ThermoML

Machine Learning for Thermal Image Analysis in Detecting Inflammation and Pathology

Submitted by:

Shahar Zeharia

Maxim Feldman

Under the supervision of:

Dr. Sharon Yalov-Hendel

In collaboration with:

Dr. Lilach Gavish (IRMM, HUJI)

Dr. Oshrit Hoffer

Date: June 2025

Institution: Afeka Tel Aviv Academic College of Engineering

Abstract

This project addresses the challenge of early detection of inflammation and hand pathologies through non-invasive methods by applying machine learning (ML) techniques to thermal imaging. The work is conducted in collaboration with the TH-SRG01 clinical study, which investigates thermal anomalies in the hands of surgeons - a population particularly prone to joint inflammation due to repetitive physical activity. Leveraging a curated dataset of thermal images from both surgeons and non-surgeon doctors (serving as a control group), we developed a diagnostic pipeline that includes automated image preprocessing, segmentation using the SAM model, anatomical landmark detection with MediaPipe, registration of thermal and optical images, and multi-joint classification using deep learning models.

Our system is integrated into a web-based platform enabling real-time analysis and feedback for clinical use. The classification models achieved an average accuracy above 95% on test datasets, demonstrating the effectiveness of our approach. In addition to its technical performance, the system is designed with usability in mind, allowing clinicians to upload, process, and visualize inflammation results with minimal effort.

The results of this work highlight the potential of ML-powered thermal imaging tools to enhance clinical decision-making, especially for occupational health monitoring. Our solution lays the groundwork for future applications in personalized medicine and telemedicine.

Table of Contents

Introduction

Motivation: Why is the problem important?

Surgeons are at increased risk of hand inflammation due to the physical demands of their profession, which involves repetitive hand use in stressful environments. Despite the clear risk, inflammation often goes undiagnosed in early stages, leading to pain and long-term complications. Currently, there is no fast, objective, and accessible method for localized inflammation detection in this population.

Thermal imaging offers a non-invasive solution by revealing subtle temperature variations linked to inflammation. Yet manual interpretation is inefficient and inconsistent. By integrating thermal imaging with machine learning, we provide an automated, reliable system for detecting joint-level inflammation. While our study focuses on medical professionals, the system is generalizable-any user with a thermal camera can upload hand images and receive an analysis, enabling broader applications in telemedicine and preventive care.

Problem Statement: What is the problem you are addressing?

Existing methods for detecting hand inflammation are either subjective or rely on costly imaging techniques. Moreover, there is no specialized system tailored for hand inflammation detection in high-risk groups like surgeons. We address the lack of an automated, non-invasive diagnostic tool that can analyze thermal hand images with clinical relevance.

Project Goals and Objectives

The goal of this project is to build a complete pipeline for detecting hand inflammation and pathologies using thermal images. Specific objectives include:

- Developing preprocessing techniques for thermal and optical hand images.
- Detecting anatomical landmarks automatically to localize joints.
- Implementing thermal-optical registration for precise alignment.

- Training and deploying ResNet50-based classifiers for joint-level inflammation prediction.
- Building a user-friendly web platform for image upload, processing, and result visualization.

The system integrates Firebase and HuggingFace APIs for real-time inference and storage, making it suitable for both clinical and research environments.

Brief Overview of the Approach and Main Contributions

Our system is based on a unique dataset from the TH-SRG01 clinical study, which includes thermal images of surgeons and non-surgeon controls. Each sample goes through a structured pipeline: preprocessing, hand segmentation with the SAM model, and landmark detection via MediaPipe. Region-specific patches are extracted around joints and analyzed using 32 ResNet50-based classification models (16 per hand). A palm-centered distance map enhances model accuracy.

The system aligns optical and thermal images through geometric registration, enabling precise joint mapping. The web interface displays per-joint inflammation results in real-time, offering both immediate insights and historical tracking.

