Enabling real-time leprosy diagnosis on mobile devices

A study on hand temperature analysis on devices with limited compute power

Daniël Franke

Supervisors: Jan van Gemert, Zhi-Yi Lin, Thomas Markhorst

EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
January 26, 2025

Name of the student: Daniël Franke
Final project course: CSE3000 Research Project
Thesis committee: Jan van Gemert, Zhi-Yi Lin, Thomas Markhorst, Kaitai Liang

Abstract

Leprosy remains a significant health challenge in developing countries, where early diagnosis is crucial to prevent severe disabilities and social stigma. Recent studies have shown that infrared imaging can be used to detect abnormalities associated with leprosy by analyzing hand temperature variations. However, existing diagnostic methods relying on manual annotation of thermal images are time-consuming, lack standardization, and require technical expertise. This research investigates methods for implementing real-time infrared video-based temperature analysis on mobile devices by focusing on hand landmark detection models, model optimization techniques, and evaluation metrics. A comprehensive literature review identified promising models such as MediaPipe Hands, OpenPose, and YOLO variants for hand landmark detection, along with optimization methods like pruning, quantization, and Neural Architecture Search (NAS) to adapt these models for mobile deployment. Furthermore, evaluation frameworks incorporating both performance and capability-oriented metrics were examined to ensure efficient and reliable deployment on resource-constrained devices. This study provides insights into developing a fully automated, mobile-based diagnostic tool for early leprosy detection, highlighting the challenges and opportunities in adapting visual AI models for infrared analysis. Future research should focus on empirical validation of optimized models on mobile platforms.

1 Introduction

Leprosy is an infectious disease that, when not treated in an early stage, can lead to serious disabilities [34]. The disease is classified as an Neglected Tropical Disease by the World Health Organization [52]. Although leprosy is not a major health issue in the western world, it still inflicts harm on more than one billion people, mainly in developing countries like Nepal [34]. In addition to physical deformities of the hands and feet and nerve impairments, patients with leprosy often suffer from discrimination, social exclusion, and even denial of human rights [52].

The treatment of leprosy consists of multi-drug therapy and early diagnosis and treatment are essential to keep the symptoms of the disease to a minimum [52]. The traditional diagnosis of leprosy is based on a clinical examination or a pathological examination [51]. Next to the clinical examination, Cavalheiro [11] has proposed another method of diagnosing leprosy. Cavalheiro [11] has shown that leprosy patients suffer from vascular abnormalities which reduces the blood flow in the limbs, leading to a lower temperature of the skin. In his research, Cavalheiro [11] assessed 17 leprosy patients and 15 people without leprosy as a control group. The result showed a significant interaction between the temperature of the hands and a clinical form of leprosy [11].

To assess the potential of the diagnosis based on hand-temperature, a research collaboration between the Leprosy Research Initiative and Delft University has been founded: Thermographic Assessment of Autonomic Impairment in Leprosy, or TAIL [42]. The goal of the initiative is to determine the value of infrared thermography (IRT) in the diagnosis of leprosy [42]. To this end, F. van den bogaert has devised an application in which one can manually annotate six fixed regions of interest and extract the temperature from infrared video [42]. Bogaert's solution uses manual annotation, which leads to poor reproducibility, introducing non-standardisation and high variability in thermographic analysis [42]. To overcome these issues, Schemkes [42], an Industrial Design student of the Technical University of Delft, has written her master thesis on developing a method for automatic temperature analysis.

In their research, Schemkes [42] adapts an existing machine learning algorithm for detecting hand landmarks. The adapted model is then used to automatically detect the regions of interest and extract the temperatures from the infrared images. While Schemkes' research has shown that this is a feasible solution to run a semi-automatic temperature analysis, it still has a few limitations. The solution still requires significant manual work, for instance by specifying what frames are to be used in the hand detection model. This manual annotation of frames currently happens through a script, which requires proficiency in running Python code on a computer.

Transforming the current solution to an application that could be run on mobile devices would greatly increase the useability of the diagnosing method. However, the implementation of such an application poses a few challenges. Mainly, running visual ai models on real-time video requires hardware with sufficient compute power. This compute power is not always available on mobile devices [3], especially on the ones used in developing countries.

Research Objective

This research paper focuses on the challenges of combining the needed resources for real-time leprosy diagnosis on mobile devices. Our research question is:

What methods are needed to enable real-time leprosy diagnosis on mobile devices through hand temperature analysis?

The goal of this research is to provide an overview of the methods that can be used to enable live visual AI applications on limited hardware. These methods could be applied to the results of our fellow researchers to create a usable application. To achieve this goal, the main research question is divided into three sub-questions:

1. What hand landmark detection methods are currently available for implementing real-time visual AI on infrared video?
2. What techniques can be used to optimize and downsize visual AI models for mobile devices?
3. What metrics can evaluate the performance and efficiency of optimized visual AI models for real-time infrared video analysis?

The first sub-question dives deeper into the current status of visual AI applications and how they can be used on infrared videos. In addition, an overview of techniques for downsizing the visual AI models will be given. Finally, various performance evaluations will be discussed, which can be used to analyze the different models.

2 Methodology

This research aims to investigate approaches and methods for running temperature analysis based on real-time infrared videos on mobile devices. The primary focus is on identifying techniques for implementing real-time visual AI, optimizing models for constrained hardware, and evaluating the performance of these solutions. This will be achieved by conducting a comprehensive literature review, analyzing existing research on real-time visual AI, model optimization techniques, and evaluation frameworks for constrained environments.

Literature Search Strategy

The literature review was conducted using three primary resources:

- TU Delft Library: A comprehensive resource searching across over 30 academic databases, including the Directory of Open Access Journals and IEEE Xplore.
- IEEE Xplore Digital Library: A specialized database for engineering, computer science, and technology research.
- ScienceDirect: A leading database for scientific and technical research, providing access to a wide range of high-quality journals and articles across multiple disciplines, particularly in engineering and applied sciences.
- Google scholar: Used as a supplementary tool to identify additional relevant articles. Its broad coverage was especially useful for identifying papers from diverse journals that may not be indexed in the other databases.

To ensure comprehensive coverage, the search keywords have been grouped by topic. An overview of the used search keywords can be found in Appendix B. Search filters were applied to focus on peer-reviewed articles published in the past ten years and to exclude irrelevant fields or applications.

