## Objective

This document evaluates three relevant research papers and proposes the optimal approach for implementation.

Overview

**Research Paper #1:** <https://www.mdpi.com/1424-8220/20/2/376>

**Title:** Trajectory-Based Air Writing Recognition Using Deep Neural Network and Depth Sensor

**Key Takeaways:**

* **Setup and Data Collection:** Used depth sensors for capturing spatial trajectory sequences.
* **Preprocessing Techniques:**
  + **Nearest Neighbor Normalization:** Smoothed trajectories by adjusting six neighboring points, reducing noise effectively. This was more accurate than the root point normalization.
  + **Root Point Normalization:** Standardized trajectories to a common origin for consistency.
* **Modeling Approach:**
  + **LSTM for Digit Recognition:** Captured sequential dependencies in trajectories and was specifically used for recognizing digits written in the air. Compared to CNN, LSTM excelled at capturing long-term dependencies inherent in sequential data.
  + **CNN:** Extracted spatial features but proved less effective for digit recognition due to its inability to fully model temporal relationships in trajectory data.
* **Findings:** LSTM performed better due to its ability to handle long-term dependencies in trajectory data.

**Relevance to Our Project:**

* The trajectory smoothing and LSTM-based modeling approaches can be adapted for LED gloves, leveraging the sequential nature of movement data. LSTM’s success in digit recognition can guide our approach to text inference.

**Research Paper #2:** <https://dl.acm.org/doi/abs/10.1145/3474085.3475694>

**Title:** Air-Text: Air-Writing and Recognition System

**Github Repo:** <https://github.com/sklee2014/Air-Text?utm_source=chatgpt.com>

**Key Takeaways:**

* **Modules:**
  + **Air-Writing Module:** Tracks fingertip locations from RGB video input and converts trajectories into binary images.
  + **Text-Recognition Module:** Utilizes a pre-trained model (TPS-ResNet-BiLSTM-Attn) for text prediction.
* **Modeling Details:**
  + **TPS-ResNet-BiLSTM-Attn:** Combines Thin Plate Spline (TPS) for alignment, ResNet for feature extraction, BiLSTM for sequence modeling, and Attention for accurate text recognition.

**Relevance to Our Project:**

* The system’s fingertip tracking can be adapted to LED brightness localization.
* The TPS-ResNet-BiLSTM-Attn model provides a strong foundation for trajectory-to-text conversion.

**Research Paper #3:** <https://ieeexplore.ieee.org/abstract/document/9257775>

**Title:** Air-Writing Recognition using Deep Convolutional and Recurrent Neural Network Architectures

**Github Repo:** <https://github.com/kosmasK/air-writing-recognition?utm_source=chatgpt.com>

**Key Takeaways:**

* **Objective:** This study focuses on recognizing air-written digits (0-9) using hand movements tracked in 3D space by a Leap Motion Controller (LMC). Both time-series (dynamic) and image-based (static) approaches were evaluated.
* **Dynamic Models:**
  + **LSTM:** Performed best for recognizing the sequential patterns in air-written digits, with high accuracy.
  + **BLSTM:** Slightly less accurate than LSTM.
  + **CNN-LSTM:** Combined spatial feature extraction with sequential modeling but was less effective than LSTM alone.
  + **TCN:** Fast and efficient but less accurate than LSTM.
* **Static Models:**
  + **CNN:** Processed binary images of digit trails with good accuracy but was less effective than LSTM for sequential data.
  + **TCN-Static:** Performed the worst due to loss of spatial details during conversion.
* **Combined Model:** Combining LSTM and CNN didn’t improve performance and slightly reduced accuracy compared to LSTM alone.

**Key Findings:**

* LSTM excelled at handling time-series data, making it the best choice for air-writing recognition.
* Static models like CNN worked well for image-based data but were less effective for sequential gestures.

**Relevance to Our Project:**

* This paper highlights that LSTM is ideal for recognizing the sequential patterns in air-written digits, which directly applies to tracking LED glove trajectories.
* The findings reinforce the importance of using dynamic models like LSTM for accurate and efficient recognition of air-writing gestures.
* Insights from combining spatial and sequential features can help refine our approach for trajectory-to-text conversion.

## Proposed Approach for Implementation

* **Preprocessing:**
  + Apply Nearest Neighbor Normalization to smooth trajectories.
* **Modeling Approach:**
  + **Trajectory Tracking:** Implement fingertip tracking using YOLO for real-time LED localization.
  + **Text Recognition:** Utilize a combination of LSTM for sequential digit recognition and TPS-ResNet-BiLSTM-Attn for more complex word inference.