



Ahsanullah University  
of Science & Technology

## Real Time Bangladeshi Sign Language Detection Using Faster R-CNN

Oishee Binty Hoque, Mohammad Imrul Jubair, Md. Saiful  
Islam, Al-Farabi Akash, Alvin Sachie Paulson

**Dept. of Computer Science & Engineering**

# Introduction



12/28/2018

International Conference on Innovation in Engineering and Technology(ICIET) 2018

# Signers Communication

How can a **non-signer** communicate with a **signer**?

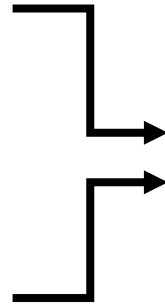
- A real-time interpreter might be a possible solution



# Signers Communication

**Our target is to construct a real-time interpreting system.**

- 🕒 To make communication easier



# Research Domain



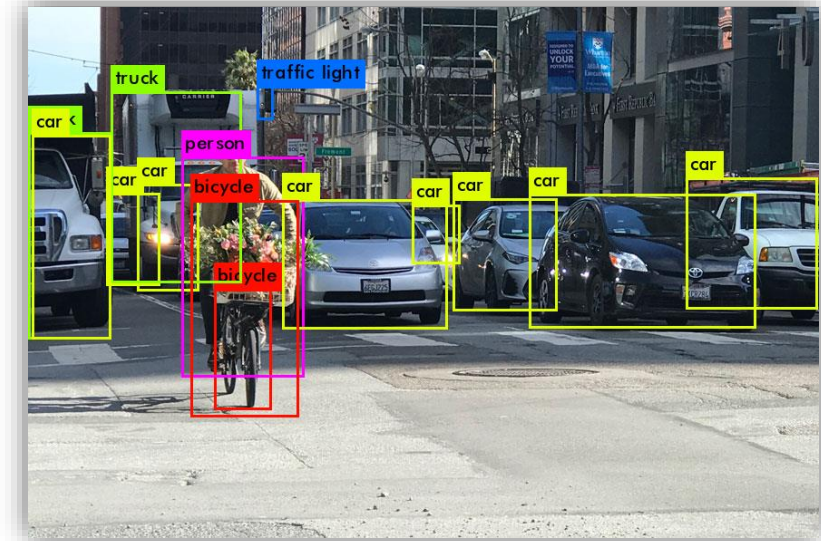
# Research Problem

## ➤ Bangladeshi Sign Letters Recognition

⊙ Real-Time

## ➤ *How can we implement that?*

⊙ Using **Deep Learning** based object detection method



➤ Object, i.e. 'car', 'bicycle', detection based on Deep Learning

➤ **Can 'signs' be detected similarly?**

# Related Works

# Existing Works

To best of our knowledge -

↳ **Rahman et al. 2018**

- ⊙ “Bangla Language Modeling Algorithm For Automatic Recognition of Hand-Signspelled Bangla Sign Language”

↳ **Yasir et al. 2018**

- ⊙ “Bangla Sign Language Recognition Using Convolutional Neural Network”

↳ **Ahmed et al. 2016**

- ⊙ “Bangladeshi Sign Language Recognition Using Fingertip Position”

↳ **Rahman et al. 2015**

- ⊙ “Computer Vision Based Bengali Sign Words Recognition Using Contour Analysis”

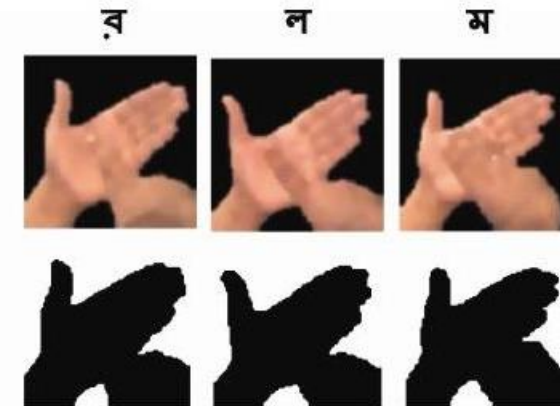
↳ **Rahman et al. 2014**

- ⊙ “Realtime Computer Vision-Based Bengali Sign Language Recognition”



# Reviews

- Additional device for input
- Dataset
  - ⊙ Requires preprocessing, such as background segmentation
  - ⊙ Absence of variations in backgrounds, illuminations, etc (which is MUST for Deep learning)
- Methodology
  - ⊙ Traditional machine learning method
  - ⊙ Beforehand feature extraction
- Output
  - ⊙ Not “completely” real time
  - ⊙ Not “enough” real time to be integrated in applications i.e. Smartphones