Inclusion and Exclusion Criteria

In selecting papers for this study, we employed specific inclusion and exclusion criteria to ensure the relevance and quality of the literature reviewed. We focused on studies that delve into real-time visual AI applications utilizing infrared video, as these are central to our research objectives. Additionally, we included papers that detail optimization techniques such as quantization and pruning, given their significance in enhancing AI model performance. Research discussing the deployment of AI on mobile devices or in similarly constrained environments was also considered important, as it aligns with our interest in practical applications of AI in resource-limited settings. Furthermore, articles providing metrics for evaluating AI performance in mobile applications were included to inform our assessment framework.

Conversely, we excluded articles that lacked experimental results or practical applications, as they do not provide empirical evidence necessary for our analysis. Non-peer-reviewed publications were also omitted to maintain the credibility and reliability of the information, ensuring that our study is grounded in strong research. By adhering to these criteria, we aimed to form a body of literature that is both relevant and robust, providing a solid foundation for our investigation into real-time visual AI applications in infrared video.

Categorization and Analysis

The collected papers were categorized into the following clusters, following the sub-questions:

- Real-Time Visual AI Implementation: Exploring existing techniques for applying visual AI on infrared video.
- Optimization Techniques: Focused on downsizing methods such as quantization and pruning.
- Evaluation Metrics: Analyzing benchmarks and metrics used for assessing AI performance on mobile devices.

This categorization allows for an organized synthesis of methods, enabling a detailed comparison of approaches across studies. Additionally, Zotero [48] was used to manage references and organize findings. Zotero is a software package that allows you to collect all literature used and easily tag and categorize them.

Justification of Methodology

A survey methodology was chosen to systematically identify, analyze, and synthesize existing research. By leveraging the TU Delft Library, IEEE Xplore, ScienceDirect and Google Scholar, a broad and reliable set of academic resources was accessed. The categorization ensures that the findings are relevant to the research questions, facilitating a clear understanding of the field's current state.

Limitations

The methodology employed in this study is subject to several limitations that could influence the findings. One potential source of bias arises from the selection of keywords, which may have inadvertently excluded relevant studies, as well as the decision to exclude non-English papers, potentially narrowing the scope of the literature reviewed. Additionally, limited access to certain databases and proprietary studies constrained the breadth of the research, leaving some potentially valuable sources unexplored. Lastly, the emphasis on recent publications, while ensuring relevance to current advancements, may have overlooked earlier foundational work that could provide historical context or insights into the evolution of the field. These limitations underscore the need for further research to address these gaps and ensure a more comprehensive understanding of the topic.

3 Findings

Hand landmark models

Hand landmark detection has seen significant advancements over recent years, with several models emerging to address challenges in accuracy, efficiency, and adaptability. These models form the foundation for applications ranging from augmented reality to gesture recognition. Below is an overview of the key models and techniques.

One widely adopted framework for hand landmark detection is the MediaPipe Hand Landmarker [41], an optimized neural network capable of identifying 21 key points on the hand in real time. Designed with efficiency in mind, MediaPipe excels at delivering accurate results on mobile and embedded devices [30]. Its lightweight architecture enables high performance without the need for intensive hardware, making it particularly attractive for applications run on devices with limited compute power. The Media Pipe hand landmarker model is used in many real-time applications, like robotics [44], recognizing sign-language [2], hand digit detection [12] and medical diagnosis of nerve injury [19]. Despite its strengths, MediaPipe's reliance on RGB training data limits its direct applicability to infrared imaging.

Another prominent model is OpenPose [9], which extends beyond single-hand detection to handle multi-person pose estimation. OpenPose employs a bottom-up approach using Part Affinity Fields to detect body parts and associate them with individual subjects in an image [9]. This capability allows for robust performance in complex scenarios, such as overlapping hands or occlusions [9]. However, its computational intensity can be a barrier to deployment on mobile devices or systems with limited resources, necessitating further optimization for real-time applications. Examples of where OpenPose is used are cattle recognition [50], athletics competitions [13] and the study of lower-limb movement [22].

In addition to these frameworks, object detection models such as those in the YOLO family [38] have been adapted for hand detection tasks. YOLO's architecture, known for its speed and simplicity, involves a single network that predicts an entire image in a single pass to identify objects and their bounding boxes [1]. Variants such as YOLOv4 and YOLOv5 offer a balance between accuracy and efficiency, making them suitable for real-time applications on constrained hardware [20]. While originally developed for general object detection, these models have been adapted to hand landmark detection as well [47]. Kristo [25] compared the thermal object detection capabilities of YOLO with other models. While the models scored comparable in detection results, YOLO achieved significantly faster inference time [25]. However, YOLO's reliance on RGB training data limits its direct applicability to infrared imaging

Advanced techniques such as multiview bootstrapping have also emerged as effective tools for enhancing hand landmark detection [14, 43]. This method leverages multiple camera views to generate robust annotations for training datasets, using triangulation to overcome challenges such as occlusions and complex hand poses [43]. By iteratively improving the accuracy of a keypoint detector through gradient descent, multiview bootstrapping has demonstrated strong potential for high-precision applications [43]. However, its reliance on multi-camera setups limits its practicality for single-camera or mobile systems, which is the current scope of this project.

While out of scope for this research, it is critical to acknowledge the RGB-infrared domain gap as a significant challenge for implementing AI-driven hand landmark detection in infrared imaging. All models discussed in this study, including MediaPipe, OpenPose and YOLO variants, are initially designed and trained in the RGB domain, where rich datasets are widely available. When applied to infrared imagery, these models face a substantial decline in performance due to the differences in image properties, texture, and lighting conditions between the two domains [6]. This domain gap limits their effectiveness without further adaptation. Bridging the RGB-infrared domain gap is essential for ensuring accurate landmark detection in infrared images. Approaches such as domain adaptation [6, 16] and style transfer [17, 24] provide promising solutions. Addressing this gap would significantly enhance the utility and reliability of AI systems for infrared-based diagnostics.

Overall, the development of hand landmark detection models reflects a dynamic field balancing accuracy, efficiency, and adaptability to diverse contexts. While frameworks like MediaPipe, OpenPose and OpenPose excel in general-purpose tasks, their adaptation to specific domains such as infrared imaging remains underexposed.

Optimization of AI handlandmark models

Optimization methods for visual AI models play a crucial role in adapting deep learning frameworks for resource-constrained environments like mobile devices. The following section explores a range of techniques, focusing on pruning, quantization, and other tailored approaches for real-time AI applications on edge devices.

Pruning Techniques

Pruning aims to reduce the complexity of neural networks by eliminating redundant parameters or neurons, thereby improving efficiency without significantly compromising performance. Traditional approaches like weight pruning [26, 28, 54] remove individual parameters with low importance, which can lead to sparsity in the network but may require specialized hardware or libraries to exploit the sparsity.

