Research Plan for CSE3000 Research Project

Vladimir Sachkov

November 17, 2024

1 Background of the research

Current visual AI for hand landmark prediction is trained on normal color images (RGB). We, however, wish to make predictions in the infrared domain. One way to handle such a domain shift is by "domain adaptation," and here we will explore its suitability. Irene Schemkes has also conducted work on this topic in her master's thesis, focusing on diagnosing leprosy with hand landmark tracking Schemkes, 2024.

2 Background of the research

Early diagnosis of leprosy, is essential to prevent transmission and permanent nerve and tissue damage. Accurate temperature measurement in regions of interest (ROIs) on the hands is a promising diagnostic method, as current techniques often detect the disease only in later stages when physical symptoms become apparent.

Manual temperature analysis involves time-consuming annotation of ROIs and subsequent measurements, leading to potential human error and inconsistencies.

Automated temperature measurement using artificial intelligence (AI) addresses these challenges by enhancing the speed and accuracy of ROI (Region of interest) detection and temperature analysis. AI-driven solutions ensure consistent results across large datasets, minimize human error, and employ real-time monitoring. Irene Schemkes conducted similar work in her master's thesis, developing a semi-automatic temperature analysis method utilizing real-time hand landmark tracking in infrared videos Schemkes, 2024. Irene used Google's MediaPipe Hand Landmarker detector to track hand landmarks in infrared videos AI, 2024, but the problem is that the model was trained on RGB images, and not infrared. Although she has used some adaptation techniques, such as adjusting sensitivity of landmark detection thresholds, or applying filters and transformations that highlight temperature variations, I belieb there is still a big room for improvement in domain adaptation part of the problem, that I am going to address and investigate in my research.

3 Research Question

Main Question: How effective is domain adaptation in enhancing the performance of visual hand landmark prediction AI when transitioning from RGB to infrared (IR) imaging

1

domains?

Addressing the domain shift from RGB to IR is crucial for leveraging AI's capabilities in medical diagnostics in general. This research aims to evaluate its suitability and effective-ness in this context. I will use Irene's work domain adaptation techniques as a baseline to compare the performance of the proposed methods. This project's peer Zofia is going to be answering a similarally stated research question, but instead of domain adaptation tecniques she is going to be using style transfer to convert IR videos to RGB, so in the later stages of the project I am planning to conduct a joint investigation on comparing the two methods, their resulting accuracy, resource usage, scalability, and speed. It's also important to note, that main objective of the problem is to optimize accuracy of predicted regions of interest on hands from the entire video, and not just single frames. For example in Irene's work, for example, Coordinate Filtering Techniques are used in the context of the entire video. But in my work I will mostly focus on single frames, and try to improve the accuracy of landmark detection. Another problem is the absence of the ground-truth landmark points in existing dataset. The established area of work totally fits the time-frame for the research project, since it utilizes established domain adaptation techniques and uses existing tools like MediaPipe and machine learning libraries like TensorFlow and PyTorch.

Sub-Questions:

1. **What are the limitations of Irene's domain adaptation techniques under current computational power constraints?**

   • *Objective Criteria:* Analyze performance bottlenecks, resource usage, and scal- ability issues in her approach.

2. **What key identifiers will describe model's ability to perform in our given scenario in Nepal?**

   • *Objective Criteria:* Discovered set of key identifiers should be able to properly compare multiple pipelines for IR to RGB conversition. For example the style transfer solution that will be produced by my peer in a later stages. Key identifiers might include: resources and time needed to retran, or fine-tune a model, inference resource usage and time, accuracy, generalizability, and scalability.

3. **How can hand landmark detection accuracy be effectively evaluated us-ing improved domain adaptation methods, considering per-frame and temporal metrics?**

   • *Objective Criteria:* Compare per-frame and temporal accuracy metrics, and identify any additional metrics required beyond those used in Irene's work.

