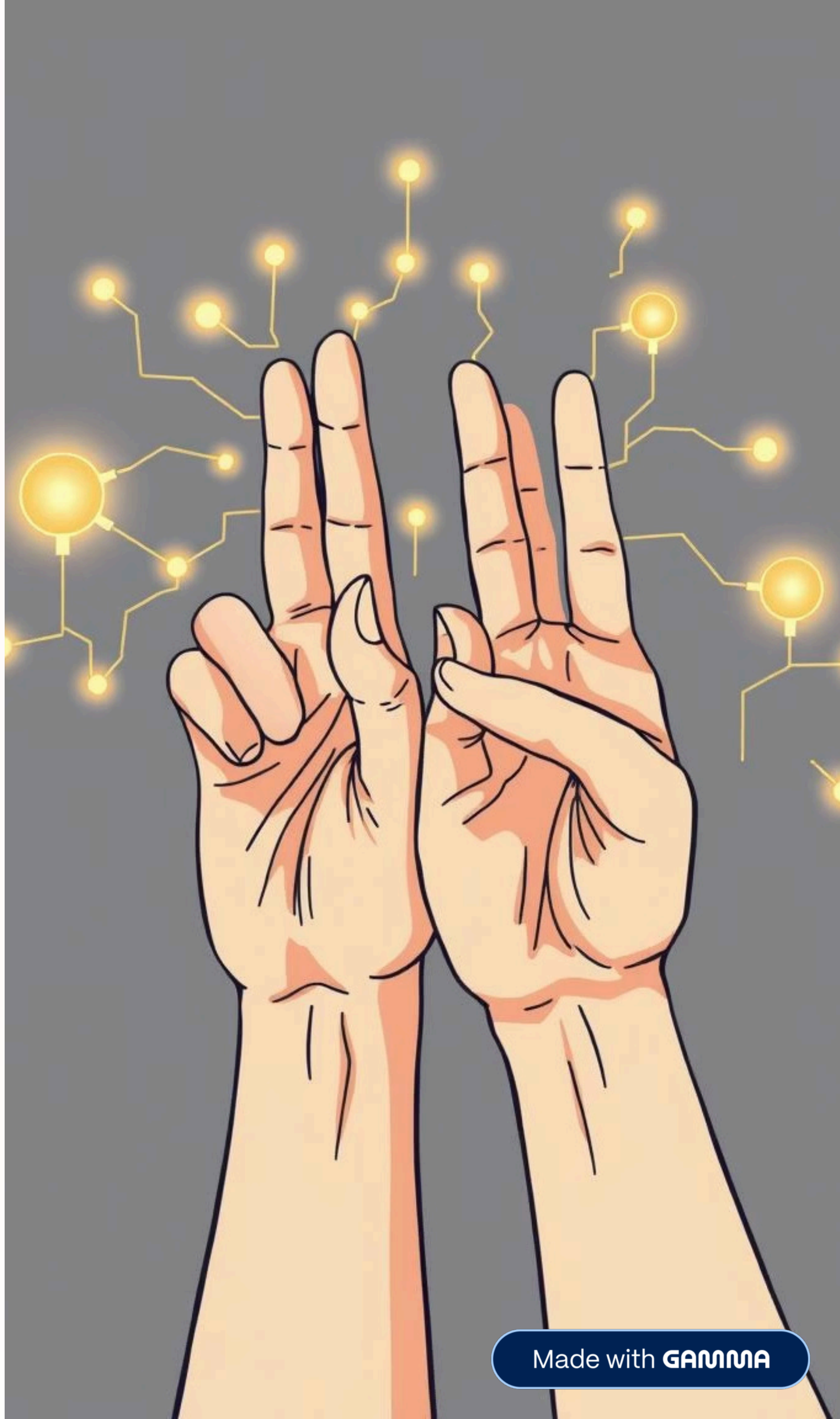
Sign Language Recognition Using MNIST Dataset

Unlocking communication through AI-powered gesture recognition

Presented by [KHUSHBAKHT JILANI]

Date: August 2025



Introduction: Bridging Communication Gaps

1

The Vital Role of Sign Language

Sign Language is a primary mode of communication for millions globally, particularly those with hearing impairments. Its importance in daily life, education, and social interaction cannot be overstated.