Recent advances have introduced structured pruning methods, such as filter pruning, which removes entire filters in convolutional layers [28, 55]. This approach simplifies implementation and maintains compatibility with standard hardware acceleration libraries, achieving substantial reductions in inference costs. Building on this, cluster pruning offers a hardware-aware strategy that removes groups of filters based on their collective impact, resulting in better alignment with AI hardware constraints and improved overall efficiency [15].

Another similar, less commonly explored approach is activation pruning, which directly reduces the number of active neurons in convolutional layers [4]. By focusing on pruning neurons that contribute minimally to the output, this method decreases both latency and power consumption, making it particularly suitable for real-time systems. Activation pruning has demonstrated improvements in efficiency while maintaining or even enhancing classification accuracy in certain resource-constrained scenarios [4].

Quantization

Quantization reduces the precision of weights and activations in neural networks, typically converting 32-bit floating-point values to lower-precision formats like 8-bit integers [28]. This technique significantly reduces model size and speeds up inference by enabling faster arithmetic operations for low-precision computation [33, 45].

Hardware considerations also influence the impact of quantization. Quantized models often perform fixed-point arithmetic operations instead of floating-point ones, significantly reducing energy consumption and computational complexity [33]. Techniques like error correction during quantization have further improved performance by minimizing errors across layers, ensuring accurate inference even in highly quantized networks [53].

Neural Architecture Search (NAS)

NAS automates the design of efficient neural network architectures optimized for specific hardware constraints and tasks [32]. By exploring a predefined search space of architectures, NAS identifies configurations that balance accuracy, speed, and resource usage [32]. While computationally intensive, NAS-generated architectures often outperform manually designed networks, making it a promising direction for resource-constrained applications.

Recent advancements have made NAS more practical. Modular search spaces, for instance, simplify the process by constructing architectures from smaller, repeatable units, such as cells [39]. These units are stacked to form the final model, reducing search complexity while maintaining flexibility for various applications [39]. Gradient-based approaches like Differentiable Architecture Search (DARTS) have further improved efficiency by enabling continuous optimization instead of relying solely on discrete searches [29].The obvious downside of Neural Architecture Search is that it would require a newly trained neural network and cannot be applied to existing models.

Frame Sampling

Currently, the data collection protocol involves recording infrared videos at a rate of one frame per second over a 15-minute recovery period following a Cold Pressure Test, a one-minute immersion in cold water (A. Knulst, personal communication, January 12, 2025). This dense sampling was initially adopted to capture any rapid thermal changes and mitigate the impact of motion artifacts. However, analysis of temperature recovery data (Appendix A) suggests that significant thermal responses occur over periods of 30–40 seconds, indicating that high-frequency sampling may be unnecessary.

Applying the Nyquist-Shannon Sampling Theorem, which states that a signal must be sampled at least twice the rate of its highest frequency to be accurately reconstructed [27], suggests that a sampling interval of 10–15 seconds would sufficiently capture meaningful temperature changes. The frame sampling could vary over time and could capture more frames in the start, obtaining more informative frames. Implementing an adaptive frame sampling strategy can reduce computational overhead. This method allows the system to skip non-informative frames and focus computational resources on critical moments, thereby improving processing efficiency without sacrificing diagnostic accuracy.

Simplifying ROI Analysis

Currently, the diagnostic protocol involves analyzing four regions of interest (ROIs) on the fingertips (A. Knulst, personal communication, January 12, 2025), with the potential to expand this number in future studies. Each ROI's mean temperature and standard deviation are calculated for every frame, though the standard deviation is not presently used in diagnosis. Eliminating the calculation of standard deviation or making it optional could reduce processing demands. On the other hand, allowing the selection of more than four ROI's would enable future research on which ROI's are the most relevant in leprosy diagnosis.

Real-time or post-hoc

When developing computer vision applications for leprosy diagnosis, two primary approaches can be considered for processing infrared (IR) video data: real-time inference and post-processing. Each method has its own set of advantages and challenges, as demonstrated by various applications in the field.

1. Real-Time Inference:

In real-time inference, the application processes each frame of the IR video as it is captured, immediately passing the data through the hand landmark detection model and calculating temperatures. This approach enables instant feedback, which is beneficial in scenarios requiring immediate analysis. Examples of applications where real-time inference is used are in manufacturing [23], self-driving cars [21], surveillance [49] and augmented reality [18]. Models like YOLO (You Only Look Once) are designed for such real-time object detection tasks, processing images in a single pass to achieve high inference speeds [49].

The advantage of real-time inference is that it provides instant analysis, allowing for prompt decision-making. The challenge of such inference is that it places a high computational load on the device running the inference. Processing high-resolution IR video frames in real-time requires significant computational resources, which can be challenging on mobile devices with limited processing power.

2. Post-Processing:

Alternatively, the application can record the IR video and perform all necessary calculations after the recording session. This method allows for more complex analyses in cases where compute power is limited. Examples of applications where post-processing is used can be found in the medical sector. In healthcare, imaging data from for instance a CT or MRI scan, is often collected and then analyzed post hoc to identify patterns or anomalies [35].

The advantage of using post-processing techniques is that it permits the use of more sophisticated algorithms that may be too computationally intensive for real-time application but provide higher accuracy in diagnostics. A disadvantage of post-processing is that results need to be computed after recording, meaning that they will not be available immediately. Moreover, post-processing requires sufficient storage capacity to save raw video data for later processing, which can be a constraint on devices with limited memory. Saving this sensitive patient data for post-processing also poses significant privacy concerns.

Evaluation of AI handlandmark models

Evaluation methods are integral to the development and deployment of AI systems, especially in resource-constrained applications such as mobile-based infrared temperature analysis. Different AI models often require tailored evaluation methods suited to their specific applications. For instance, generative AI models benefit from qualitative evaluation techniques, such as human assessments of output quality and creativity [5], whereas medical AI systems often prioritize metrics like interpretability and transparency to ensure that clinicians can trust and understand the model's predictions [46]. This section explores existing evaluation frameworks, highlighting their strengths and limitations in assessing the performance, reliability, and scalability.

Performance-Oriented Evaluation
Traditional AI evaluation methods predominantly focus on performance-oriented metrics, which measure the system's ability to achieve specific outcomes on predefined benchmarks [7]. These include accuracy and precision, which are commonly used to assess classification and detection tasks [7]. For example, in their paper on clinical decision support, Magrabi [31] uses accuracy scores to balance precision and recall, ensuring robust performance across diverse datasets.