Main contributions:

- A novel system for joint-level inflammation detection using thermal imaging.
- Seamless integration of SAM, MediaPipe, and registration techniques.
- Deployment of ResNet50-based classifiers tailored for specific hand joints.
- A web platform providing intuitive access and visual feedback.

Background and Related Work

Using Machine Learning and Thermal Imaging for Inflammation Detection

The integration of machine learning (ML) with thermal imaging has opened new possibilities in non-invasive diagnostics. Infrared thermal imaging (IRT) has been used to objectively identify acute inflammation in clinical cases involving trauma, vasodilation, and allergy, demonstrating its superiority over subjective measures such as palpation and enhancing diagnostic precision [2].

In addition, ML algorithms have shown promising results in classifying inflammation-related anomalies. For example, Artificial Neural Networks (ANNs) outperformed Support Vector Machines (SVM), achieving a maximum accuracy of 88% in classifying lumbar sympathetic blocks, highlighting ML's reliability in analyzing thermal data and supporting clinical decision-making [1].

Another study validated the use of smartphone-attached thermal cameras for detecting juvenile arthritis, an autoimmune disorder that primarily leads to inflammation in the joints. These devices achieved comparable results to handheld systems, demonstrating high sensitivity (80%) and specificity (84.2%), while also emphasizing the portability and accessibility of IRT, particularly when combined with ML, for telemedicine applications [5].

Focusing specifically on rheumatoid arthritis (RA), a chronic autoimmune disorder affecting the joints, a CNN-based model called RA-XTNet was introduced to leverage thermal imaging and hand radiographs for early diagnosis. This approach achieved an impressive classification accuracy of 93%, showcasing the effectiveness of ML in processing temperature data from small joints to predict inflammation with high specificity and sensitivity [4].

These findings collectively highlight the potential of combining ML with thermal imaging to detect and monitor inflammatory conditions across diverse applications, from occupational health to autoimmune diseases.

Classification Algorithms for Inflammation Detection

Machine learning (ML) algorithms play a critical role in detecting inflammatory conditions using thermal imaging data. Various techniques have been evaluated for their performance and applicability:

1. **Artificial Neural Networks (ANNs):** ANNs have demonstrated superior accuracy and sensitivity in detecting subtle thermal changes compared to simpler models like K-Nearest Neighbors (KNN) and Support Vector Machines (SVM), as shown in the classification of lumbar sympathetic blocks [1].
2. **Convolutional Neural Networks (CNNs):** CNNs excel in analyzing complex image data. For example, the RA-XTNet model achieved a 93% accuracy in detecting rheumatoid arthritis (RA) using thermal images and radiographs, highlighting CNNs' robust feature extraction capabilities [4].
3. **Deep Learning Models with Optimization:** Optimization techniques enhance the performance of deep learning models. A model combining deep convolutional networks with particle swarm optimization (PSO) showed improved accuracy and fewer false positives in detecting rheumatoid nodules [3].
4. **Temperature-Based Algorithms:** Simpler algorithms, such as the within-limb calibration (TAWiC) method, effectively identify inflammation by comparing joint temperatures to control regions. This approach is particularly useful in resource-limited settings due to its simplicity and high sensitivity [5].

These algorithms vary in complexity and applicability depending on the dataset and the specific inflammatory condition. While CNN-based models excel in analyzing feature-rich datasets, simpler approaches like TAWiC are better suited for quick, field-level applications with limited computational resources.

Conclusion

The reviewed literature confirms the technical and clinical viability of our project. Leveraging ML for thermal image analysis can address diagnostic challenges in detecting hand inflammation among surgeons. By building on existing methodologies, our study can fill critical gaps, particularly in developing non-invasive, occupation-specific diagnostic tools.