2

The Accessibility Challenge

A significant barrier exists due to the limited number of people who understand sign language, leading to communication gaps and reduced accessibility for the deaf community. This creates isolation and hinders inclusion.

3

Our AI-Driven Objective

Our goal is to leverage advanced machine learning techniques to automatically recognise hand gestures used in sign language. This aims to create a universal translator, breaking down communication barriers.

4

Leveraging MNIST for Benchmarking

We utilise MNIST-based datasets, renowned for providing a standardised, accessible benchmark for image classification. This allows for rigorous testing and comparison of our model's performance.

The Problem: Why Sign Language Recognition?

Costly Human Interpreters

Manual interpretation relies on skilled human translators, who are often scarce and expensive. This limits immediate and widespread access to sign language communication in everyday scenarios.

Enabling Inclusive Communication

Real-time automated recognition can foster inclusive communication across vital sectors like education, healthcare, and digital platforms. Imagine seamless interactions in a world without communication barriers.

Complexity of Hand Gestures

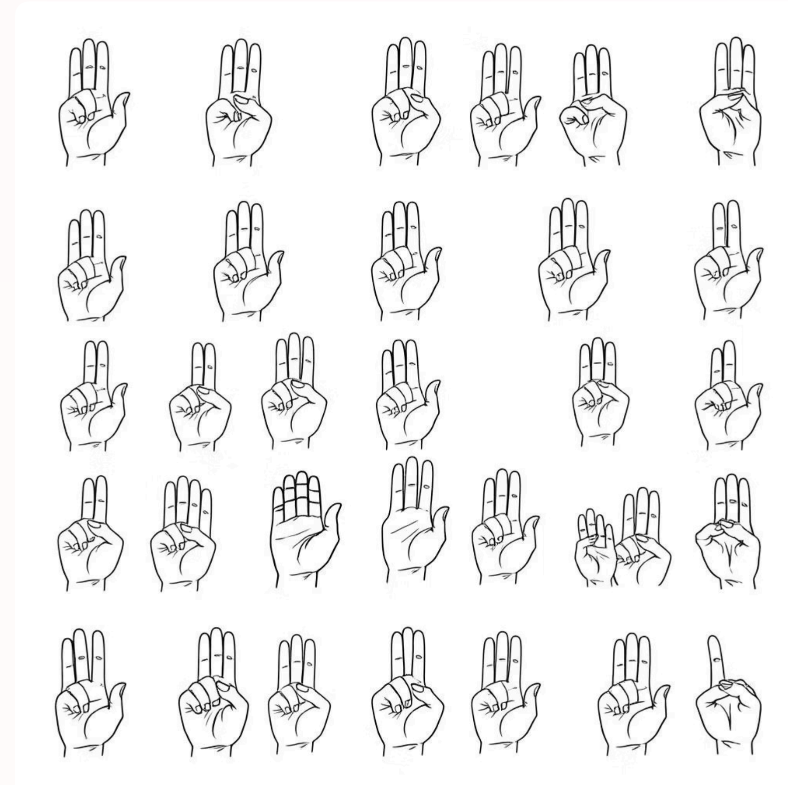
Sign language involves intricate hand shapes and subtle movements, making accurate recognition a significant technical challenge. Distinguishing between similar signs requires highly sophisticated models.

Need for Robust Models

There's a critical need for highly accurate and robust models that can generalise effectively across diverse users, varying environments, and different lighting conditions, ensuring reliability in real-world use.

