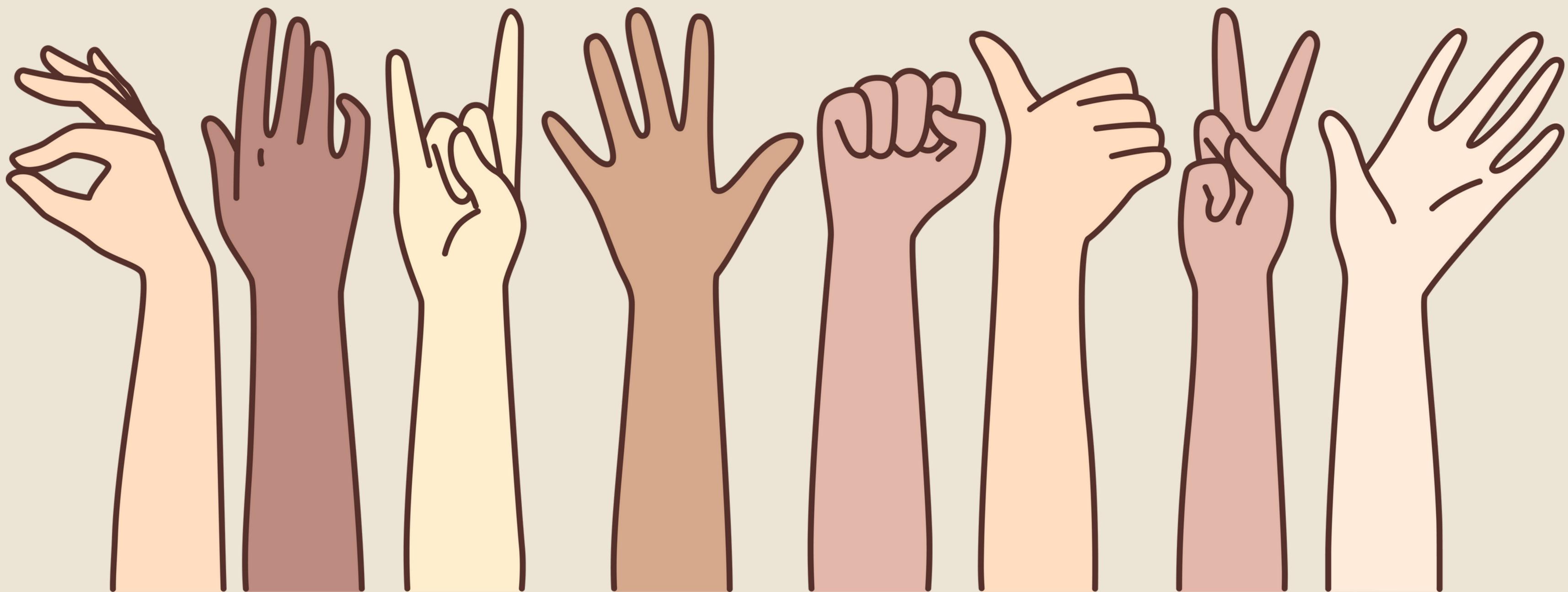


Hello!



Only 0.8% of the world's population uses sign language.



If more people knew sign language...





If only there
were some
way to
translate sign
language in
real time...