*\*For detailed review, please go through our paper \*\*Images are collected from Internet and Rahman et al.*

# Challenges and Contributions of our Work



# Challenges

The method must be -

- Device Independent
- Real-Time in erratic background

# Challenges

- Needs Deep Learning based recognition methods
  - ⊙ Dependent on Robust Dataset
    - Large number of images/class
    - Enough variation in input data in terms of
      - ⊙ Background, Gesture Angle, Age and Gender

# Contribution

| Domain     | Our Contribution            |   |
|------------|-----------------------------|---|
| Dataset    | Open for Community          | ✓ |
|            | Dynamic Lightning Condition | ✓ |
|            | Dynamic Background          | ✓ |
|            | Similar Gesture             | ✓ |
| Our System | Device Based Input          | ✗ |
|            | Real Time                   | ✓ |

# Methodology (Faster R-CNN)



# Faster R-CNN

## ➤ Basic Concepts

- ⊙ Neural Network & Convolutional Neural Network (CNN)
- ⊙ Anchors
- ⊙ Regional Proposal Network (RPN)

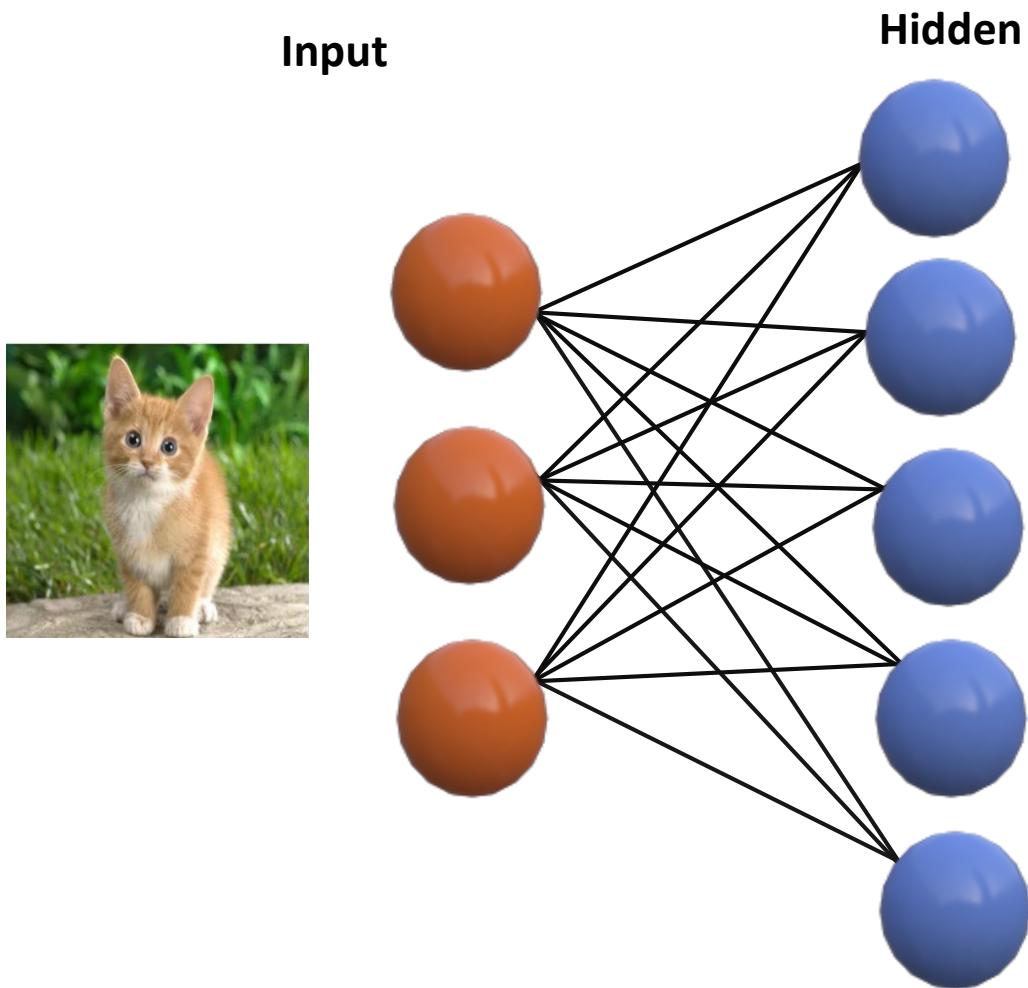
# Neural Network

Input

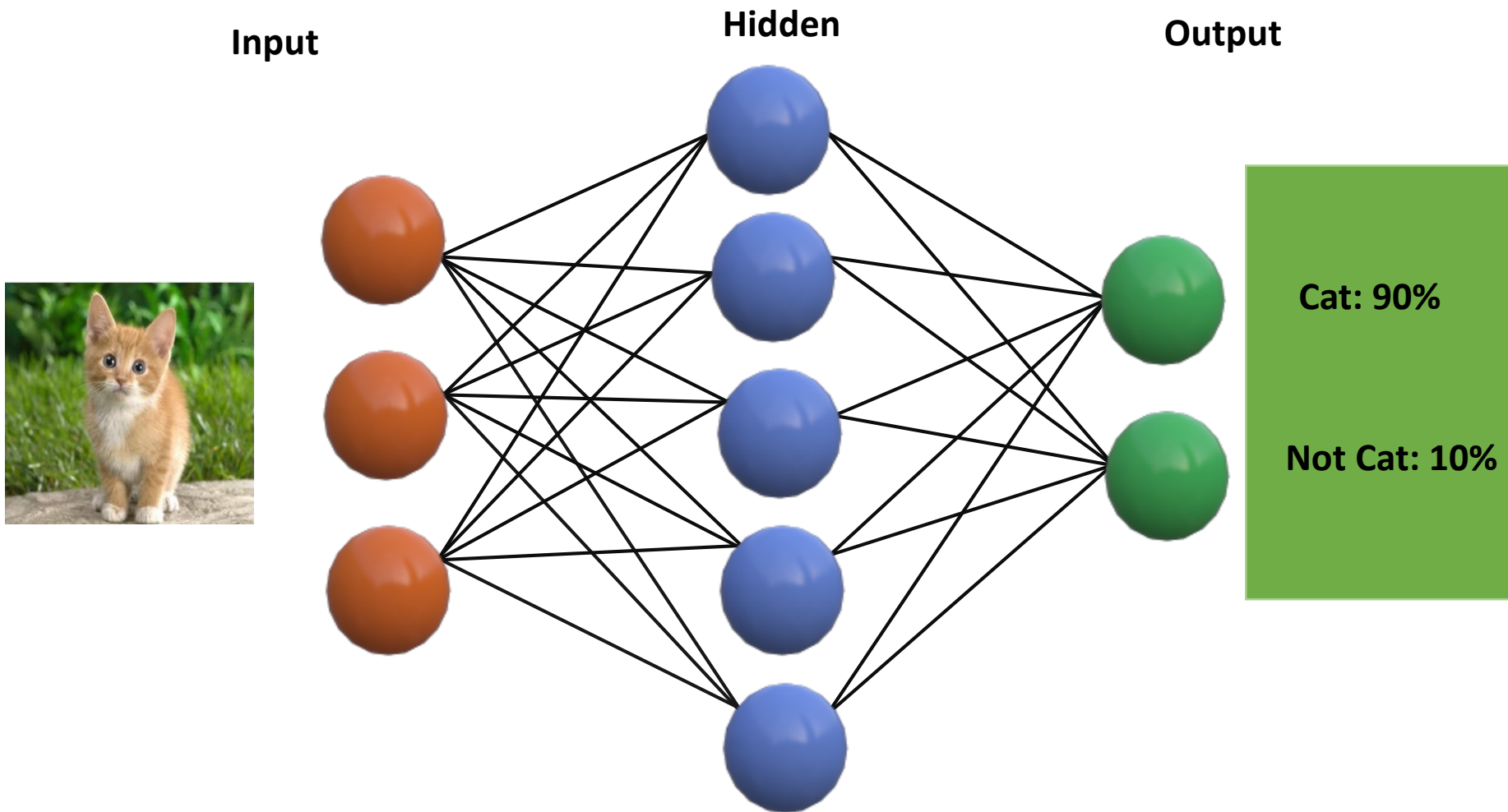




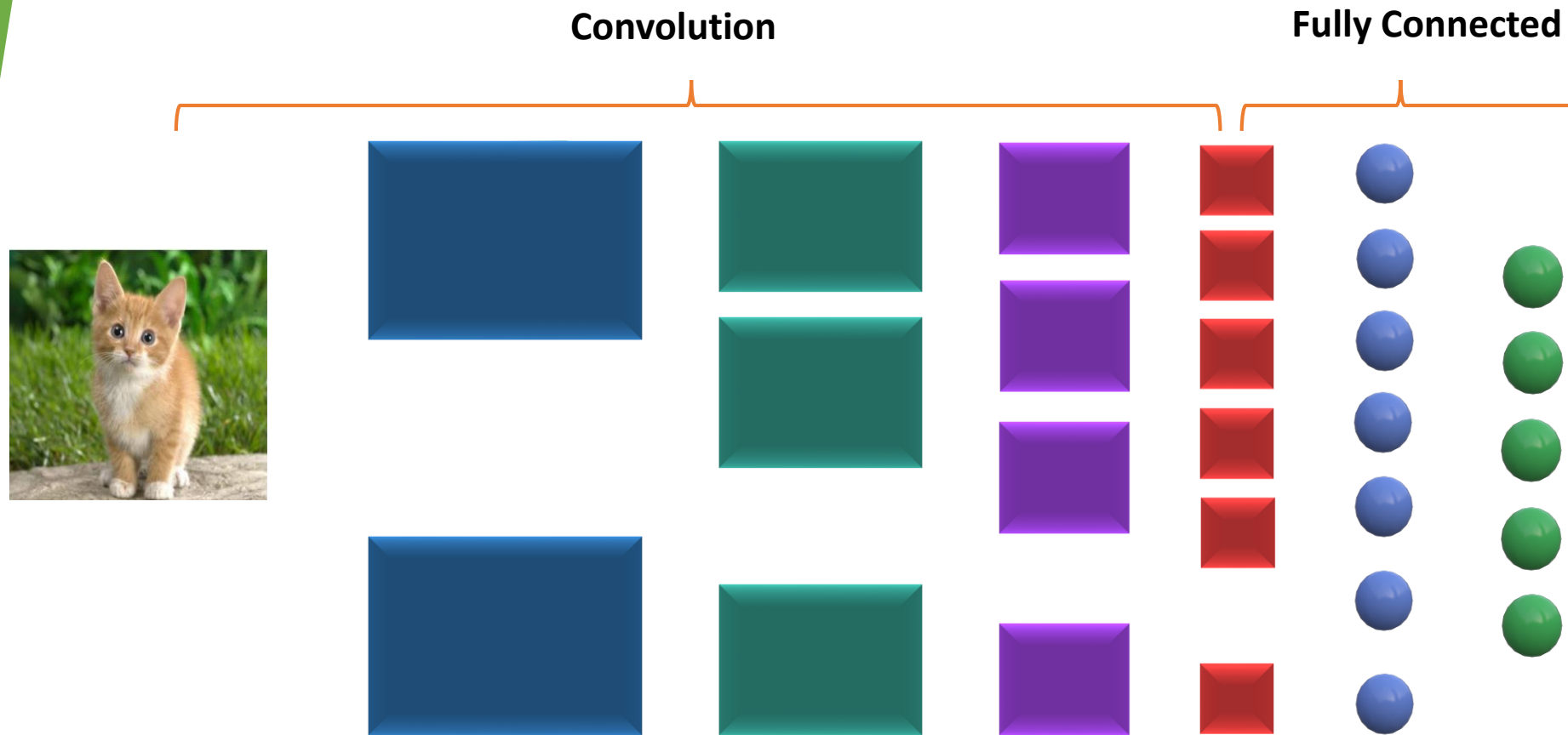
# Neural Network



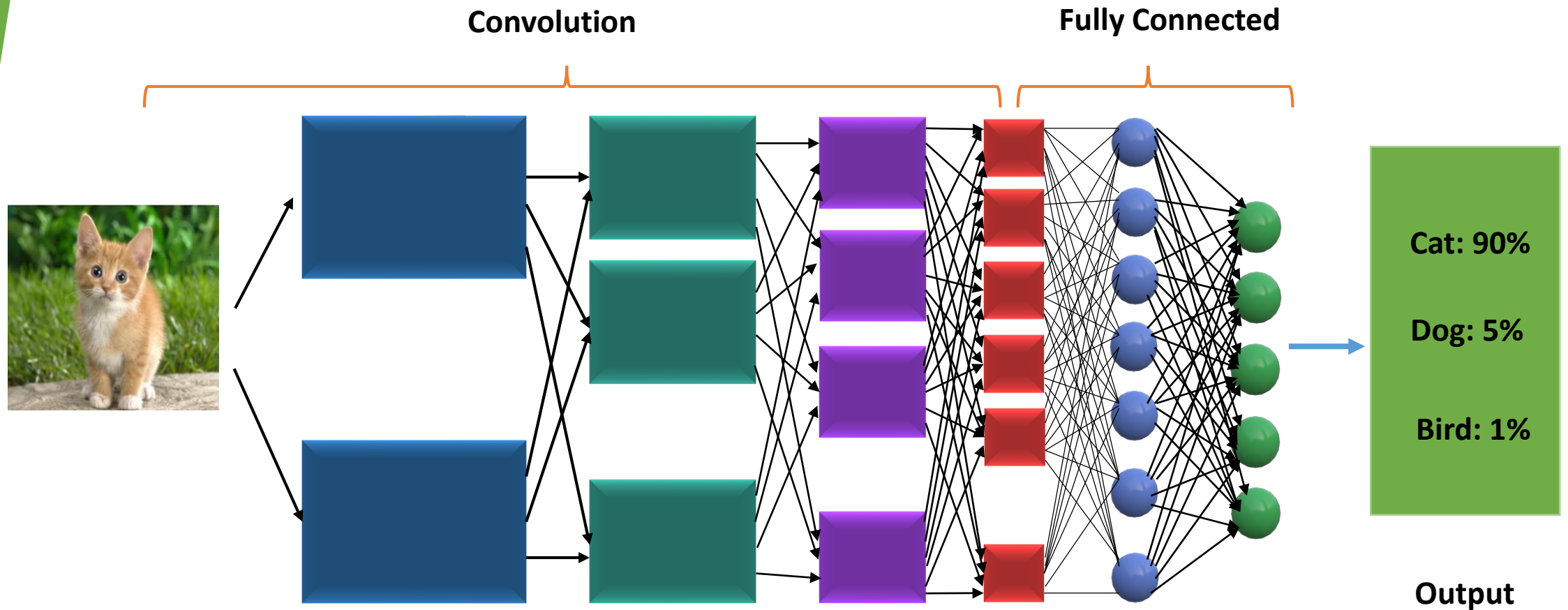
# Neural Network



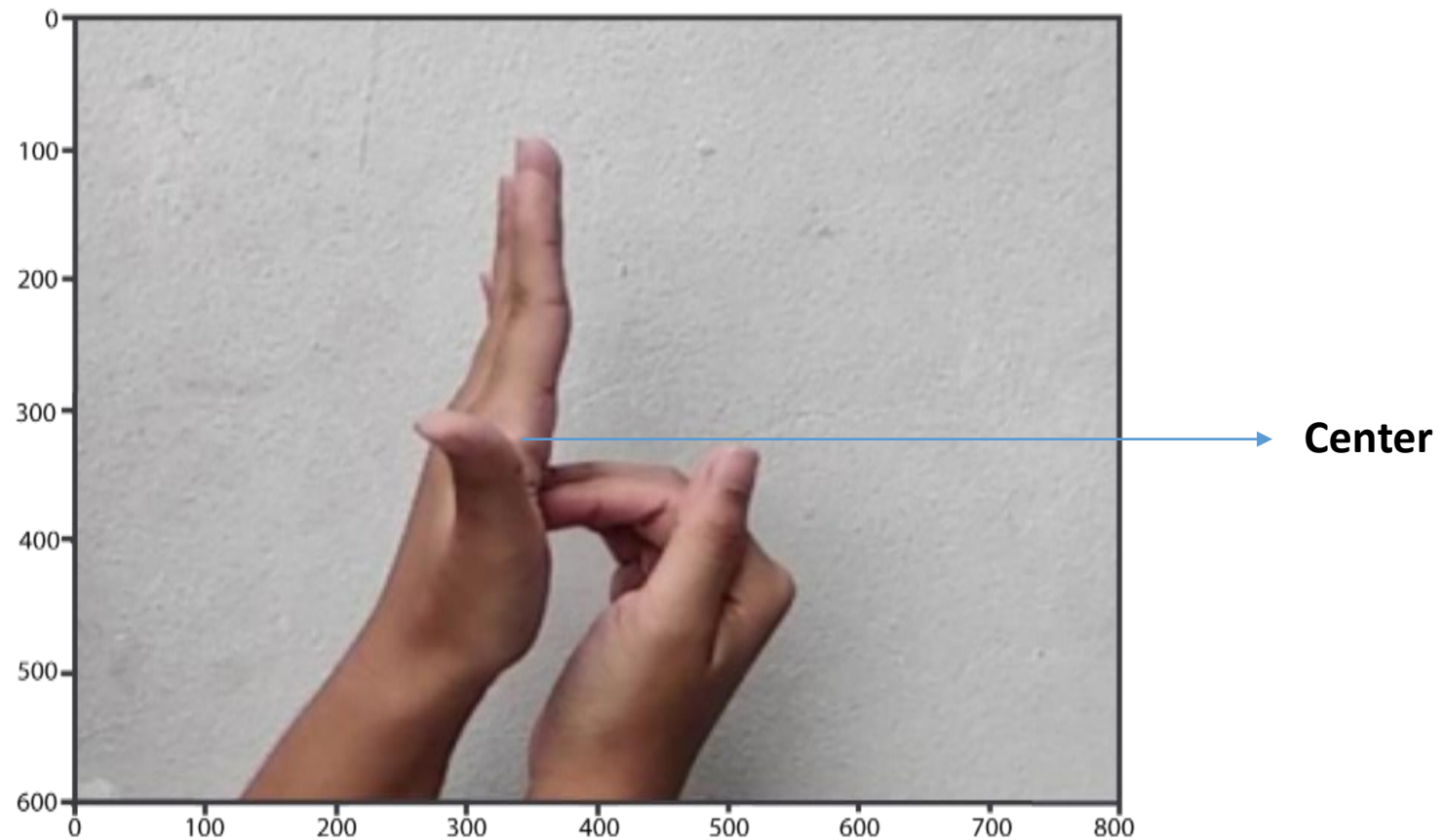
# Convolutional Neural Network (CNN)