   - Cross-validation using expert-annotated subset of frames
   - Temporal consistency metrics across video sequences
   - relative positions and distances between landmarks

- Stability analysis across different thermal conditions and hand positions

4. **Will the low-cost methods (Feature Alignment Using Maximum Mean Discrepance, Fine-Tuning with Pseudo-Labeling, or Adaptive Batch Nor-**

2

malization) be able to improve the performance of hand landmark pre-diction in infrared videos compared to Irene's baseline?

- *Objective Criteria:* The low-cost methods implemnted only using compute re-sources available online will provide better results than Irene's work, given only infrared videos dataset.

5. **Will the medium-cost method (Domain-Adversarial Neural Networks domain adaptation method, Correlation Alignment, or Self-Supervised Learning with Pretext Tasks) produce better and faster inference results than Irene's baseline or cheaper methods?**

- *Objective Criteria:* Implement and assess specific domain adaptation tech- niques to enhance model performance on infrared (IR) data using comparative analysis with metrics such as accuracy, precision, recall, and F1-score.

6. **Will the domain adaptation method result in a model that maintains rea-sonable inference time for real-time hand landmark prediction in infrared videos?**

- *Objective Criteria:* The adapted model's inference time and resource utiliza-tion should not exceed predefined thresholds that ensure real-time processing capabilities. Irene's method averaged 42 minutes per video processing time.

7. **What are the main bottlenecks in landmark detection on IR videos with Google's MediaPipe Hand Landmarker?** • *Objective Criteria:* Is the palm detection or the landmark detection more problematic? Determine accuracies by isolating and combining these components on labeled data.

4 Method

To tackle the research aims, I'll break down the method into smaller, doable tasks, tackling each sub-question one by one in the order they're listed.

1. Setting Up the Development Environment First, I need to figure out the best development environment for this project. I'll evaluate different machine learning libraries to ensure they have all the necessary data processing methods and advanced mathematical models. Libraries like TensorFlow, PyTorch, and scikit-learn will be considered based on their capabilities and compatibility with our goals.

2. Organizing and Cleaning Data Next, I'll organize and clean the existing infrared video data. This involves manually labeling the points of interest on the hands to create accurate datasets. Additionally, I'll add metadata to videos and frames, such as relative hand positions and other relevant information, to enrich the dataset.

3. Feature Extraction with Unsupervised Learning I'll run some unsupervised ma-chine learning methods to identify unique features in the videos. Techniques like clustering and dimensionality reduction will help in uncovering patterns and features that are not immediately apparent.

4. Developing a Custom Evaluation Framework Implementing a custom evaluation framework is crucial for testing solutions. This framework will assess accuracy, perform cross-validation, and more, addressing the second sub-question. To ensure robustness, I might include testing on custom videos captured with our own camera to check if the models generalize well.

5. Implementing Irene's Methods I'll start by implementing Irene's domain adapta-tion methods, including Google's MediaPipe Hand Landmarker. Running these methods

3

through my evaluation pipeline will allow me to compare the performance metrics with those reported in Irene's paper.

6. Identifying and Addressing Limitations After evaluating Irene's methods, I'll pin-point their limitations. Based on these insights, I'll begin implementing low-cost domain adaptation techniques specifically designed to overcome these issues. If certain challenges require more advanced solutions, such as deep learning or fine-tuning, I'll seek access to a computation cluster from Delft.

7. Experimenting with Various Domain Adaptation Techniques I'll explore a range of domain adaptation methods, from cheap options like Feature Alignment using Maximum Mean Discrepancy and Fine-Tuning with Pseudo-Labeling to medium-cost methods such as Domain-Adversarial Neural Networks and Self-Supervised Learning with Pretext Tasks. Each method will be evaluated based on accuracy, hardware resource usage, and processing time.

8. Collaboration and Optimization Throughout the project, I'll collaborate with peers experienced in machine learning to troubleshoot and optimize the models. Using version control tools like Git will help manage the codebase efficiently and track changes.

9. Validation and Documentation Each task's completion will be validated by com-paring the performance metrics against Irene's baseline using scikit-learn. Finally, I'll document the entire process, ensuring that each step meets the predefined objectives and preparing the results for further analysis. Once having propotypes of my domain adap-tation methods, I will be ready to compare their performance with Zofia's style transfer technique.