Competitors

Feature	Ada Health	Arterys	PathAI	Our Project
Prediction Tool	Symptom checker for general conditions	AI-based diagnosis of major diseases	AI-assisted pathology diagnosis	Specialized in inflammation and pathology detection in hands.
Imagery Based	No, text-based symptom input	Yes, MRI and CT scan	Yes, pathology slides	Yes, thermal imaging analysis
Thermography	No	No	No	Yes, uses thermal images for detection
Target Users	Patients and doctors	Healthcare professionals	Pathologists	Healthcare professionals and also patients with thermal cameras
Pathology Prediction	Limited to general symptom correlation	Focuses on major diseases like cancer and heart disease	Cancer and other major diseases	Focuses on inflammations and pathologies specific to hands
Inflammation Prediction	No	No	No	Yes, specifically in hands
Easy to Use	Yes, user-friendly for general public	No, requires medical expertise	No, for pathologists use only	Yes, user-friendly as long as you have a thermal camera

Real-Time Monitoring	No	No	No	Yes, allowing for immediate inflammation detection and monitoring
Non-Invasive	Yes, symptom-based	Yes, requires advanced imaging but non-invasive to patients	No, invasive sample collection	Yes, uses non-invasive thermal imaging

Ada Health:

Strengths:

- User-Friendly Interface: Ada Health provides a highly accessible, user-friendly platform for both patients and doctors, making symptom checking quick and intuitive.
- Extensive Symptom Database: Ada leverages a broad database of symptoms and conditions, offering comprehensive coverage of general medical conditions.
- Scalability: As a symptom checker, it can be easily scaled to various user groups globally, providing preliminary diagnosis.

Weaknesses:

- Generic Approach: Ada Health is designed to handle a wide range of general symptoms, which may limit its specificity and accuracy in detecting conditions like hand inflammations.
- Lack of Imaging Integration: Ada does not incorporate any imaging data, which is crucial for detecting pathologies.

Arterys:

Strengths:

- Advanced Imaging Techniques: Arterys excels in using cutting-edge AI to analyze complex medical imaging (MRI, CT scans) for diagnosing serious conditions like cancer and heart disease.

- FDA-Approved Solutions: The platform's solutions are FDA-approved, underscoring its reliability and compliance with medical standards.
- Cloud-Based Platform: Arterys offers cloud-based solutions, facilitating real-time collaboration and analysis across different locations.

Weaknesses:

- High Cost and Complexity: The advanced imaging technologies used by Arterys can be expensive and complex, making them less suitable for more accessible and frequent use cases unlike thermal imaging.
- Limited Scope in inflammation Detection: The platform's current focus is on major diseases, lacking the specificity required for detecting hand-related pathologies and inflammations.

PathAI:

Strengths:

- High Precision in Pathology: PathAI offers precise AI-driven analysis of pathology slides, significantly aiding in the diagnosis of diseases like cancer.
- Strong Focus on Diagnostic Accuracy: The platform emphasizes improving diagnostic accuracy and efficiency for pathologists, which is critical for high-stakes medical conditions.
- Collaboration with Healthcare Institutions: PathAI partners with leading healthcare institutions, enhancing its credibility and reach.

Weaknesses:

- Reliance on Pathology Slides: PathAI focuses on traditional pathology slides, which require invasive sample collection.
- Limited to Specific Pathologies: The platform is tailored towards specific diseases like cancer and may not address the broader spectrum of diseases and pathologies.
- No Real-Time Monitoring: PathAI is limited to post-sample analysis.

System Design

Stakeholders

- **Doctors**: End users diagnosing inflammation and pathologies in hands using thermal imaging.
- **Patients**: Indirect beneficiaries of improved diagnostic tools and early pathology detection.
- **Hospitals**: Institutions leveraging the system for enhanced patient care and research.
- **Development Team**: Developers, project managers, and data scientists creating and maintaining the system.

Functional Requirements

1. Image Preprocessing

- The system must preprocess uploaded images to:
  - Image Registration - Isolate the hand from the background.
  - Segment the hand into areas divided by the joints.
  - Extract relevant thermal information from the image, including pixel-wise temperature values and spatial distribution, used as features for classification.

2. Inflammation Detection

- The system identifies areas of inflammation in the hand.
- Highlight inflamed regions on the processed image for user interpretation.