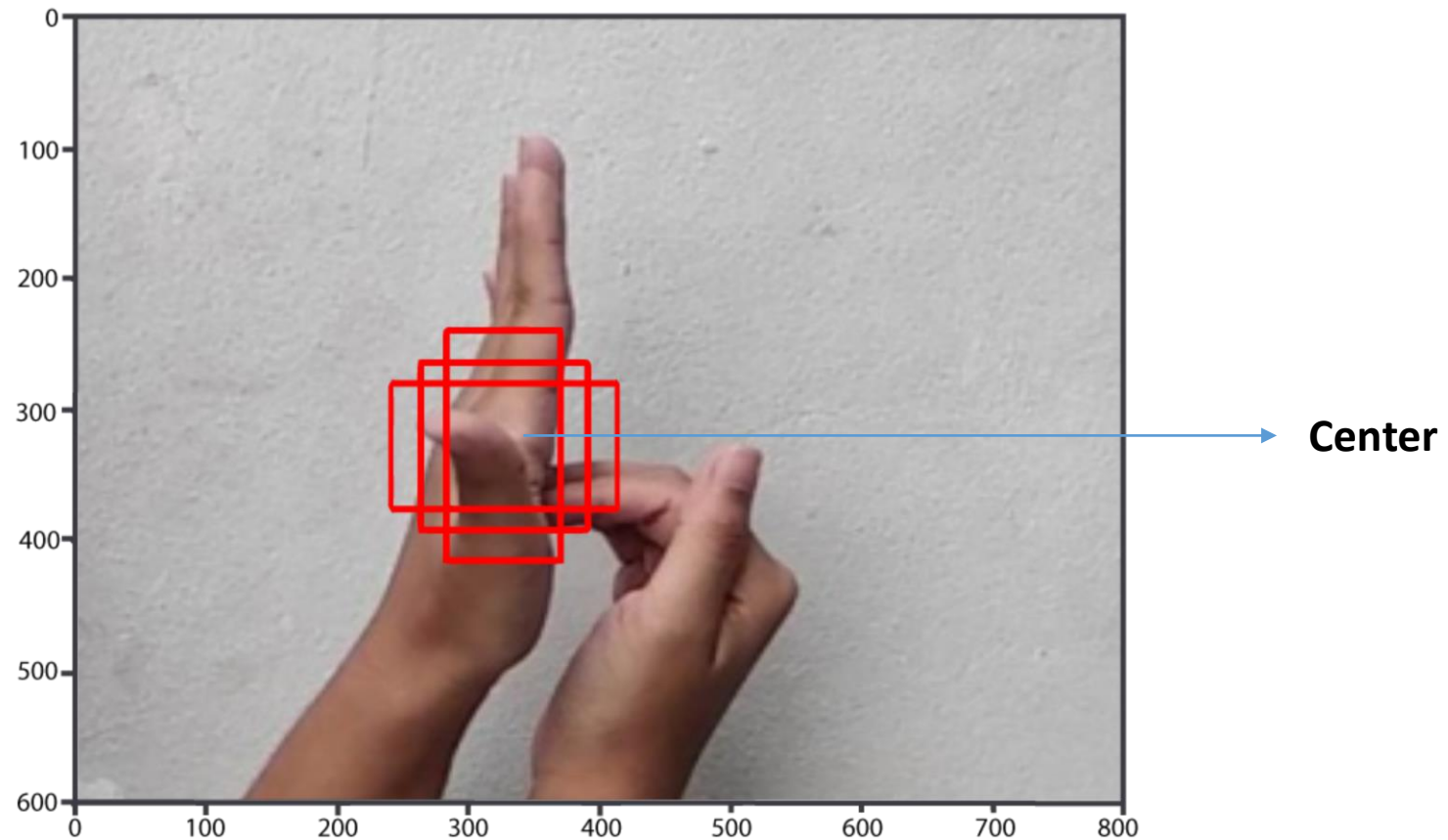
# Convolutional Neural Network (CNN)