Bridging Silence

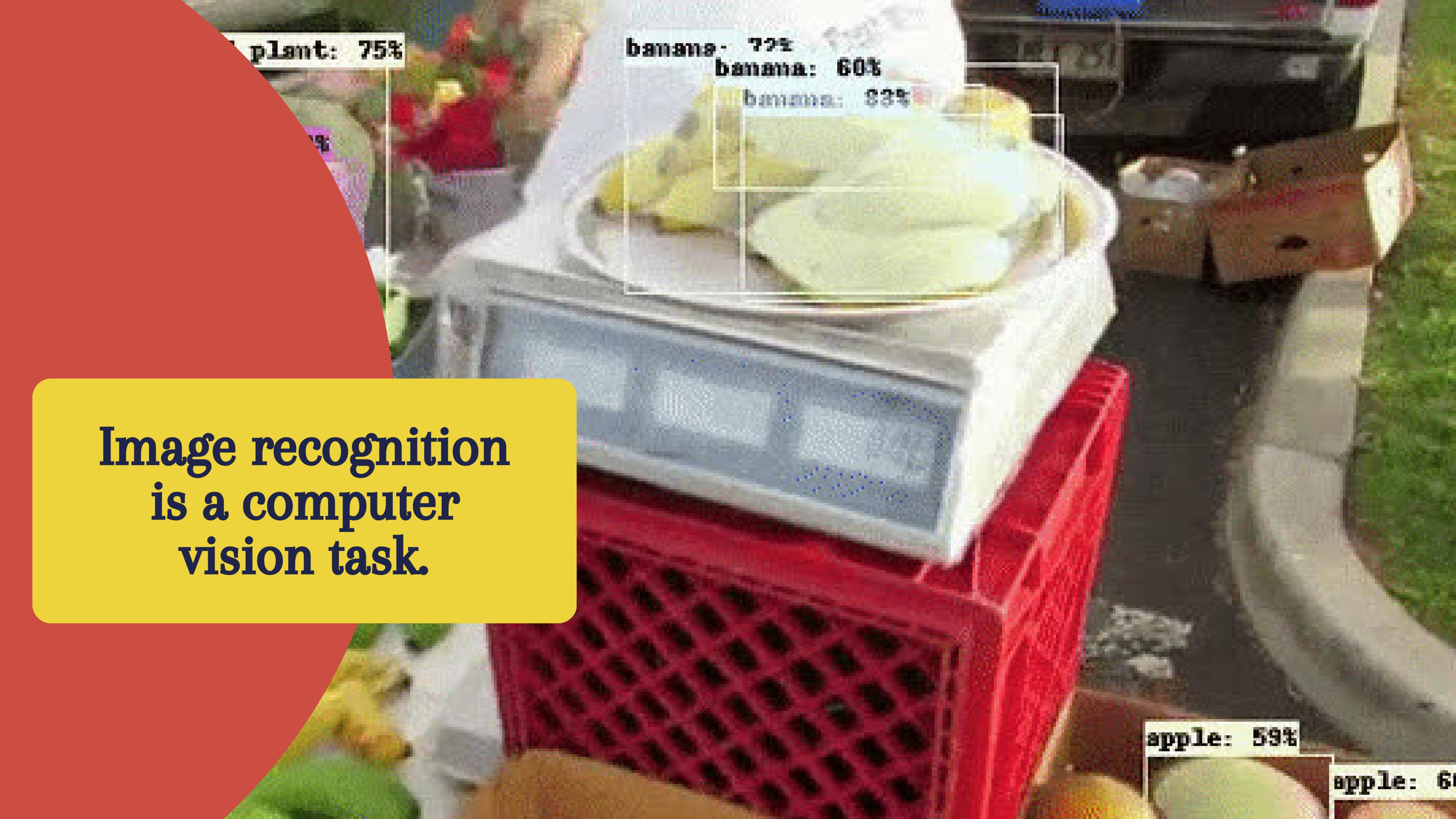


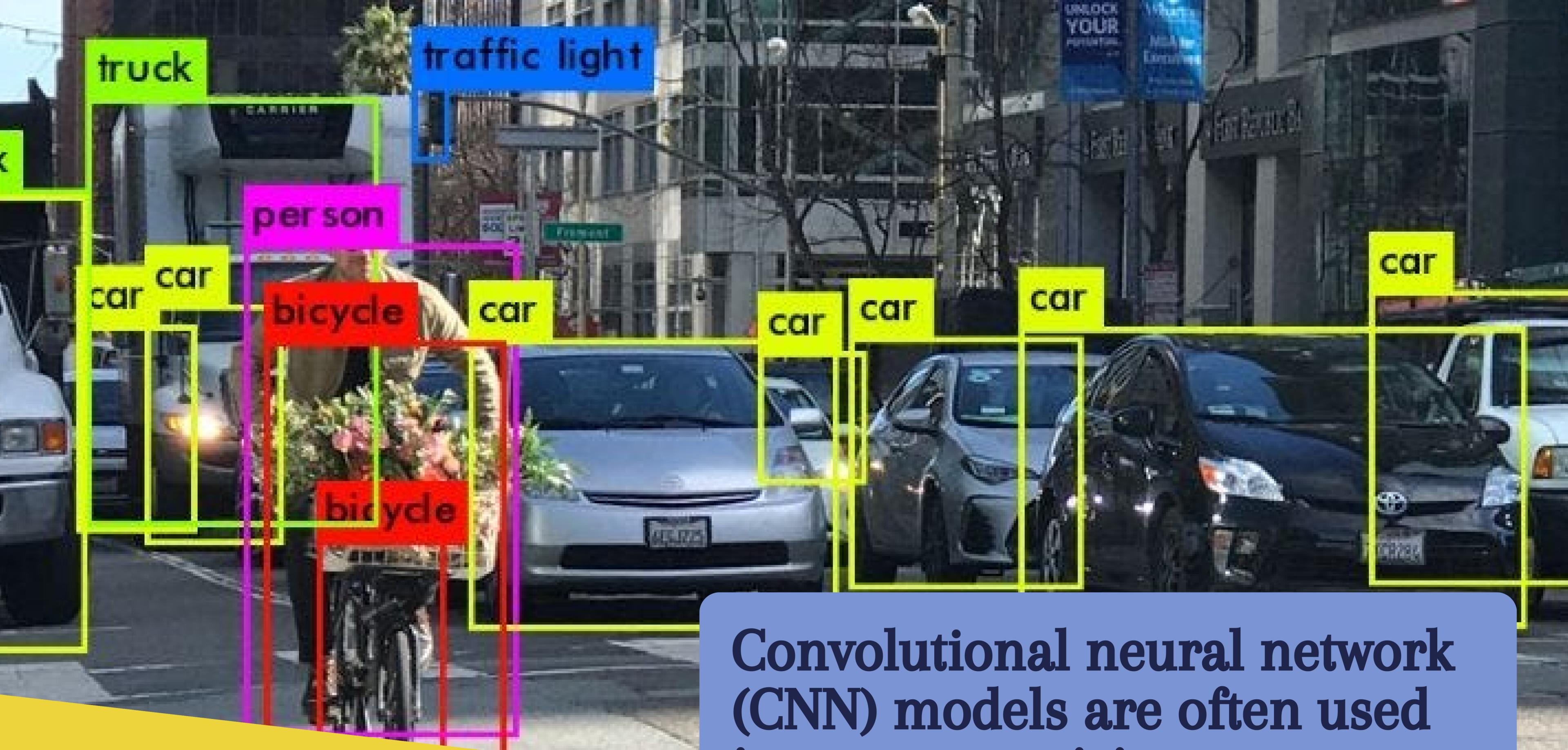
AI-Powered American
Sign Language Detection