Capability-Oriented Evaluation
In contrast to performance-oriented approaches, capability-oriented evaluation focuses on the underlying capabilities that enable AI systems to generalize across diverse tasks and environments [7, 40]. This paradigm assesses the system's adaptability, robustness, and potential for safe deployment. However, capability-oriented evaluation remains underutilized in practical AI development due to its complexity and reliance on comprehensive testing environments [7].

This paradigm is particularly relevant in our research, where generalizing hand landmark detection models from the RGB to the infrared domain is crucial. Traditional benchmarks often fail to capture this capability because they rely on narrow, domain-specific datasets. For example, RGB-trained models, such as MediaPipe and YOLO, excel in environments closely aligned with their training data but struggle when applied to thermal imaging due to the differences between RGB and infrared data. Capability-oriented evaluation addresses this gap by emphasizing robust generalization, even when there is a domain shift [8].

The TEHAI framework
Another framework that focuses on the underlying capabilities of AI models is the TEHAI framework [37]. The TEHAI framework emphasizes evaluating capability alongside utility and adoption to ensure AI systems are effective in real-world applications. This framework highlights the importance of dataset integrity, ensuring training and validation data align with the system's intended use, and internal and external validity, ensuring the model's generalizability to diverse conditions. Such evaluation frameworks provide the structure necessary to assess how AI systems perform in dynamic healthcare environments. Examples of where TEHAI is used are in COVID-19 studies [10] and in evaluating large language models for use in healthcare [36].

4 Responsible Research

Ethical Considerations

The development and deployment of AI-driven applications for medical diagnostics raises critical ethical concerns. Addressing these issues is essential to ensure that research is conducted responsibly and contributes positively to healthcare outcomes. This section discusses key ethical considerations, including data privacy, fairness, and the impact of AI-driven decision-making.

Data Privacy and Security
The application involves processing and potentially storing sensitive patient data, such as infrared (IR) thermographic videos, which may be traceable to individual patients. Ensuring compliance with data protection regulations, such as the General Data Protection Regulation (GDPR), is important. Specific measures include anonymization. Where possible, patient-related data should be de-identified to prevent linking the data to individual identities. Secure Storage should be used to encrypt and store the data safely, with access restricted to authorized personnel. Finally, patients should provide explicit consent for the collection and use of their data in the research and diagnostic processes.

Bias and Fairness
AI models trained on biased datasets may produce inadequate outcomes, potentially leading to misdiagnosis or exclusion of certain populations. For instance, if the training data primarily includes images from one demographic group, the model may underperform on others. Addressing this concern requires diversifying training datasets to include a range of skin tones, hand shapes, and thermal imaging patterns. The models should be regularly audited to detect biases.

Impact of AI on medical decision making
Automating the diagnostic process carries the risk of over-reliance on AI models, which could lead to reduced clinical oversight. To mitigate this, the application should be designed to assist, not replace, healthcare professionals. Providing clear outputs ensures that clinicians understand the model's decisions and can intervene when necessary.

Reproducibility of Method

The following measures have been taken to increase reproducibility in this study:

Transparent Methodology
The literature review process, including database selection (TU Delft Library, IEEE Xplore, and Google Scholar), search

keywords, and inclusion/exclusion criteria, has been clearly documented. This enables other researchers to replicate the review process and verify the findings.

Non-proprietary literature and open-source tools
This research does not include proprietary literature and exclusively uses open-source tools, such as MediaPipe, OpenPose and YOLO, to ensure accessibility and facilitate further development.

5 Discussion

The objective of this research was to explore methods for implementing real-time infrared video-based temperature analysis on mobile devices to support leprosy diagnosis. This discussion analyzes the findings related to hand landmark detection models, optimization techniques, and evaluation methods, contextualizing them within existing research and identifying their practical implications for mobile deployment.

Interpretation of Findings

The review of existing hand landmark detection models, including MediaPipe Hands, OpenPose, and adaptations of YOLO, highlighted the diversity of approaches in balancing accuracy, computational efficiency, and adaptability. MediaPipe demonstrated strong potential for mobile deployment due to its lightweight design and real-time performance capabilities. However, its reliance on RGB data limits its direct use for infrared imaging, necessitating domain adaptation techniques. OpenPose, while highly accurate in complex scenarios, is computationally intensive and thus less suitable for resource-constrained devices without significant optimization. YOLO models, known for their high inference speed, present a viable option for real-time applications, but adapting them to detect hand landmarks in infrared images poses the same challenges as for MediaPipe and OpenPose.

Optimization techniques such as pruning, quantization, and Neural Architecture Search (NAS) were found to be essential for adapting complex models to mobile platforms. Pruning methods effectively reduce computational load while preserving model performance. Quantization further enhances efficiency by lowering model precision, leading to reduced memory and faster inference times, though it can compromise accuracy if not carefully implemented. NAS offers promising results in discovering architectures optimized for specific hardware constraints but introduces significant computational overhead during the search process, making it less practical for existing models Just as important, adaptive frame sampling and posthoc processing were discussed as strategies for efficiency. By reducing the frequency of frame analysis and delegating intensive computations to post-processing, these methods address the practical limitations of mobile hardware without sacrificing diagnostic accuracy. Together, these findings suggest that combining traditional optimization techniques with adaptive sampling and posthoc processing can enable scalable and resource-efficient solutions for mobile infrared temperature analysis.

Evaluation of these models revealed that traditional performance-oriented metrics such as accuracy, precision, and latency are insufficient when assessing models for mobile deployment. Capability-oriented evaluations, which assess adaptability and robustness across tasks and environments, provide a more comprehensive understanding of model performance. This is particularly relevant for adapting RGB-trained models to infrared imaging, where generalization across domains is critical. Furthermore, context-specific evaluations that account for mobile constraints, such as energy efficiency, are essential for ensuring practical utility.

Comparison with Previous Work

Compared to existing solutions in temperature-based leprosy diagnosis, this research advances the field by systematically reviewing methods for automating the process through hand landmark detection in infrared imagery. Previous work, such as Schemkes' semi-automatic annotation tool, addressed some of the reproducibility and standardization issues in temperature analysis but still required manual frame selection and Python scripting expertise. By exploring fully automated detection methods and mobile deployment strategies, this research bridges the gap between proof-of-concept models and scalable, user-friendly diagnostic tools.

In contrast to existing studies focused on performance improvements for AI models in general object detection tasks, this research contextualizes optimization techniques specifically for mobile infrared applications.