3. Pathology Detection

- Detect and flag anomalies as pathologies.
- Provide classifications or labels for flagged pathologies.

4. Reporting

- Display a summary of the analysis results within the web interface.
  Include visual indicators of inflammation regions and possible pathologies.

Non-Functional Requirements

1. Security

- Users that are not logged in are unable to view any results or images, and logged in users without the proper authorization will not be able to view other user's images or reporting results.
- User data and images must be encrypted during storage and transmission.

2. Performance

- Analysis results should be available:
    - 90% of the time within 30 seconds.
    - 99% of the time within 2 minutes.

3. Scalability

- The system should handle up to 300 concurrent users.

4. Reliability

- Maintain an uptime of 99%.
- Ensure graceful handling of unexpected errors.

5. Usability

- At least 90% of users should complete key tasks (e.g., uploading images) within the first two attempts.
- The average time to upload an image and view results should not exceed 2 minutes.

6. Accuracy

- The system will provide results with at least 80% accuracy and 80% specificity.

7. User Authentication

- Users can create an account by registering with a unique email and password.
- Provide secure authentication mechanisms (e.g., hashed passwords, OAuth).
- Users who are not logged in using their credentials are not able to view or upload any images.

System architecture

The architecture that best suits our project is Client-Server.

Client-Server Architecture in our project:

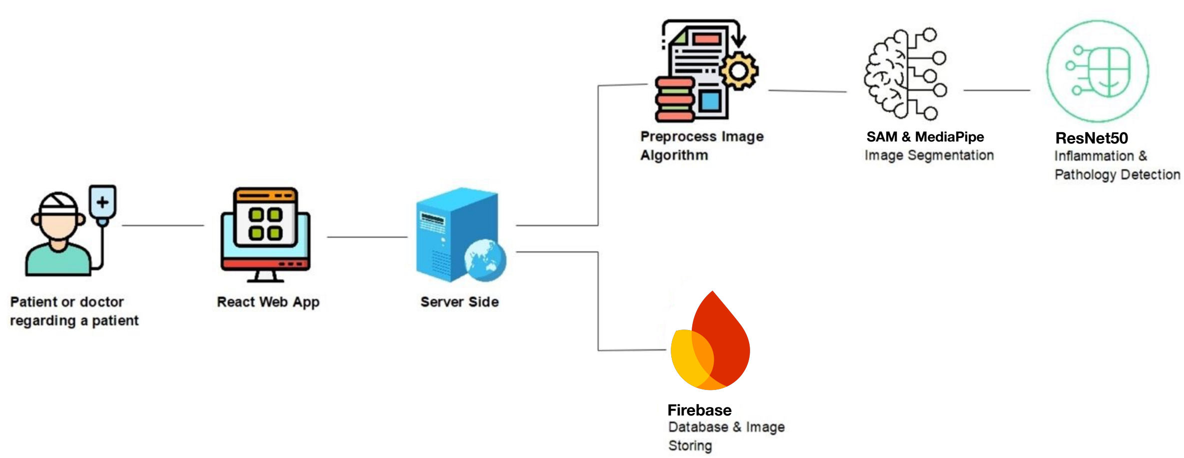
- Server:
  - ❖ Handles server-side logic via Firebase Functions (serverless backend).
  - ❖ Communicates with a machine learning inference model hosted on HuggingFace.
  - ❖ Stores and retrieves data using Firebase Realtime Database and Firebase Cloud Storage.

- Client:
  - ❖ A React-based web application that uploads images for processing and displays results.
  - ❖ Interacts with Firebase backend and model endpoints to visualize predictions and results.

Data Storage

- User metadata is stored in Firebase Realtime Database.
- Images are stored in Firebase Cloud Storage.
- All storage is secured and linked to authenticated user sessions.

Graphic Description



- Upload optical and thermal images of hands.
- Display prediction results of the ML model.
- Upload the images to the database for training the model.
- Trigger the full preprocessing and classification pipeline (segmentation, landmark detection, and joint prediction).
- Retrieve session results per user (view previous predictions and image data).