Bulatao, Pati, Servanez, Rian



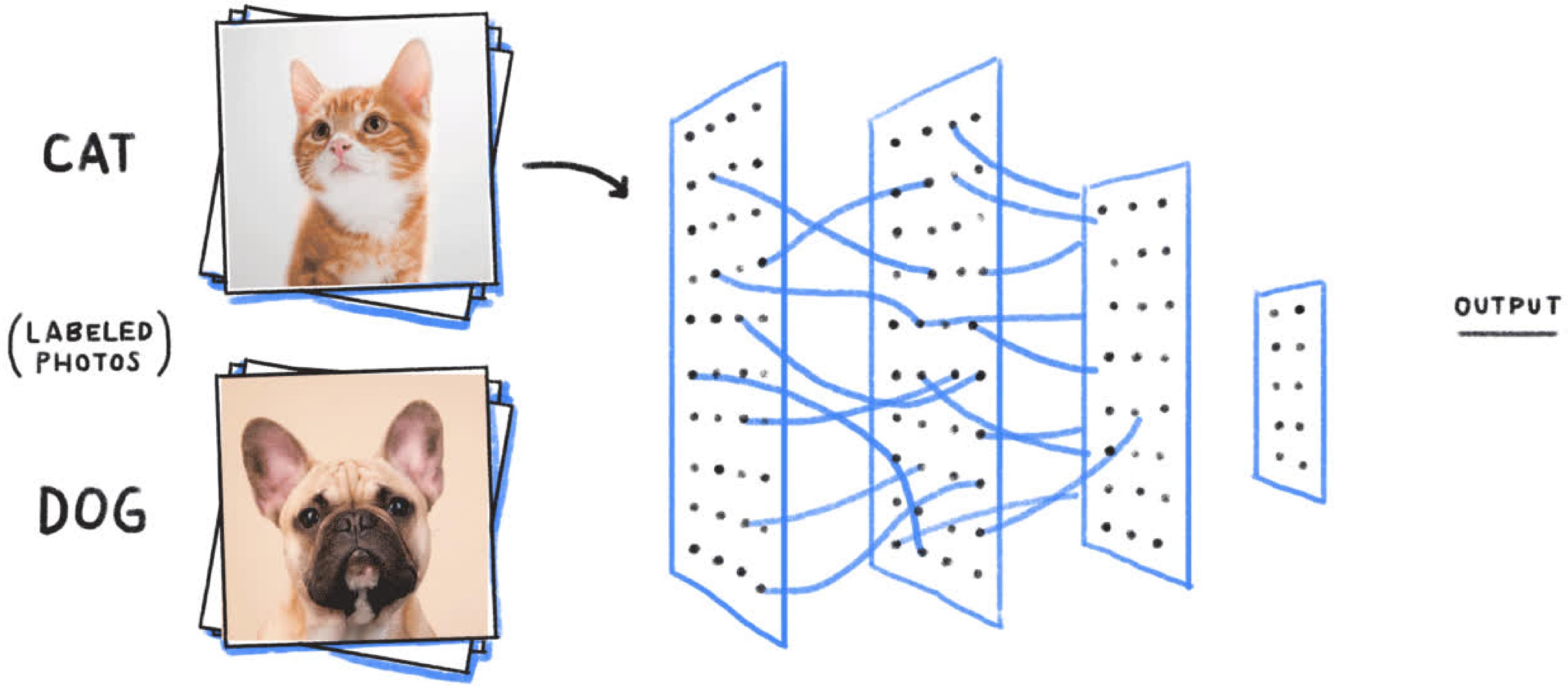
Image recognition
is a computer
vision task.





Convolutional neural network
(CNN) models are often used
in computer vision.

How do Convolutional Neural Networks (CNN) work?



Objectives



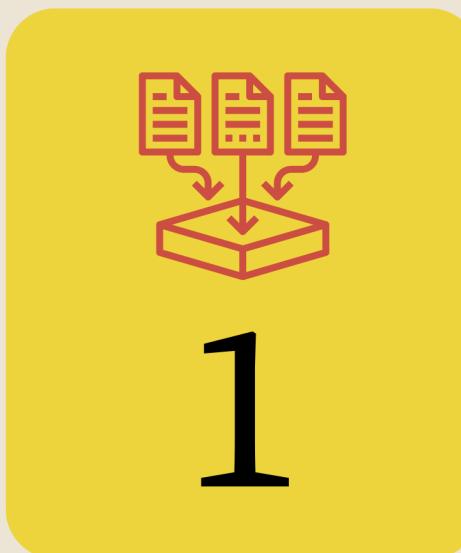
1

Create a CNN model that can accurately recognize letters of the American Sign Language (ASL) alphabet and translate them into English text.

2

Implement and optimize the model to function in real-time on various devices, such as smartphones, tablets, and computers.

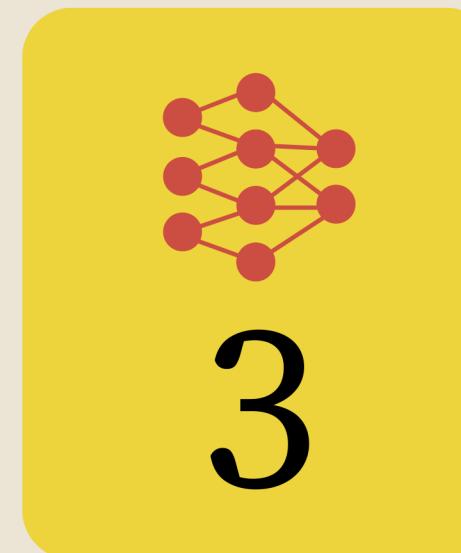
Methodology



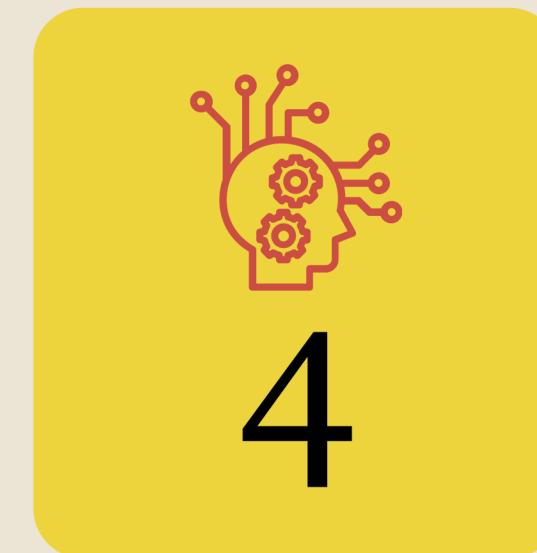
1
Load the Data



2
Preprocessing



3
CNN
Architecture



4
Fit and Train
Model



5
Store the Model



6
Live Prediction



7
Results

Data Description

Sign Language MNIST

- format is patterned closely after the original MNIST dataset, comprised of handwritten digits from 0 to 9
- images of ASL signs from A-Z (excluding J and Z)
- training = 27,455 cases; test = 7172 cases