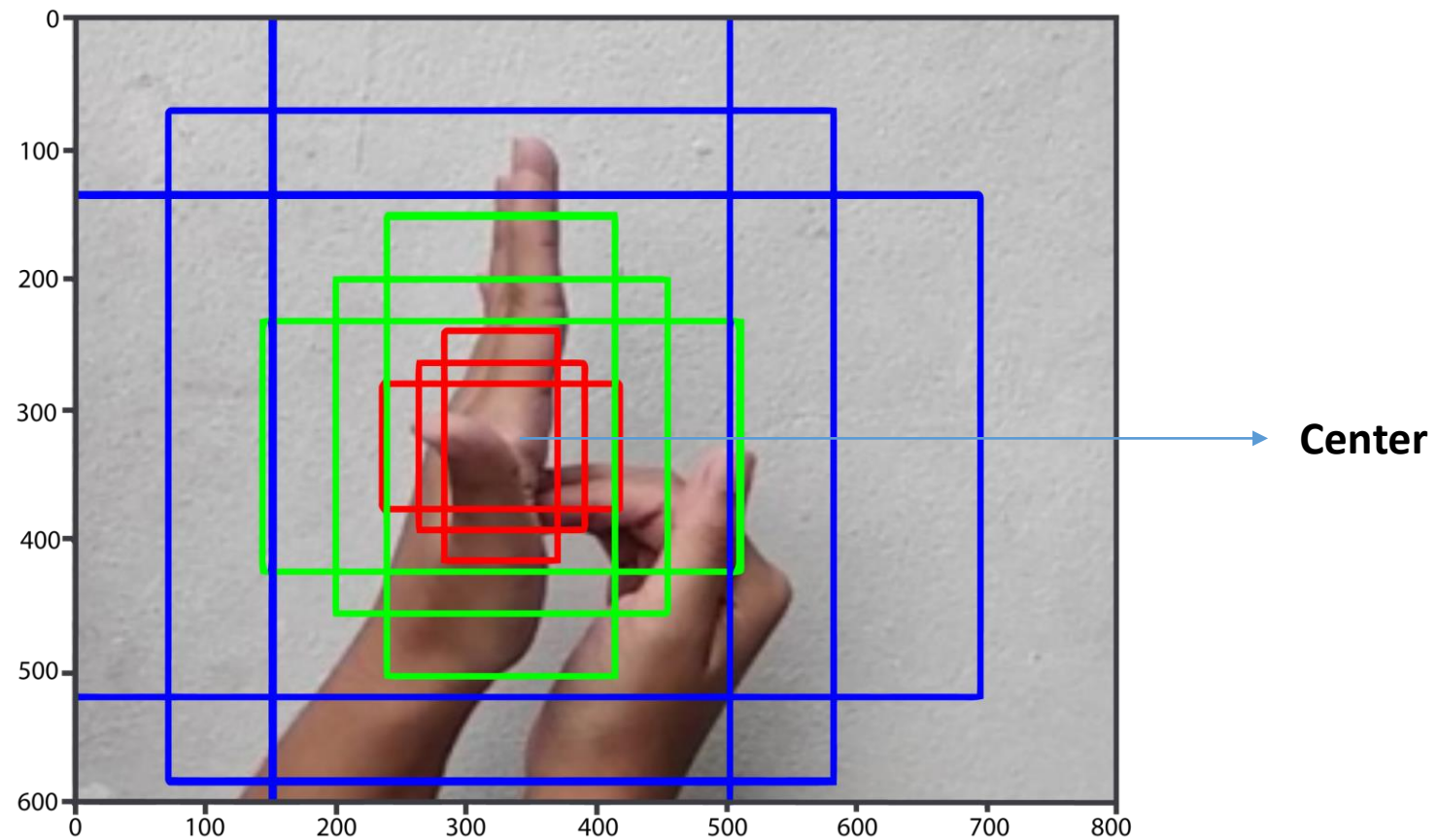
# Faster R-CNN (Anchor)



