was unable to fully integrate it into the system. With more time, I would have collected a larger number of demonstrations and captured all relevant auscultation positions, thereby creating a more diverse and robust dataset for training and evaluating the policy.

One promising direction for future work is the **automation of blood glucose monitoring** using glucometers. In this scenario, the robotic arm could be responsible for handling the lancet device, extracting the blood sample, and processing the reading—thus expanding the range of robotics-assisted medical applications.

Additional Resources and Related Work:

- Greeper and Manipulator – GitHub
- Intelligent Health Care Beacon – Hackster
- Robot Arm Like a Doctor – Hackster