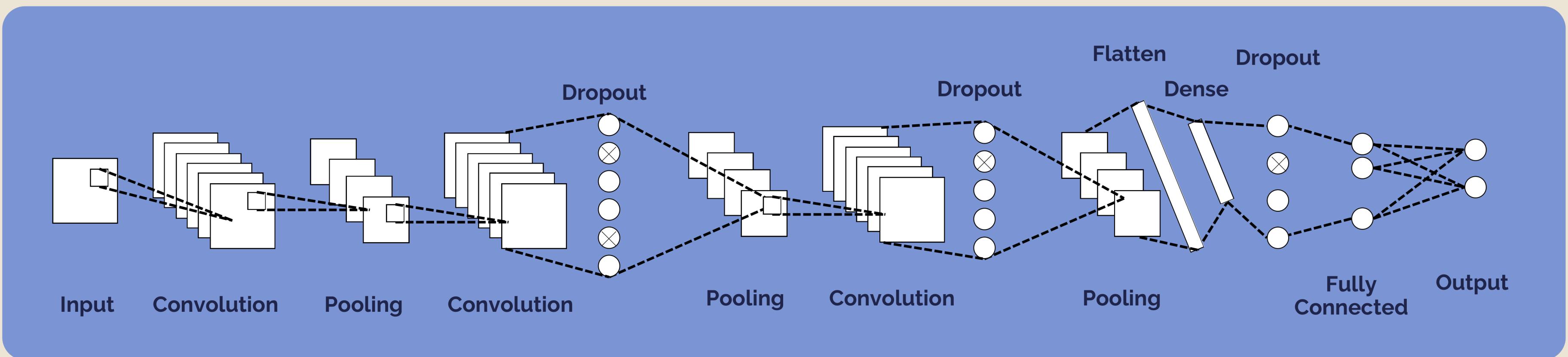


Model Architecture

```
model = Sequential([
    Conv2D(32, (3, 3), activation='relu', input_shape=(28, 28, 1), padding='same'),
    MaxPooling2D((2, 2)),
    Conv2D(64, (3, 3), activation='relu', padding='same'),
    MaxPooling2D((2, 2)),
    Conv2D(128, (3, 3), activation='relu', padding='same'),
    MaxPooling2D((2, 2)),
    Flatten(),
    Dense(512, activation='relu'),
    Dropout(0.5),
    Dense(num_classes, activation='softmax')
])
```