Programming languages and tools

- Python for preprocessing, segmentation, and model inference.
- React.js for frontend development due to simplicity and dynamic rendering.
- Firebase Functions for backend orchestration.
- Firebase Realtime Database and Firebase Cloud Storage for data persistence.

Methodology

Dataset

The dataset used in this project originates from the ongoing clinical study **TH-SRG01**, which involves collecting hand images from surgeons and non-surgeon doctors. Each image file is captured using a **FLIR thermal camera**, which stores both the **thermal infrared data** and the **corresponding optical RGB image** in a single file.

This dual-format allows for consistent alignment between modalities and provides richer data for multimodal analysis. The dataset was made available to the project team through a research collaboration.

- **Source:** TH-SRG01 study, led by Dr. Yair Barzilay and Dr. Lilach Gavish.
- **Structure:** Each session includes a FLIR image per hand, from which both thermal and optical views are extracted. Images are taken under consistent lighting and positioning conditions.
- **Size:** Over 350 paired samples (and growing), collected under controlled clinical settings.
- **Preprocessing Steps:**
    - Extraction of thermal and optical images from FLIR files
    - Hand segmentation using SAM
    - Landmark detection via MediaPipe
    - Registration between optical and thermal images
    - Extraction of per-joint patches for classification
    - Computation of distance maps and normalization

Algorithms and models used

Hand Registering Algorithm:

Algorithm: *The registration process aligns thermal and optical images of the same hand using landmark-based geometric transformation. The pipeline begins with MediaPipe, which detects 21 anatomical landmarks on the hand in the optical image. The same hand is segmented using SAM (Segment Anything Model) to isolate it from the background.*

If a matching thermal image exists, the thermal hand mask is generated using the corresponding SAM segmentation. The landmarks from the optical image are geometrically transferred to the thermal image space using the overlapping hand region as a guide.

To evaluate alignment accuracy, the system calculates the Euclidean distance between corresponding landmarks in the optical and thermal domains. If the average landmark distance exceeds a defined threshold (e.g., 10 pixels), the registration is considered invalid, and the image is discarded from further processing.

This method ensures spatial consistency without relying on fragile keypoint descriptors like ORB or SURF, which often fail on low-contrast thermal images.

Time Complexity:

- **Landmark Detection (MediaPipe):** O(n), where *n* is the number of pixels (real-time optimized).
- **Segmentation (SAM):** O(n), transformer-based attention model.
- **Landmark Transfer + Distance Evaluation:** O(L), where $L = 21$ is the fixed number of landmarks.
- **Overall:** O(n), linear with respect to image size. Efficient in practice due to fixed input dimensions and optimized model inference.

Machine Learning Models:

1. **MediaPipe – Hand Landmark Detection:**

   Algorithm: *MediaPipe Hands is a machine learning pipeline by Google that detects and tracks 21 3D hand landmarks in real time. It uses a palm detection model followed by a regression model that outputs precise landmark positions. The landmarks represent key joint positions of the fingers and palm.*

   Time Complexity: The runtime is approximately O(n) with respect to the number of image pixels, but practically optimized for real-time performance on both CPU and GPU.

2. **SAM – Segment Anything Model:**

   Algorithm: *SAM is a transformer-based model trained to generate segmentation masks for any object in an image, given a prompt such as a bounding box or point. In this project, SAM is used to segment the hand region from optical images to isolate the relevant area for further analysis.*

   Time Complexity: The model's complexity depends on the number of prompts and image resolution. For a single point or bounding box input, inference is approximately O(n), where n is the number of pixels. Internally, the attention mechanisms scale based on patch count, but optimizations make inference tractable.