Dataset Overview: Sign Language MNIST

- **Image Count:** Contains **27,455 grayscale 28x28 pixel images** of hand signs.
- **Content:** Represents **24 ASL letters** (A-Y, excluding J and Z, which are motion-based).
- **Preprocessing:** Images are preprocessed to focus solely on the hand region, aligning with the format of the classic MNIST digit dataset.
- **Split:** Approximately **27,455 images** for training and **7,172 images** for testing.
- **Balance:** Classes are well-balanced, ensuring fair representation for each sign, apart from the intentionally excluded letters.
- **Source:** Publicly available and widely used benchmark dataset for ASL recognition.



> "The Sign Language MNIST dataset provides a clean, preprocessed foundation for static sign recognition, making it an excellent starting point for deep learning models."

Methodology: Convolutional Neural Networks (CNN)

CNNs for Image Recognition

Convolutional Neural Networks (CNNs) are chosen for their exceptional ability to learn spatial hierarchies and patterns in image data, making them ideal for gesture recognition.

Core Architecture

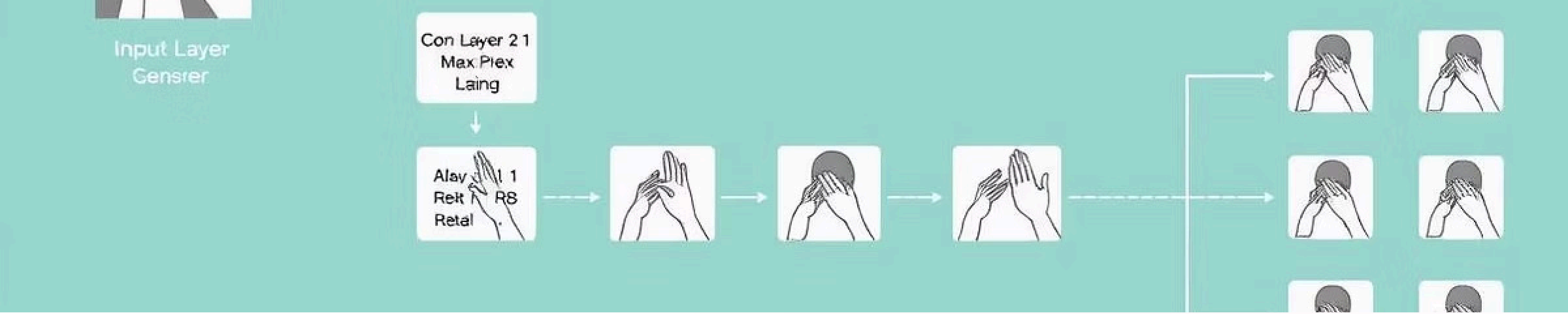
The model features multiple convolutional layers for feature extraction (edges, shapes, textures), pooling layers for dimensionality reduction, and fully connected layers for final classification into 24 ASL classes.

Training Stability & Prevention of Overfitting

Batch normalization layers enhance training stability and speed, while dropout layers are crucial for preventing overfitting, ensuring the model generalises well to unseen data.

Optimisation & Augmentation

The model is trained using TensorFlow/Keras, employing categorical cross-entropy loss and the Adam optimizer for efficient learning. Data augmentation (rotations, noise) is applied to improve robustness and generalisability.



Visual Diagram: CNN Architecture for ASL Recognition

This diagram illustrates the sequential flow of data through our CNN model, from raw image input to the final classification of a sign language gesture.

Results: Model Performance

~95%

Test Accuracy

Achieved on the Sign Language MNIST dataset, demonstrating high effectiveness in recognising static hand signs.

High

Precision & Recall

Consistently strong across most classes, indicating excellent ability to correctly identify and distinguish between different signs.