Model Architecture



Model Performance - Baseline

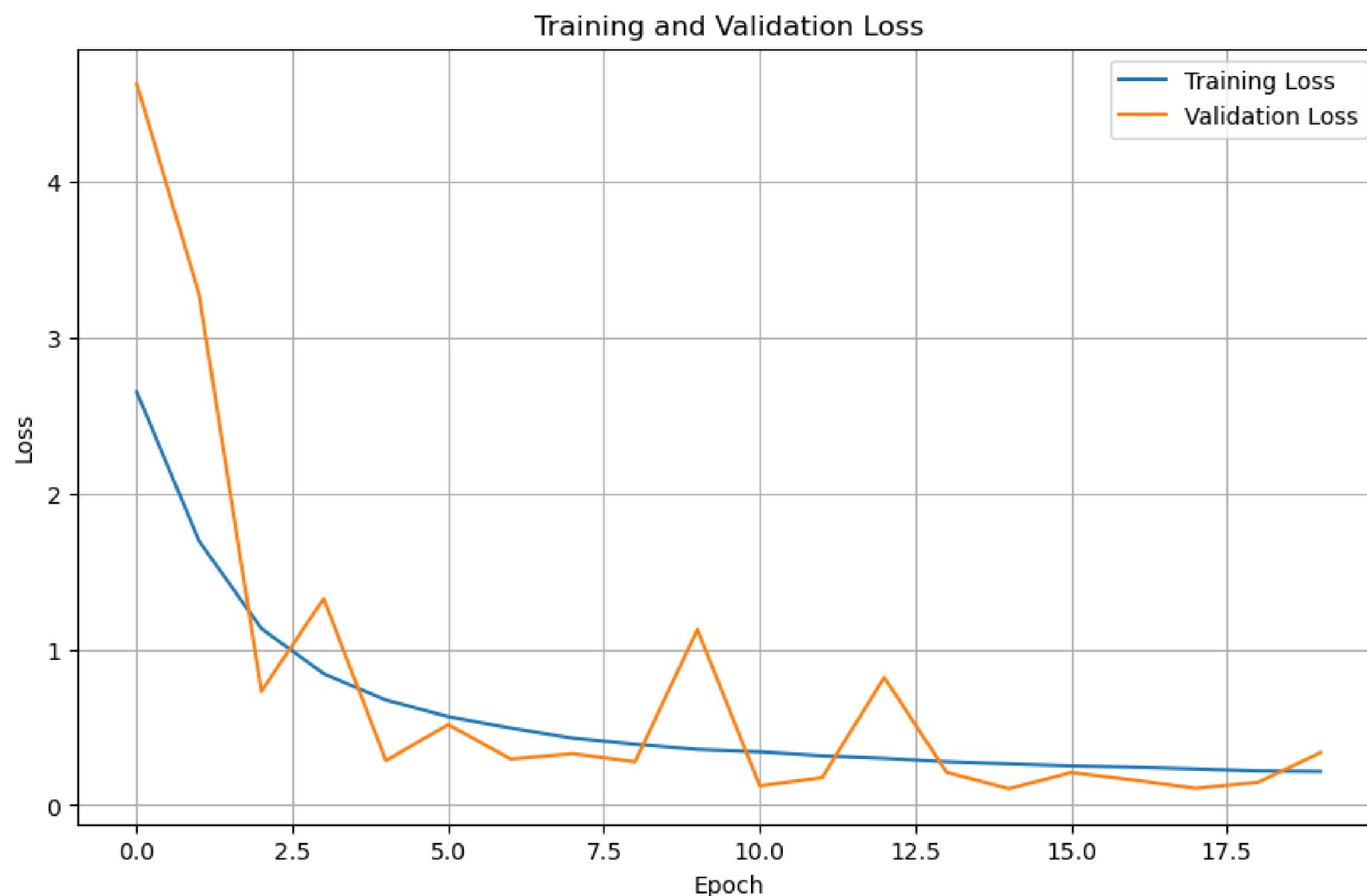
Baseline Accuracy - Kaggle 97%

Training Accuracy 95%

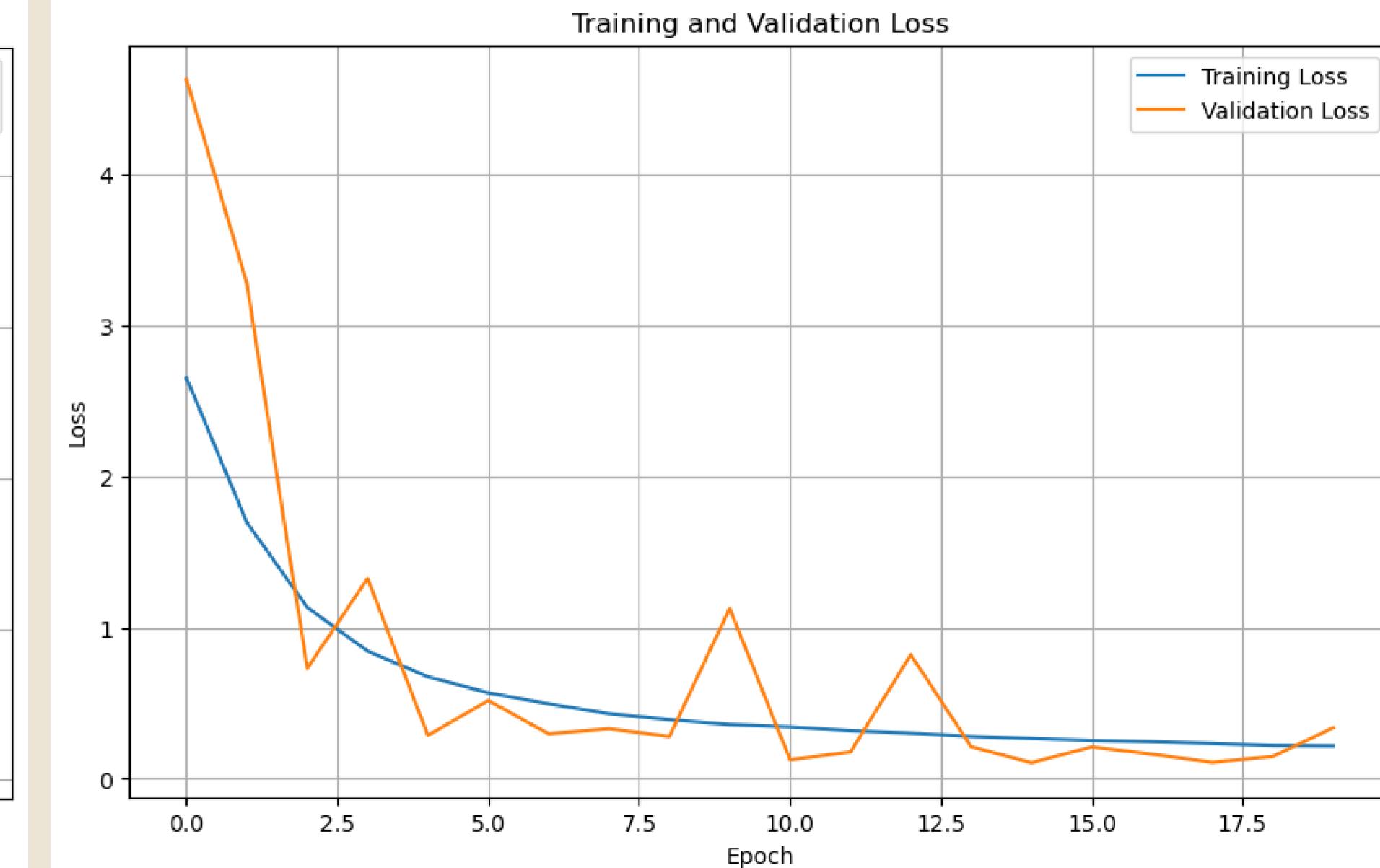
Validation Accuracy 94%



Model Performance - Baseline



Accuracy over Epochs



Loss over Epochs

HOLD ON!