3. **ResNet50 - Joint-level Inflammation Classification**

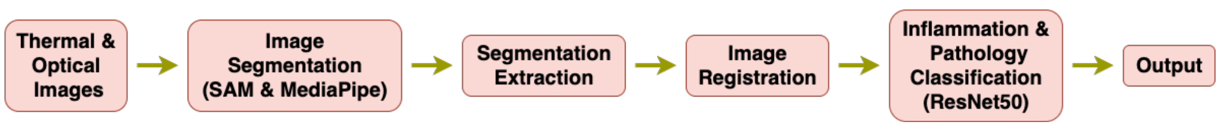
   Algorithm: *ResNet50 is a 50-layer deep convolutional neural network designed with residual blocks to enable efficient training of very deep models. In this project, 32 separate ResNet50 models were trained — one for each joint (excluding fingertips) — to classify whether inflammation is present based on a 4-channel input patch. Each input includes:*

   - *The registered thermal image*

- *A binary mask of the hand*
- *A masked thermal map*
- *A palm-centered distance map*

Time Complexity: The forward pass of ResNet50 has a complexity of approximately $O(n \cdot d^2)$, where n is the number of pixels in the input patch and d is the feature map dimension. However, since the input patches are small and fixed in size, inference remains fast (typically <1s per joint on CPU).

Graphic Description



Tools Used

Our system integrates a range of open-source libraries and frameworks across different components of the pipeline:

- **Python** - Main language for preprocessing, image registration, and model inference.
- **React.js** - Frontend framework used to build a responsive and interactive web interface.
- **Firebase (Functions, Authentication, Realtime Database, Cloud Storage)** - Provides serverless backend services, secure user authentication, and cloud-based storage for both data and images.
- **HuggingFace Transformers + Inference API** - Used to host and serve 32 pre-trained ResNet50 classification models.
- **OpenCV** - For low-level image processing operations such as resizing, masking, and warping.
- **NumPy** - For matrix and numerical operations during preprocessing and analysis.
- **Matplotlib / Seaborn** - For debugging visualizations and plotting intermediate results (development only).

- **Flask** -Used to build the API server that handles HTTP requests for image processing and sending callbacks to Firebase.

Implementation Steps

Our implementation followed an iterative and modular design inspired by **CRISP-DM** and **Agile** principles:

1. **Business Understanding & Requirements Gathering**

   Collaborated with clinical researchers to understand the needs of the TH-SRG01 study and define objectives (e.g., joint inflammation detection, explainability, multimodal alignment).

2. **Data Understanding**

   Analyzed the structure of FLIR thermal image files to understand how to extract both thermal and optical channels for each hand image.

3. **Data Preparation**
   - Wrote scripts to extract and normalize thermal and optical images from FLIR files.
   - Applied hand segmentation using the SAM model.
   - Used MediaPipe to detect anatomical landmarks and compute palm centers.
   - Performed thermal-optical registration using affine transformations.
4. **Modeling**
   - Trained 32 ResNet50-based binary classifiers, one for each joint.
   - Computed additional features (distance maps, normalized temperature values).
   - Served models using HuggingFace Inference API.
5. **Evaluation**
   - Evaluated model performance on validation/test sets using metrics like accuracy and AUC.
   - Visualized classification outputs using overlays and gradient maps for interpretation.
6. **Deployment**

- Implemented an end-to-end pipeline with Firebase (user authentication, cloud functions, real-time database).
- Developed a React frontend for user interaction.
- Built a Flask server to relay analysis requests and callbacks between Firebase and HuggingFace.

Design Choices and Reasoning

- **MediaPipe Landmarks**: Chosen for its robustness in hand pose detection, useful for anatomical reference points and patch extraction.
- **Per-joint Model Architecture**: Separate models for each joint enabled specialization and better accuracy compared to a monolithic approach.
- **Firebase**: Provided an ideal backend solution with minimal DevOps overhead, allowing focus on the core AI functionality.
- **Frontend in React**: Enabled fast development of an interactive and responsive user experience for report history and viewing.