Limitations

Despite providing a comprehensive overview of existing models and optimization techniques, this study faces certain limitations. The literature review was constrained by access to certain proprietary resources and primarily focused on recent publications, possibly excluding foundational research in AI optimization. Additionally, the practical evaluation of the discussed models and optimization techniques was not conducted, limiting the findings to theoretical analysis. Real-world testing of these models on infrared data and mobile devices is necessary to validate the conclusions drawn from the literature.

6 Conclusions and Future Work

This research investigated methods to enable real-time infrared video-based temperature analysis for mobile devices, focusing on hand landmark detection models, optimization techniques, and evaluation frameworks. The findings highlight opportunities to optimize AI systems for constrained environments, such as those used for leprosy diagnosis.

Conclusions

The study demonstrates that the performance demands of the landmark detection model can be significantly reduced by leveraging a combination of traditional and context-specific optimization techniques. Traditional techniques such as pruning and quantization enable downsizing of AI models for mobile deployment. These methods allow for reductions in model size and computational demands while preserving accuracy, making them vital for resource-constrained devices. Context-specific techniques like adaptive frame sampling and post-processing strategies can further reduce the required compute power. Specifically, analyzing the infrared video at 15-second intervals instead of 1-second intervals allows for substantial reductions in computational load without compromising diagnostic accuracy. Combining adaptive frame sampling with making the calculations after recording ensures that even devices with limited processing power, such as commonly available smartphones, can feasibly run the temperature analysis.

Given this reduction in performance requirements, the choice of the specific hand landmark detection model should prioritize Capability-Oriented Evaluation metrics rather than traditional Performance-Oriented metrics like speed or latency. The key Capability-Oriented mertics are:

- Generalization across domains: The ability of the model to adapt from its RGB-trained domain to the infrared domain is critical. Models must effectively identify relevant hand landmarks in infrared images despite domain-specific challenges.
- Robustness: Ensuring consistent performance in varying conditions, such as different lighting environments or temperature ranges, is key for real-world deployment.
- Practical adaptability: Lightweight models that are compatible with mobile hardware and require minimal customization will be preferred.

Future Work

Future research should focus on addressing several key areas to build upon the findings of this study. One important direction is the adaptation of hand landmark detection models, such as MediaPipe and YOLO, from RGB-trained domains to infrared imagery. Additionally, the theoretical advantages of adaptive frame sampling and post-processing need to be validated through real-world testing to ensure their impact on computational efficiency and diagnostic accuracy under practical conditions. Furthermore, additional region-of-interest selection should be explored as a means to investigate the diagnostic relevance of specific ROI's. These future efforts will contribute to the development of scalable, efficient, and reliable tools for infrared-based healthcare diagnostics in resource-constrained settings.

.

References

[1] Aabidah Nazir and M. Arif Wani, eds. *You Only Look Once - Object Detection Models: A Review*. Meeting Name: INDIACom. Piscataway, NJ: IEEE, 2023. 1 p. ISBN: 978-93-80544-47-2.

[2] Shahan Ahmed, Sumit Kumar Kar, and Sarnali Basak. "A Novel Approach for Recognizing Real-Time American Sign Language (ASL) Using the Hand Landmark Distance and Machine Learning Algorithms". In: *2023 IEEE 9th International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*. 2023 IEEE 9th International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE). Thiruvananthapuram, India: IEEE, Nov. 25, 2023, pp. 171–176. ISBN: 9798350319651. DOI: 10.1109/WIECON-ECE60392.2023.10456414. URL: https://ieeexplore.ieee.org/document/10456414/ (visited on 01/07/2025).

[3] Ali Almadan and Ajita Rattani. "Compact CNN Models for On-device Ocular-based User Recognition in Mobile Devices". In: *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. 2021 IEEE Symposium Series on Computational Intelligence (SSCI). Orlando, FL, USA: IEEE, Dec. 5, 2021, pp. 1–7. ISBN: 978-1-72819-048-8. DOI: 10.1109/SSCI50451.2021.9660033. URL: https://ieeexplore.ieee.org/document/9660033/ (visited on 01/26/2025).

[4] Arash Ardakani, Carlo Condo, and Warren J. Gross. "Activation pruning of deep convolutional neural networks". In: *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. 2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP). Montreal, QC: IEEE, Nov. 2017, pp. 1325–1329. ISBN: 978-1-5090-5990-4. DOI: 10.1109/GlobalSIP.2017.8309176. URL: http://ieeexplore.ieee.org/document/8309176/ (visited on 01/07/2025).

[5] Ajay Bandi, Pydi Venkata Satya Ramesh Adapa, and Yudu Eswar Vinay Pratap Kumar Kuchi. "The Power of Generative AI: A Review of Requirements, Models, Input–Output Formats, Evaluation Metrics, and Challenges". In: *Future Internet* 15.8 (July 31, 2023), p. 260. ISSN: 1999-5903. DOI: 10.3390/fi15080260. URL: https://www.mdpi.com/1999-5903/15/8/260 (visited on 01/25/2025).

[6] Jürgen Beyerer, Miriam Ruf, and Christian Herrmann. "CNN-based thermal infrared person detection by domain adaptation". In: *Autonomous Systems: Sensors, Vehicles, Security, and the Internet of Everything*. Autonomous Systems: Sensors, Vehicles, Security and the Internet of Everything. Ed. by Michael C. Dudzik and Jennifer C. Ricklin. Orlando, United States: SPIE, May 3, 2018, p. 8. ISBN: 978-1-5106-1797-1 978-1-5106-1798-8. DOI: 10.1117/12.2304400. URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10643/2304400/CNN-based-thermal-infrared-person-detection-by-domain-adaptation/10.1117/12.2304400.full (visited on 01/25/2025).

[7] John Burden. *Evaluating AI Evaluation: Perils and Prospects*. Version Number: 1. 2024. DOI: 10.48550/ARXIV.2407.09221. URL: https://arxiv.org/abs/2407.09221 (visited on 01/07/2025).

[8] Ryan Burnell et al. "Not a Number: Identifying Instance Features for Capability-Oriented Evaluation". In: *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*. Thirty-First International Joint Conference on Artificial Intelligence {IJCAI-22}. Vienna, Austria: International Joint Conferences on Artificial Intelligence Organization, July 2022, pp. 2827–2835. ISBN: 978-1-956792-00-3. DOI: 10.24963/ijcai.2022/392. URL: https://www.ijcai.org/proceedings/2022/392 (visited on 01/25/2025).

[9] Zhe Cao et al. *OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*. May 30, 2019. DOI: 10.48550/arXiv.1812.08008. arXiv: 1812.08008[cs]. URL: http://arxiv.org/abs/1812.08008 (visited on 11/26/2024).