Cost-sensitive learning

Model Performance - Improved

Baseline Accuracy - Kaggle 97%

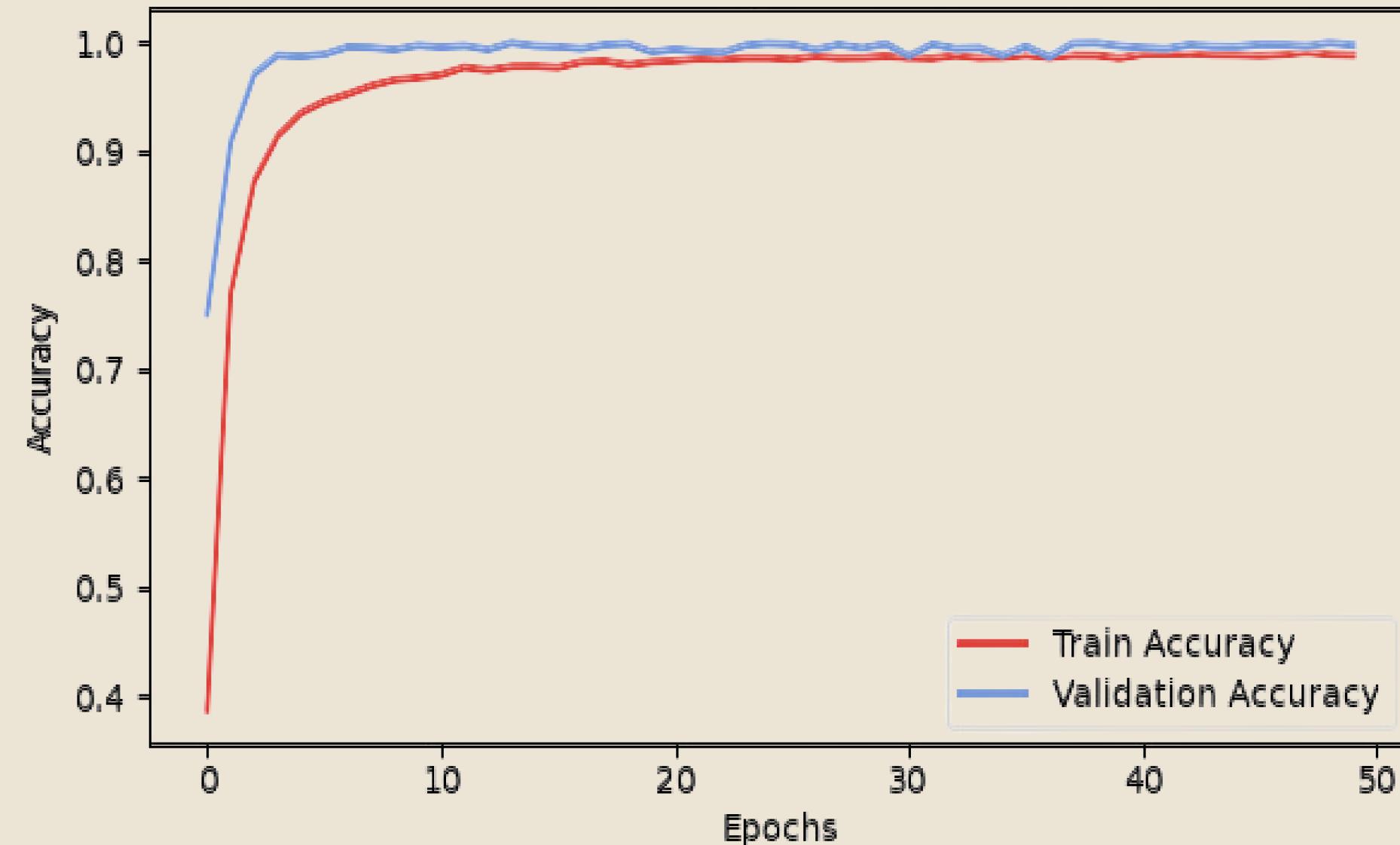
Training Accuracy 99.78%

Validation Accuracy 98.7%

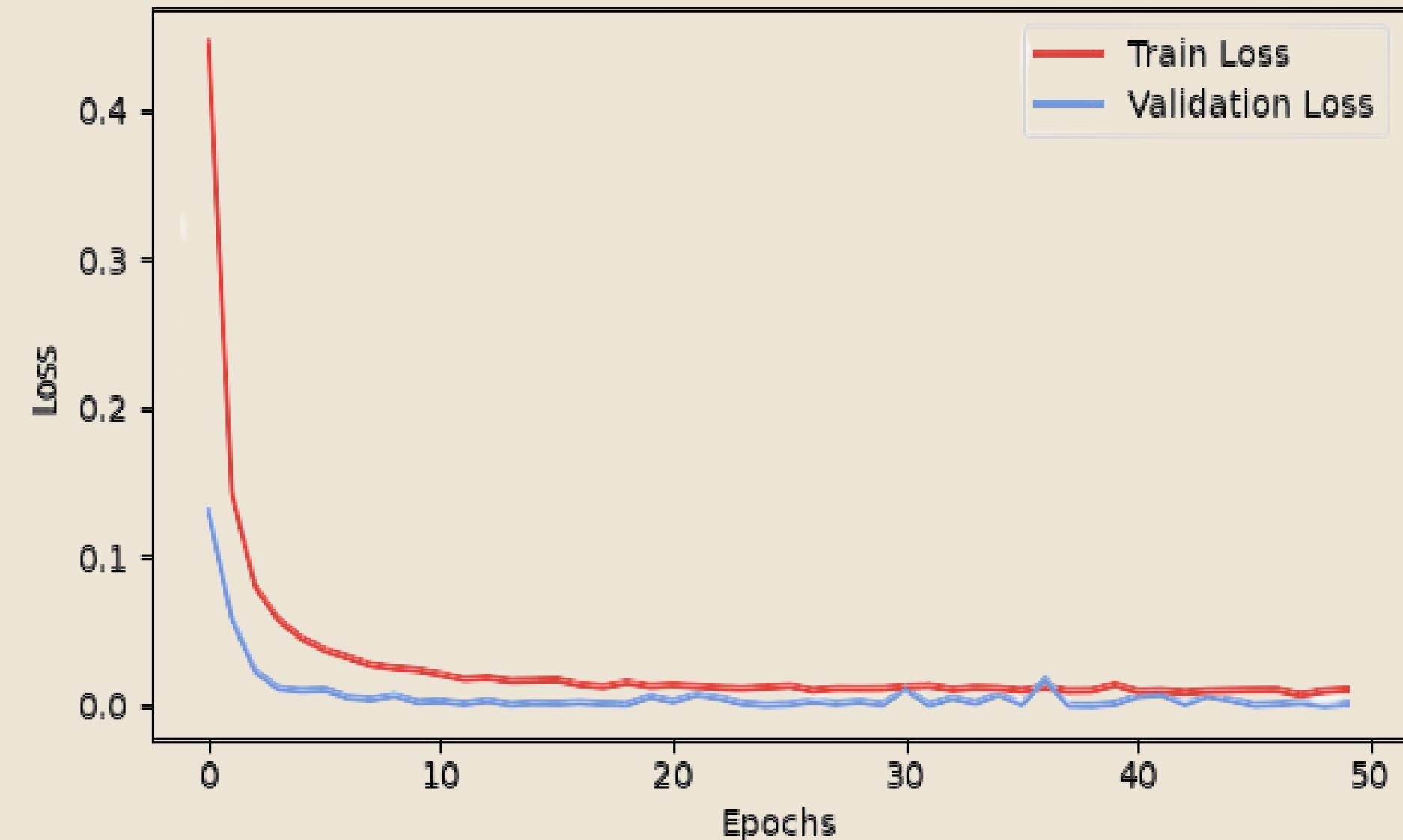