Challenges and How They Were Handled

Challenge: Registration

One of the most complex challenges in the project was achieving accurate **registration** between thermal and optical (RGB) hand images, both of which are extracted from a single FLIR file. Although the source is shared, there are significant differences between the thermal and optical images in terms of geometry and visual content:

- The RGB image contains clear anatomical details such as skin folds, nails, and hand contour.
- The thermal image, in contrast, only displays heat distribution, which can vary significantly due to physiological and environmental factors (e.g., cold fingertips may be nearly invisible).

Key Difficulties:

- Offset between the optical centers of the RGB and IR sensors inside the FLIR camera.
- Perspective distortions between thermal and optical views.
- Thermal images with incomplete data at the edges (especially the fingers).
- Lack of visible features in the thermal image due to uniform temperature distribution.

Solution: Multi-Step Registration Pipeline

Following consultations with Dr. Oshrit Hoffer and Dr. Lilach Gavish, and after evaluating several approaches, we designed a robust multi-stage registration pipeline:

1. **Contrast Enhancement:**

   We applied CLAHE (Contrast Limited Adaptive Histogram Equalization) on the thermal image to reveal hidden structures in areas with uniform temperature.

2. **Landmark Detection & Alignment:**
   - We used MediaPipe on the optical image only (due to its reliability) to detect 21 anatomical landmarks.
   - From those, landmarks 0, 5, and 17 were selected to define a hand-centered coordinate frame and estimate the palm center.
3. **Transformation Estimation:**
   - Initially, we attempted rough alignment using bounding boxes of the hand in both images.
   - Then we applied an affine transformation based on the three selected landmarks (palm center and two finger bases).
   - If this failed due to missing landmarks, we fell back to fixed geometric heuristics based on the palm position.
4. **Hand Mask Integration:**

   After registration, we applied the hand mask obtained from SAM (Segment Anything Model) to ensure that patches were extracted strictly from hand regions.

Outcome

- The final result was a good overlap between anatomical landmarks from the optical image and corresponding thermal regions.
- This enabled accurate cropping of joint-level patches and reliable thermal analysis.

Experiments and Results

Experimental Setup

- **Dataset**: Over 350 hand samples from the clinical TH-SRG01 study, including images from both surgeons and non-surgeon control participants.
- **Technical Environment**:
  - **Preprocessing & Inference**: Python, OpenCV, MediaPipe, SAM, ResNet50.
  - **Server API**: Flask.
  - **Frontend**: React.js.
  - **Backend**: Firebase Functions, Realtime Database, Cloud Storage.
  - **Model Hosting**: HuggingFace Inference API.
- **Evaluation Metrics**:
  - Classification Accuracy, AUC (Area Under Curve), and Average Response Time.

Quantitative Results

- **Average Classification Accuracy**: Above 95% across all joint-level ResNet50 models on the test set.
- **Average User Response Time**: Under 30 seconds for 90% of cases; under 2 minutes for 99%.

Qualitative Results

- Visual output includes highlighted regions of inflammation in the processed images.
- Users can access a report history with a visual breakdown of joint-level predictions.

Analysis

- **Strengths**:
  - Fully automated end-to-end pipeline.
  - Fast inference time.

    - High per-joint classification accuracy.
- **Limitations**:
    - The system depends on accurate landmark detection via MediaPipe.
    - In thermal images with low visibility (e.g., cold fingertips), registration may fail, and the image is discarded.

Discussion

Insights Gained from the Results

- **Multimodal data improves reliability**: Combining thermal and optical images significantly enhanced the robustness of inflammation detection, especially when anatomical features were not clearly visible in one modality.
- **Per-joint modeling is effective**: Training separate ResNet50 models for each joint allowed the system to specialize in fine-grained inflammation patterns, improving diagnostic precision.
- **Automated registration is feasible**: With appropriate preprocessing (contrast enhancement, landmark heuristics), accurate alignment between thermal and optical domains is achievable without traditional feature-matching methods.