[10] Aaron Edward Casey et al. "Application of a Comprehensive Evaluation Framework to COVID-19 Studies: Systematic Review of Translational Aspects of Artificial Intelligence in Health Care". In: *JMIR AI* 2 (July 6, 2023), e42313. ISSN: 2817-1705. DOI: 10.2196/42313. URL: https://ai.jmir.org/2023/1/e42313 (visited on 01/25/2025).

[11] Aretusa Lopes Cavalheiro et al. "Thermographic analysis and autonomic response in the hands of patients with leprosy". In: *Anais Brasileiros De Dermatologia* 91.3 (2016), pp. 274–283. ISSN: 1806-4841. DOI: 10.1590/abd1806-4841.20164612.

[12] Rung-Ching Chen et al. "Automatic Digit Hand Sign Detection With Hand Landmark". In: *2022 International Conference on Machine Learning and Cybernetics (ICMLC)*. 2022 International Conference on Machine Learning and Cybernetics (ICMLC). Japan: IEEE, Sept. 9, 2022, pp. 6–11. ISBN: 978-1-66548-832-7. DOI: 10.1109/ICMLC56445.2022.9941325. URL: https://ieeexplore.ieee.org/document/9941325/ (visited on 01/07/2025).

[13] Neil J. Cronin et al. "Feasibility of OpenPose markerless motion analysis in a real athletics competition". In: *Frontiers in Sports and Active Living* 5 (Jan. 5, 2024), p. 1298003. ISSN: 2624-9367. DOI: 10.3389/fspor.2023.1298003. URL: https://www.frontiersin.org/articles/10.3389/fspor.2023.1298003/full (visited on 01/23/2025).

[14] Mosam Dabhi et al. *MBW: Multi-view Bootstrapping in the Wild*. Version Number: 1. 2022. DOI: 10.48550/ARXIV.2210.01721. URL: https://arxiv.org/abs/2210.01721 (visited on 01/25/2025).

[15] Chinthaka Gamanayake et al. "Cluster Pruning: An Efficient Filter Pruning Method for Edge AI Vision Applications". In: *IEEE Journal of Selected Topics in Signal Processing* 14.4 (May 2020), pp. 802–816. ISSN: 1932-4553, 1941-0484. DOI: 10.1109/JSTSP.2020.2971418. URL: https://ieeexplore.ieee.org/document/8979425/ (visited on 01/07/2025).

[16] Lu Gan, Connor Lee, and Soon-Jo Chung. "Unsupervised RGB-to-Thermal Domain Adaptation via Multi-Domain Attention Network". In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. 2023 IEEE International Conference on Robotics and Automation (ICRA). London, United Kingdom: IEEE, May 29, 2023, pp. 6014–6020. ISBN: 9798350323658. DOI: 10.1109/ICRA48891.2023.10160872. URL: https://ieeexplore.ieee.org/document/10160872/ (visited on 01/25/2025).

[17] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. "Image Style Transfer Using Convolutional Neural Networks". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, June 2016, pp. 2414–2423. ISBN: 978-1-4673-8851-1. DOI: 10.1109/CVPR.2016.265. URL: http://ieeexplore.ieee.org/document/7780634/ (visited on 01/25/2025).

[18] Vladimir Geroimenko. *Augmented Reality and Artificial Intelligence: The Fusion of Advanced Technologies*. 1st ed. Springer Series on Cultural Computing Series. Cham: Springer, 2023. 1 p. ISBN: 978-3-031-27166-3.

[19] Fanbin Gu et al. "Automatic detection of abnormal hand gestures in patients with radial, ulnar, or median nerve injury using hand pose estimation". In: *Frontiers in Neurology* 13 (Dec. 7, 2022), p. 1052505. ISSN: 1664-2295. DOI: 10.3389/fneur.2022.1052505. URL: https://www.frontiersin.org/articles/10.3389/fneur.2022.1052505/full (visited on 01/07/2025).

[20] Zeyu Guan. "Real time object recognition based on YOLO model". In: *Theoretical and Natural Science* 28.1 (Dec. 26, 2023), pp. 137–143. ISSN: 2753-8818, 2753-8826. DOI: 10.54254/2753-8818/28/20230450. URL: https://www.ewadirect.com/proceedings/tns/article/view/9302 (visited on 01/07/2025).

[21] Abhishek Gupta et al. "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues". In: *Array* 10 (July 2021), p. 100057. ISSN: 25900056. DOI: 10.1016/j.array.2021.100057. URL: https://linkinghub.elsevier.com/retrieve/pii/S2590005621000059 (visited on 01/26/2025).

[22] Roxane Henry et al. "Comparison of the OpenPose system and the reference optoelectronic system for gait analysis of lower-limb angular parameters in children". In: *Orthopaedics & Traumatology: Surgery & Research* (Nov. 2024), p. 104044. ISSN: 18770568. DOI: 10.1016/j.otsr.2024.104044. URL: https://linkinghub.elsevier.com/retrieve/pii/S1877056824003426 (visited on 01/23/2025).

[23] Rudolf Hoffmann and Christoph Reich. "A Systematic Literature Review on Artificial Intelligence and Explainable Artificial Intelligence for Visual Quality Assurance in Manufacturing". In: *Electronics* 12.22 (Nov. 8, 2023), p. 4572. ISSN: 2079-9292. DOI: 10.3390/electronics12224572. URL: https://www.mdpi.com/2079-9292/12/22/4572 (visited on 01/26/2025).

[24] Yongcheng Jing et al. *Neural Style Transfer: A Review*. Version Number: 7. 2017. DOI: 10.48550/ARXIV.1705.04058. URL: https://arxiv.org/abs/1705.04058 (visited on 01/25/2025).

[25] Mate Kristo, Marina Ivasic-Kos, and Miran Pobar. "Thermal Object Detection in Difficult Weather Conditions Using YOLO". In: *IEEE Access* 8 (2020), pp. 125459–125476. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3007481. URL: https://ieeexplore.ieee.org/document/9133581/ (visited on 01/23/2025).

[26] Namhoon Lee, Thalaiyasingam Ajanthan, and Philip H. S. Torr. *SNIP: Single-shot Network Pruning based on Connection Sensitivity*. Version Number: 2. 2018. DOI: 10.48550/ARXIV.1810.02340. URL: https://arxiv.org/abs/1810.02340 (visited on 01/25/2025).

[27] Luc Lévesque. "Nyquist sampling theorem: understanding the illusion of a spinning wheel captured with a video camera". In: *Physics Education* 49.6 (Nov. 2014), pp. 697–705. ISSN: 0031-9120, 1361-6552. DOI: 10.1088/0031-9120/49/6/697. URL: https://iopscience.iop.org/article/10.1088/0031-9120/49/6/697 (visited on 01/18/2025).