Model Performance - Improved

Accuracy Evolution



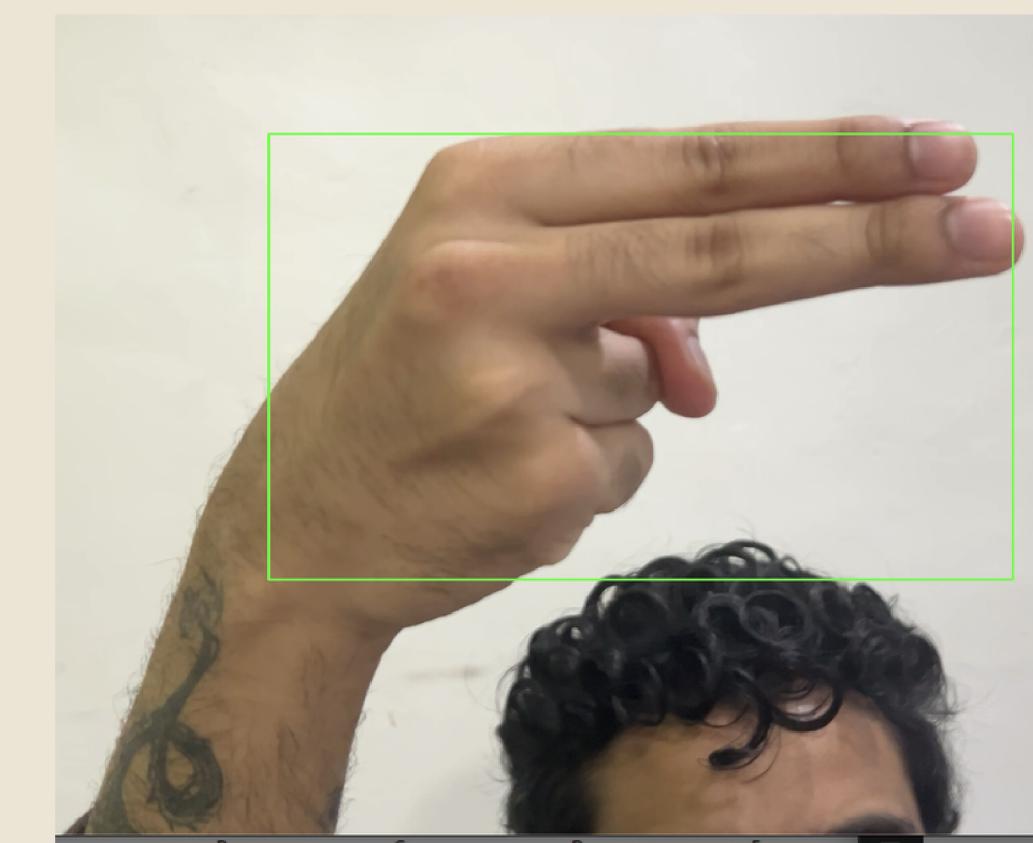
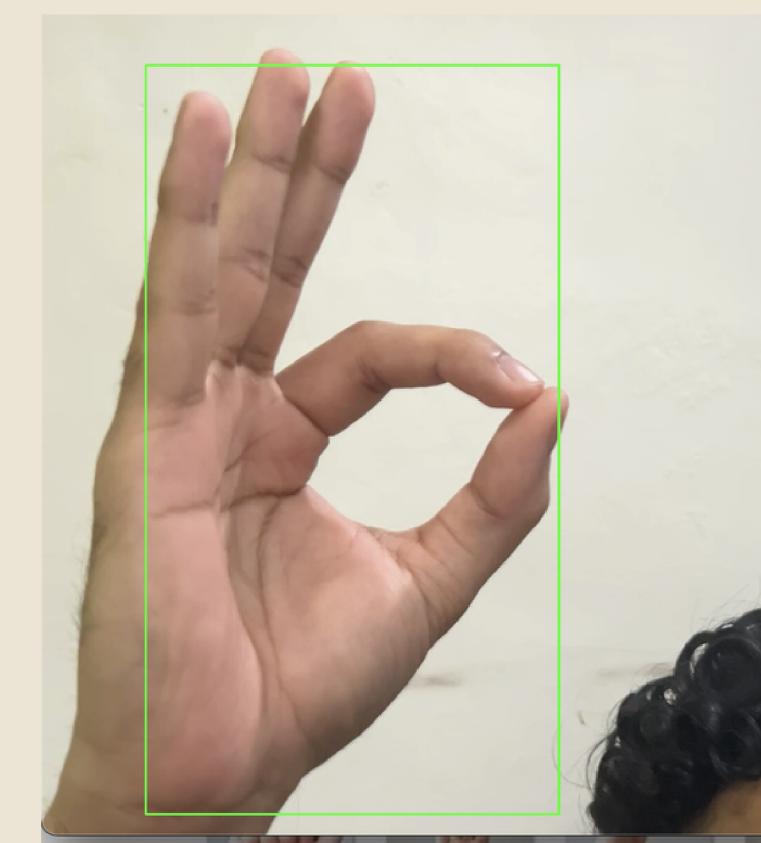
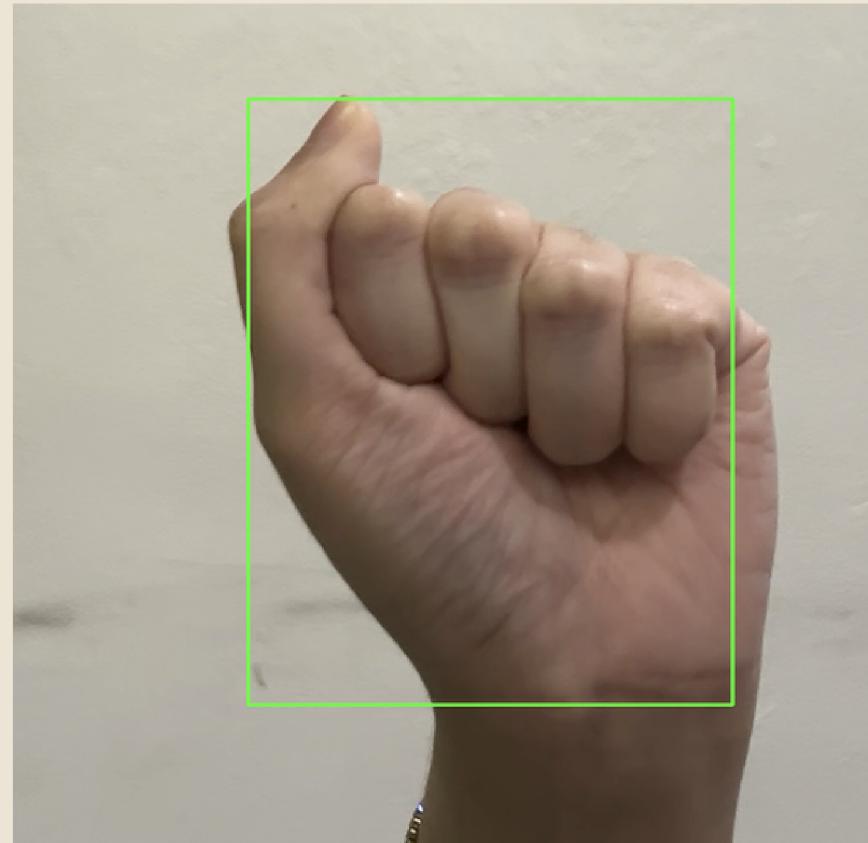
Loss Evolution