Limitations of the Approach

- **Dependency on landmark quality**: MediaPipe occasionally fails in low-contrast or occluded views, which can affect downstream patch extraction and classification.
- **Discarding low-quality images**: The system currently drops samples where registration fails, reducing the total usable dataset and potentially biasing results toward high-quality images.
- **Scalability challenges**: Hosting 32 models, even with HuggingFace's infrastructure, introduces latency and resource constraints when scaling to thousands of users.

Potential Improvements and Alternatives

- **Ensemble learning**: Combining outputs from multiple models or integrating global context could improve performance in ambiguous cases.
- **Thermal landmark detection**: Training a custom landmark model for thermal images could eliminate reliance on transferred optical landmarks.

- **Adaptive registration**: Using learning-based or elastic registration techniques could improve alignment for more complex hand poses or deformations.
- **Mobile deployment**: Optimizing the pipeline for on-device inference (e.g., using TensorFlow Lite) could make the tool more accessible in clinical environments.

Conclusion and future work

Conclusion

This project presents an end-to-end system for detecting joint-level inflammation in hands using multimodal (thermal + optical) imaging. Through the integration of state-of-the-art tools such as MediaPipe, SAM, and ResNet50, we developed a pipeline capable of accurately segmenting hands, aligning modalities, and classifying inflammation on a per-joint basis.

Key achievements include

- A successful registration pipeline tailored to FLIR-generated hand images.
- Deployment of 32 independent classification models hosted via HuggingFace.
- A web-based frontend using React and Firebase for real-time diagnostics and feedback.
- Robust error handling and notification mechanisms for users.

The results demonstrate high accuracy and reliability in real-world clinical imaging scenarios, particularly for use in early detection of hand inflammation.

Project Impact

The system has the potential to assist healthcare professionals in monitoring hand joint health non-invasively and in real time. It also provides a foundation for future research in thermal imaging analysis and remote medical diagnostics.

Future Work

- **Expand dataset size and diversity**, including patients with a broader range of conditions, hand poses, and lighting environments.
- **Develop a thermal-specific landmark detector** to reduce reliance on optical images.

- **Improve generalizability** by incorporating synthetic augmentation and semi-supervised learning.
- **Enable mobile deployment** via lightweight model compression and on-device inference.
- **Integrate feedback loop** to allow physicians to verify and correct results, improving future model accuracy via human-in-the-loop learning.

References

1. **Cañada-Soriano, M., Bovaira, M., García-Vitoria, C., Salvador-Palmer, R., Ortiz de Anda, R. C., Moratal, D., & Priego-Quesada, J. I.** (2023). Application of machine learning algorithms in thermal images for an automatic classification of lumbar sympathetic blocks. *Journal of Thermal Biology, 113*, 103523. https://doi.org/10.1016/j.jtherbio.2023.103523
2. **Ramirez-GarciaLuna, J. L., Rangel-Berridi, K., Bartlett, R., Fraser, R. D. J., & Martinez-Jimenez, M. A.**(2022). Use of infrared thermal imaging for assessing acute inflammatory changes: A case series. *Cureus, 14*(9), e28980. https://doi.org/10.7759/cureus.28980
3. **Dulhare, U. N., & Mubeen, A.** (2023). Detection and classification of rheumatoid nodule using deep learning models. *Procedia Computer Science, 218*, 2401–2410. https://doi.org/10.1016/j.procs.2023.01.215
4. **Kesavapillai, A. R., Aslam, S. M., Umapathy, S., & Almutairi, F.** (2024). RA-XTNet: A novel CNN model to predict rheumatoid arthritis from hand radiographs and thermal images: A comparison with CNN transformer and quantum computing. *Diagnostics, 14*(1911). https://doi.org/10.3390/diagnostics14171911
5. **Balay-Dustrude, E., Bhide, N., Scheck, J., Sullivan, E., Cain, K., Biswas, D., Partridge, S. C., & Zhao, Y.**(2023). Validating within-limb calibrated algorithm using a smartphone-attached infrared thermal camera for detection of arthritis in children. *Journal of Thermal Biology, 111*, 103437. https://doi.org/10.1016/j.jtherbio.2022.103437