Minor

Misclassifications

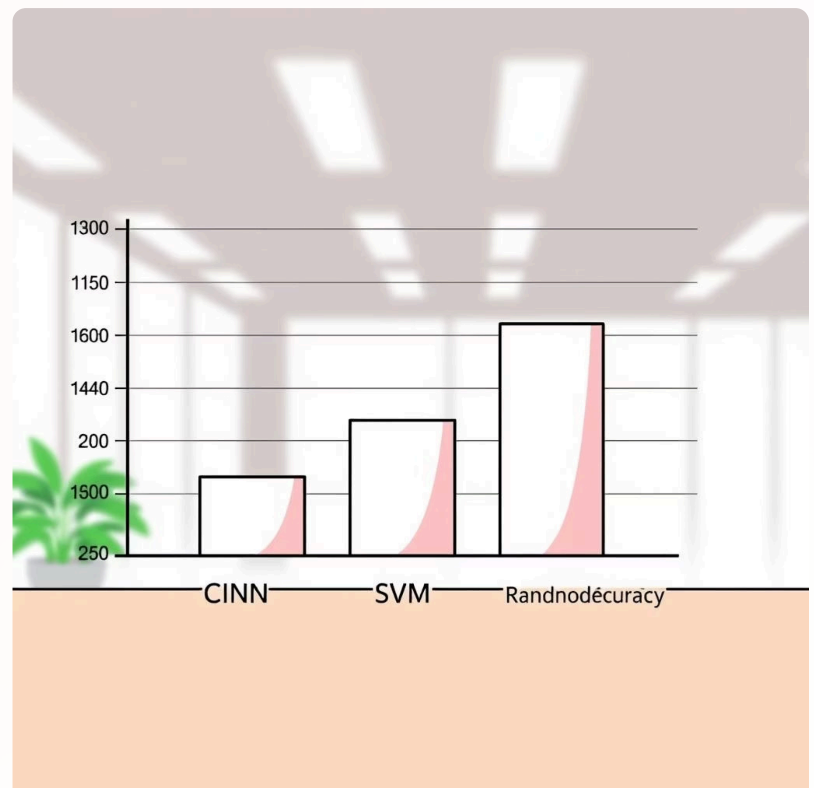
Primarily observed between visually similar signs, as shown in the confusion matrix, highlighting areas for future refinement.

Real-time Prediction

The model supports real-time prediction, with a possible demo using webcam input integrated via OpenCV, showcasing its practical utility.

Outperformance

Our CNN model significantly outperforms traditional machine learning classifiers, such as Support Vector Machines (SVM) and Random Forests, in both accuracy and robustness.



Applications: Real-World Impact

Assistive Technology

Empowering deaf and hard-of-hearing individuals with seamless daily communication, providing independence and ease of interaction.

Video Conferencing

Integrating live captioning and interaction capabilities into virtual meetings, ensuring full participation for all attendees.

Educational Tools

Providing interactive platforms for learning and practicing sign language, making it more accessible to a wider audience.

Smart Home Control

Enabling intuitive control of smart devices and appliances through gestures, enhancing accessibility for disabled users.