5 Planning of the Research Project

References

AI, G. (2024). Mediapipe: Hand landmarker [Accessed: 2024-05-21].

Schemkes, I. (2024). *Diagnosing leprosy with hand landmark tracking* [Master's thesis, Delft University of Technology] [Unpublished master's thesis].

4

Date	Activity	Type
11 Nov 2024 - 17 Nov 2024	Conduct systematic literature review and develop evaluation framework	Personal Work
11 Nov 2024	Lecture - Information Literacy Assignment	Lecture
11 Nov 2024	Lecture - Kick-Off	Lecture
18 Nov 2024 - 24 Nov 2024	Finalize systematic literature review and complete evaluation framework	Personal Work
12 Nov 2024	Student Deadline - First Week Plan	Student Deadline
12 Nov 2024	Supervisor - Meeting	Meeting
15 Nov 2024	Lecture - ACS Assignment 1	Lecture
21 Nov 2024	Research plan meeting	Meeting
22 Nov 2024	Responsible Professor - Research plan presentation	Meeting
25 Nov 2024	Lecture - Session Responsible Research	Lecture
25 Nov 2024	Lecture - Session ACS (paper)	Lecture
25 Nov 2024 - 1 Dec 2024	Develop code for testing and comparing models and prepare clean annotated data	Personal Work
29 Nov 2024	Student Deadline - Set Date for Final Presentation	Student Deadline
29 Nov 2024	Supervisor - Meeting	Meeting
29 Nov 2024	Responsible Professor - Meeting	Meeting
29 Nov 2024	Examiner - Meeting	Meeting
2 Dec 2024 - 8 Dec 2024	Implement and evaluate cheaper methods (and Irene's domain adaptation techniques), and assess feasibility of medium resource demand methods	Personal Work
2 Dec 2024	Lecture - Session Responsible Research	Lecture
6 Dec 2024	Lecture - ACS Assignments 2a and 2b	Lecture
9 Dec 2024	Lecture - Session ACS (poster)	Lecture
10 Dec 2024	Midterm meeting	Meeting
11 Dec 2024	Student Deadline - Midterm Presentations	Student Deadline

Date	**Activity**	**Type**
13 Dec 2024	Lecture - ACS Assignment 3	Lecture
13 Dec 2024 - 15 Dec 2024	Write introduction and methodology sections; prepare rough outline for midterm presentation	Personal Work
16 Dec 2024	Lecture - Session ACS (paper)	Lecture
16 Dec 2024 - 22 Dec 2024	Make custom model functional before Midterm Presentation	Personal Work
23 Dec 2024	Lecture - Coaching ACS	Lecture
24 Dec 2024	Student Deadline - Paper Draft v1	Student Deadline
25 Dec 2024	Conduct experiments and fix bugs in time for Paper Draft v1	Personal Work
26 Dec 2024	Student Deadline - Peer Review Draft v1	Student Deadline
30 Dec 2024	Lecture - Coaching ACS	Lecture
30 Dec 2024 - 5 Jan 2025	Rewrite paper using peer review feedback; amend potential shortcomings	Personal Work
1 Jan 2025	Student Deadline - Paper Draft v2	Student Deadline
6 Jan 2025	Lecture - Coaching ACS	Lecture
6 Jan 2025 - 12 Jan 2025	Ensure proper styling, citation, and minor details for Paper Draft v2; prepare rough draft of the poster	Personal Work
7 Jan 2025	Responsible Professor - Feedback v2	Meeting
8 Jan 2025	Student Deadline - Submit Final Paper	Student Deadline
13 Jan 2025	Lecture - Session ACS (poster)	Lecture
13 Jan 2025 - 17 Jan 2025	Prepare rough draft of the poster and prepare for the presentation	Personal Work
13 Jan 2025	Student Deadline - Submit Poster	Student Deadline
17 Jan 2025	Student Deadline - Poster Presentations	Student Deadline

Table 1: Project Timeline with Personal Work Deadlines