Accuracy over Epochs

Loss over Epochs

Live Predictions



Predicted Character: A

Confidence:
93.49%

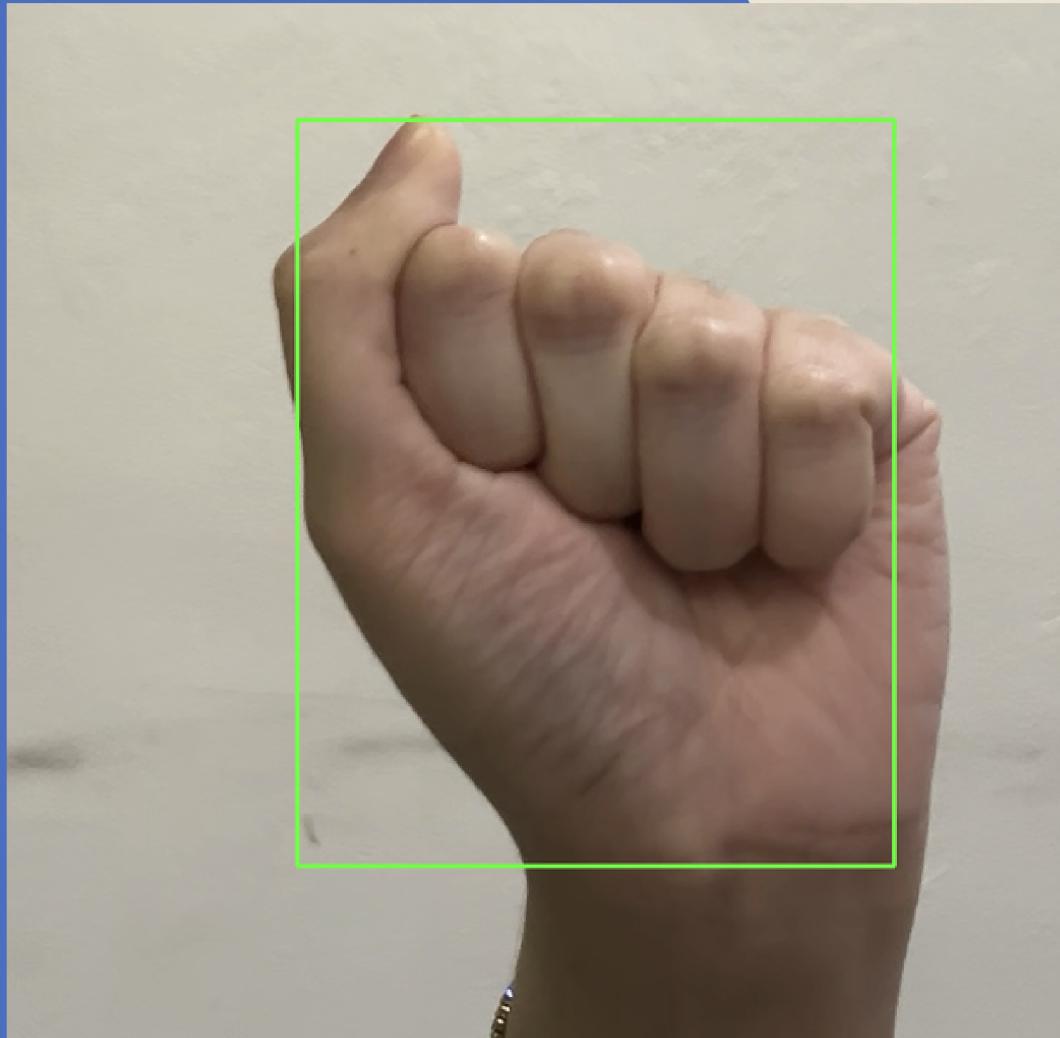
Predicted Character: F

Confidence:
94.55%

Predicted Character: H

Confidence:
99.94%

Bounding Box



A rectangular frame used to define the location of an object within an image, helping to identify and track the object.

It can accurately locate and isolate hand and finger movements, allowing the model to focus on the relevant gestures for precise interpretation and translation.

Conclusions

Our improved CNN model with cost-sensitive learning was able to recognize ASL alphabet gestures with greater accuracy.



Recommendations

- 1 Implement Annotated results to help instant results
- 2 Refine the model to eliminate biases towards skin colors and environments, enhancing accuracy, fairness, and robustness.



Live Demo

@ the parallel session later!





Thanks
for
“listening!”