[28] Tailin Liang et al. *Pruning and Quantization for Deep Neural Network Acceleration: A Survey*. Version Number: 3. 2021. DOI: 10.48550/ARXIV.2101.09671. URL: https://arxiv.org/abs/2101.09671 (visited on 01/07/2025).

[29] Hanxiao Liu, Karen Simonyan, and Yiming Yang. *DARTS: Differentiable Architecture Search*. Version Number: 2. 2018. DOI: 10.48550/ARXIV.1806.09055. URL: https://arxiv.org/abs/1806.09055 (visited on 01/24/2025).

[30] Camillo Lugaresi et al. *MediaPipe: A Framework for Building Perception Pipelines*. June 14, 2019. DOI: 10.48550/arXiv.1906.08172. arXiv: 1906.08172[cs]. URL: http://arxiv.org/abs/1906.08172 (visited on 11/26/2024).

[31] Farah Magrabi et al. "Artificial Intelligence in Clinical Decision Support: Challenges for Evaluating AI and Practical Implications: A Position Paper from the IMIA Technology Assessment & Quality Development in Health Informatics Working Group and the EFMI Working Group for Assessment of Health Information Systems". In: *Yearbook of Medical Informatics* 28.1 (Aug. 2019), pp. 128–134. ISSN: 0943-4747, 2364-0502. DOI: 10.1055/s-0039-1677903. URL: http://www.thieme-connect.de/DOI/DOI?10.1055/s-0039-1677903 (visited on 01/07/2025).

[32] Arnab Neelim Mazumder et al. "A Survey on the Optimization of Neural Network Accelerators for Micro-AI On-Device Inference". In: *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 11.4 (Dec. 2021), pp. 532–547. ISSN: 2156-3357, 2156-3365. DOI: 10.1109/JETCAS.2021.3129415. URL:

https://ieeexplore.ieee.org/document/9627710/ (visited on 01/07/2025).

[33] Markus Nagel et al. *A White Paper on Neural Network Quantization*. June 15, 2021. DOI: 10.48550/arXiv.2106.08295. arXiv: 2106.08295[cs]. URL: http://arxiv.org/abs/2106.08295 (visited on 01/24/2025).

[34] Enrico Nunzi, Cesare Massone, and Françoise Portaels, eds. *Leprosy and Buruli Ulcer: A Practical Guide*. Cham: Springer International Publishing, 2022. ISBN: 978-3-030-89703-1 978-3-030-89704-8. DOI: 10.1007/978-3-030-89704-8. URL: https://link.springer.com/10.1007/978-3-030-89704-8 (visited on 01/07/2025).

[35] Meghavi Rana and Megha Bhushan. "Machine learning and deep learning approach for medical image analysis: diagnosis to detection". In: *Multimedia Tools and Applications* 82.17 (July 2023), pp. 26731–26769. ISSN: 1380-7501, 1573-7721. DOI: 10.1007/s11042-022-14305-w. URL: https://link.springer.com/10.1007/s11042-022-14305-w (visited on 01/20/2025).

[36] Sandeep Reddy. "Evaluating large language models for use in healthcare: A framework for translational value assessment". In: *Informatics in Medicine Unlocked* 41 (2023), p. 101304. ISSN: 23529148. DOI: 10.1016/j.imu.2023.101304. URL: https://linkinghub.elsevier.com/retrieve/pii/S2352914823001508 (visited on 01/25/2025).

[37] Sandeep Reddy et al. "Evaluation framework to guide implementation of AI systems into healthcare settings". In: *BMJ Health & Care Informatics* 28.1 (Oct. 2021), e100444. ISSN: 2632-1009. DOI: 10.1136/bmjhci-2021-100444. URL: https://informatics.bmj.com/lookup/doi/10.1136/bmjhci-2021-100444 (visited on 01/25/2025).

[38] Joseph Redmon et al. *You Only Look Once: Unified, Real-Time Object Detection*. Version Number: 5. 2015. DOI: 10.48550/ARXIV.1506.02640. URL: https://arxiv.org/abs/1506.02640 (visited on 01/07/2025).

[39] Pengzhen Ren et al. "A Comprehensive Survey of Neural Architecture Search: Challenges and Solutions". In: *ACM Computing Surveys* 54.4 (May 31, 2022), pp. 1–34. ISSN: 0360-0300, 1557-7341. DOI: 10.1145/3447582. URL: https://dl.acm.org/doi/10.1145/3447582 (visited on 01/24/2025).

[40] Rutar, D., Cheke, L. G., Hernández-Orallo, J., Markelius, A., & Schellaert, W. "General Interaction Battery: Simple Object Navigation and Affordances (GIBSONA)". In: (2024).

[41] Guillermo Sánchez-Brizuela et al. "Lightweight real-time hand segmentation leveraging MediaPipe landmark detection". In: *Virtual Reality* 27.4 (Dec. 2023), pp. 3125–3132. ISSN: 1359-4338, 1434-9957. DOI: 10.1007/s10055-023-00858-0. URL: https://link.springer.com/10.1007/s10055-023-00858-0 (visited on 01/07/2025).

[42] Schemkes, I. A. S. "Semi-Automatic temperature analysis based on Real-Time hand landmark tracking in infrared videos". In: (2024).

[43] Tomas Simon et al. "Hand Keypoint Detection in Single Images Using Multiview Bootstrapping". In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI: IEEE, July 2017, pp. 4645–4653. ISBN: 978-1-5386-0457-1. DOI: 10.1109/CVPR.2017.494. URL: http://ieeexplore.ieee.org/document/8099977/ (visited on 11/26/2024).

[44] Sreehari Sreenath et al. "Monocular Tracking of Human Hand on a Smart Phone Camera using MediaPipe and its Application in Robotics". In: *2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC)*. 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC). Bangalore, India: IEEE, Sept. 30, 2021, pp. 1–6. ISBN: 978-1-66543-240-5. DOI: 10.1109/R10-HTC53172.2021.9641542. URL: https://ieeexplore.ieee.org/document/9641542/ (visited on 01/07/2025).

[45] Pierre Stock et al. *And the Bit Goes Down: Revisiting the Quantization of Neural Networks*. Nov. 9, 2020. DOI: 10.48550/arXiv.1907.05686. arXiv: 1907.05686[cs]. URL: http://arxiv.org/abs/1907.05686 (visited on 01/24/2025).