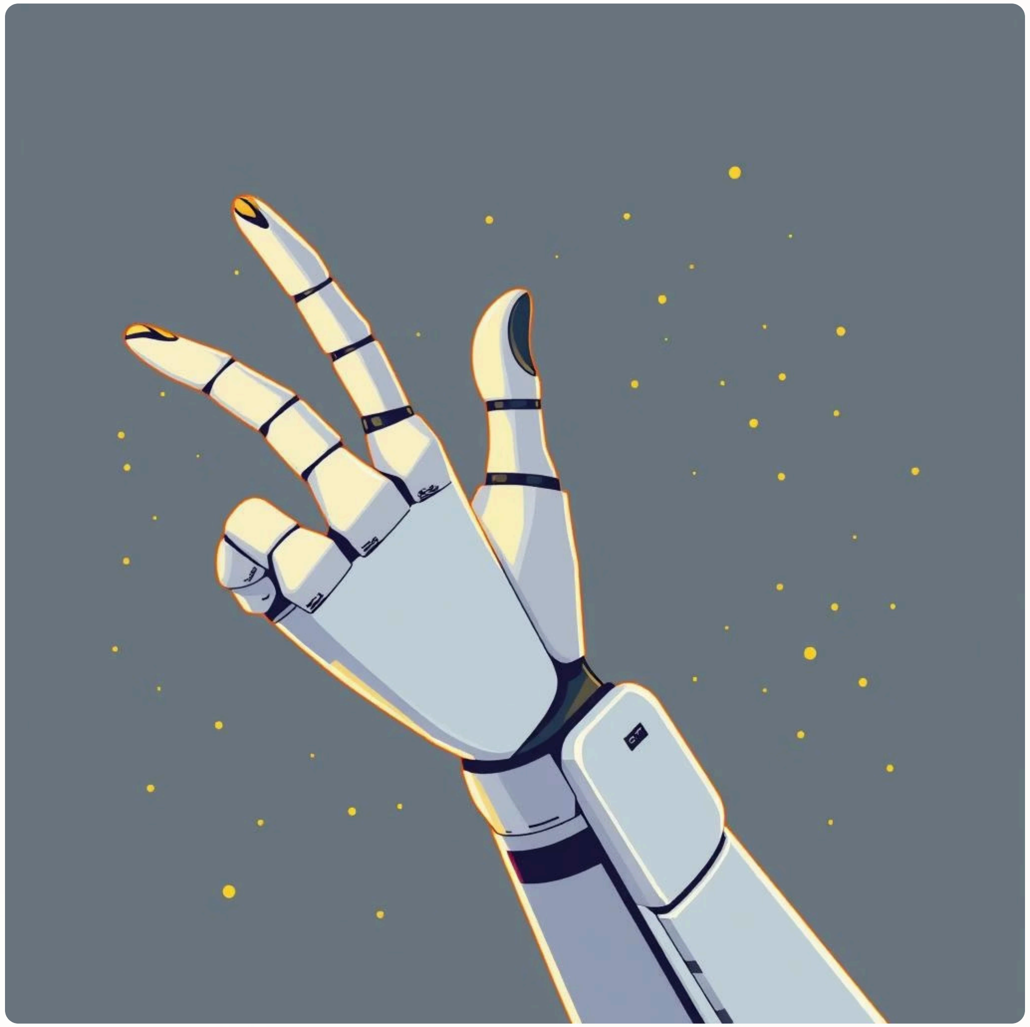
Dynamic Gesture Foundation

Laying the groundwork for future expansion into more complex, dynamic gesture recognition, moving beyond static signs.

Future Scope & Challenges

Expanding Dataset & Dynamic Gestures

The next step involves significantly extending the dataset to include dynamic gestures (e.g., J and Z) and continuous sign language sequences, capturing the fluidity of real-world communication.



Robustness & Environment Adaptability

Improving the model's resilience to diverse backgrounds, varying lighting conditions, and different hand shapes is crucial for practical deployment in varied real-world settings.

Multimodal Data Integration

Incorporating additional data streams, such as hand pose estimation, depth sensor data, and facial expressions, will provide a richer context and significantly enhance recognition accuracy.

Deployment on Edge Devices

Developing lightweight models optimised for mobile and embedded devices will enable widespread adoption and real-time processing without relying on cloud infrastructure.

Community Collaboration

Crucial collaboration with linguists and the deaf community will ensure the development of culturally accurate and user-centric recognition systems that truly meet their needs.

Conclusion & Call to Action

Promising Accessibility

Sign Language Recognition using the MNIST dataset showcases the profound potential of AI in driving accessibility and fostering inclusivity.

High Accuracy

CNN-based models deliver high accuracy, making practical applications a reality today, bridging crucial communication gaps.

Future Inclusivity

Continued advancements promise a future where communication is effortless and inclusive for everyone, worldwide.

Let's innovate together to break communication barriers and empower all voices.

Thank you!

Questions & Discussion