[46] Teja Reddy Gatla. "COMPARATIVE EVALUATION OF AI MODELS FOR PREDICTING STROKE RISK USING GENETIC AND LIFESTYLE FACTORS". In: *International Journal of Innovations in Engineering Research and Technology* 11.5 (May 31, 2024), pp. 37–49. ISSN: 2394-3696. DOI: 10.26662/ijiert.v11i5.pp37-49. URL: https://repo.ijiert.org/index.php/ijiert/article/view/3946 (visited on 01/25/2025).

[47] Ultralytics. *Ultralytics YOLO docs*. URL: https://docs.ultralytics.com/.

[48] Thomas E. Vanhecke. "Zotero". In: *Journal of the Medical Library Association : JMLA* 96.3 (July 2008), pp. 275–276. ISSN: 1536-5050, 1558-9439. DOI: 10.3163/1536-5050.96.3.022. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2479046/ (visited on 01/26/2025).

[49] Ajantha Vijayakumar and Subramaniyaswamy Vairavasundaram. "YOLO-based Object Detection Models: A Review and its Applications". In: *Multimedia Tools and Applications* 83.35 (Mar. 14, 2024), pp. 83535–83574. ISSN: 1573-7721. DOI: 10.1007/s11042-024-18872-y. URL: https://link.springer.com/10.1007/s11042-024-18872-y (visited on 01/20/2025).

[50] Jianping Wang et al. "Open Pose Mask R-CNN Network for Individual Cattle Recognition". In: *IEEE Access* 11 (2023), pp. 113752–113768. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2023.3321152. URL: https://ieeexplore.ieee.org/document/10268406/ (visited on 01/23/2025).

[51] World Health Organization. *Guidelines for the diagnosis, treatment and prevention of leprosy*. 2018.

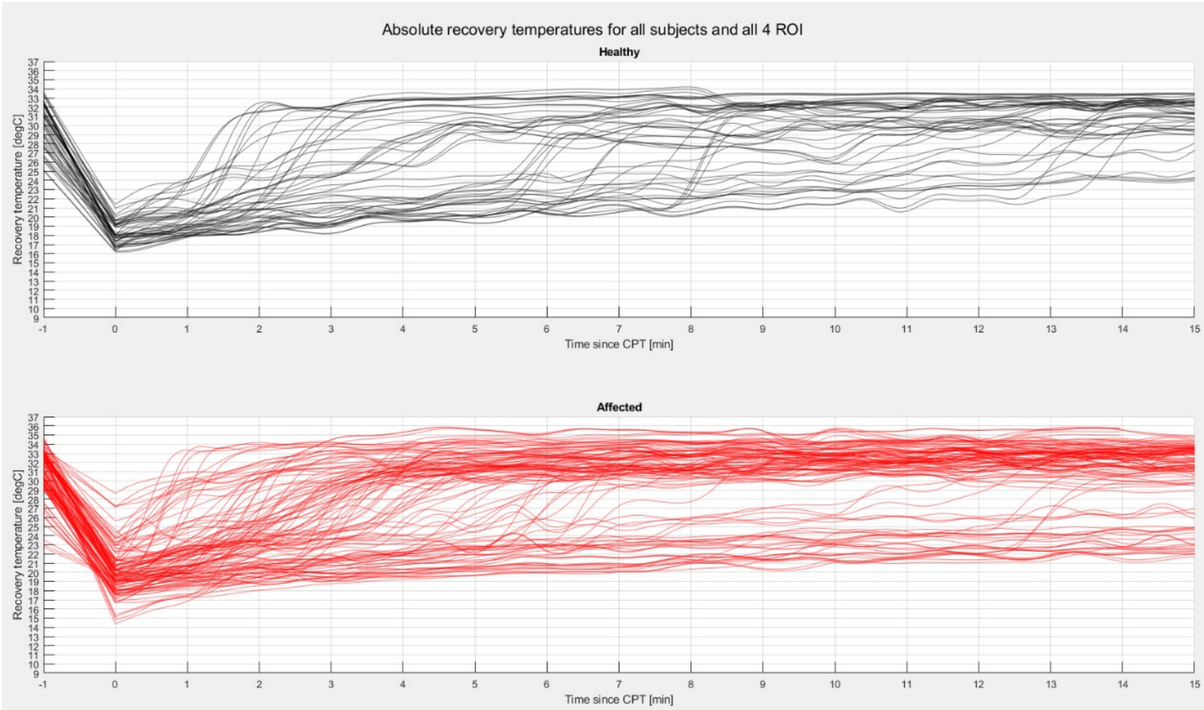
[52] World Health Organization. *Towards zero leprosy*. 2021.

[53] Jiaxiang Wu et al. *Quantized Convolutional Neural Networks for Mobile Devices*. Version Number: 3. 2015. DOI: 10.48550/ARXIV.1512.06473. URL: https://arxiv.org/abs/1512.06473 (visited on 01/24/2025).

[54] Haichuan Yang, Yuhao Zhu, and Ji Liu. *Energy-Constrained Compression for Deep Neural Networks via Weighted Sparse Projection and Layer Input Masking*. Version Number: 3. 2018. DOI: 10.48550/ARXIV.1806.04321. URL: https://arxiv.org/abs/1806.04321 (visited on 01/25/2025).

[55] Seul-Ki Yeom et al. “Pruning by explaining: A novel criterion for deep neural network pruning”. In: *Pattern Recognition* 115 (July 2021), p. 107899. ISSN: 00313203. DOI: 10.1016/j.patcog.2021.107899. URL: https://linkinghub.elsevier.com/retrieve/pii/S0031320321000868 (visited on 01/25/2025).

Appendix A



Recovery temperatures of healthy and affected patients. t = -1 is the baseline temperature, t = 0 is the temperature right after cooling. Significant thermal responses occur over periods of 30–40 seconds (A. Knulst, personal communication, January 12, 2025)

Appendix B

The following search keywords, grouped by topic, have been used in the literature study:

Topic	Search keywords
Real-time visual AI on infrared video	hand landmark detection, visual AI, infrared video analysis, real-time AI applications, MedidaPipe, OpenPose, YOLO multiview bootstrapping
Optimizing AI models for mobile devices	AI model downsizing techniques, quantization, pruning, AI optimization for constrained hardware. visual AI on mobiel devices weight pruning, activation pruning, filter pruning, neural architecture search Nyquist-Shannon sampling theorem, real-time visual AI, post-hoc processing
Performance metrics and evaluation	AI performance evaluation, metrics for real-time AI, mobile AI benchmarks Performance-Oriented Evaluation, Capability-Oriented Evaluation, medical AI evaluation, TEHAI framework