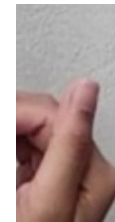
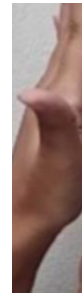
# Faster R-CNN (Anchor)



# Faster R-CNN (Anchor)



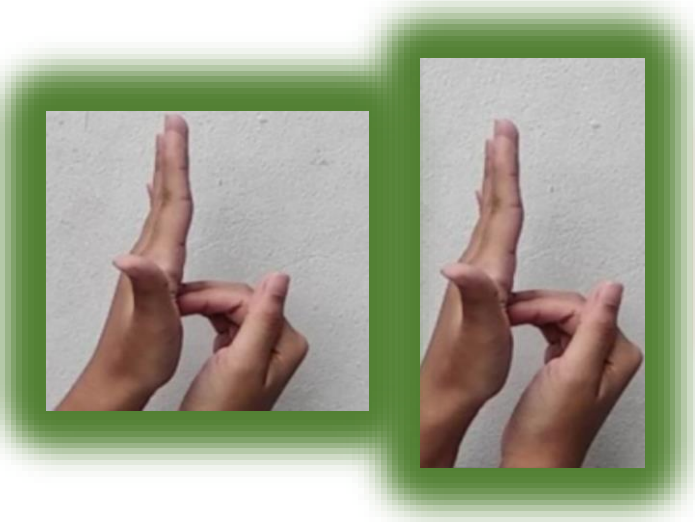
# Some Example of Anchors



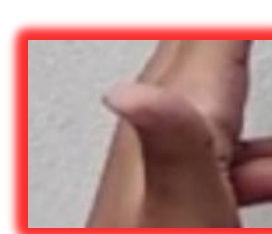


# Region Proposal Network

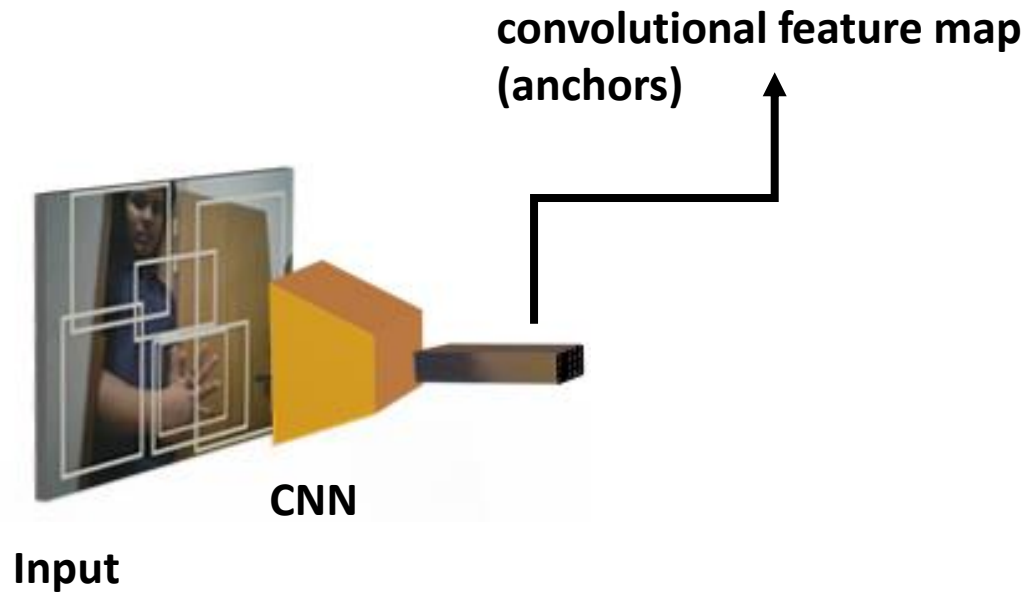
Selected



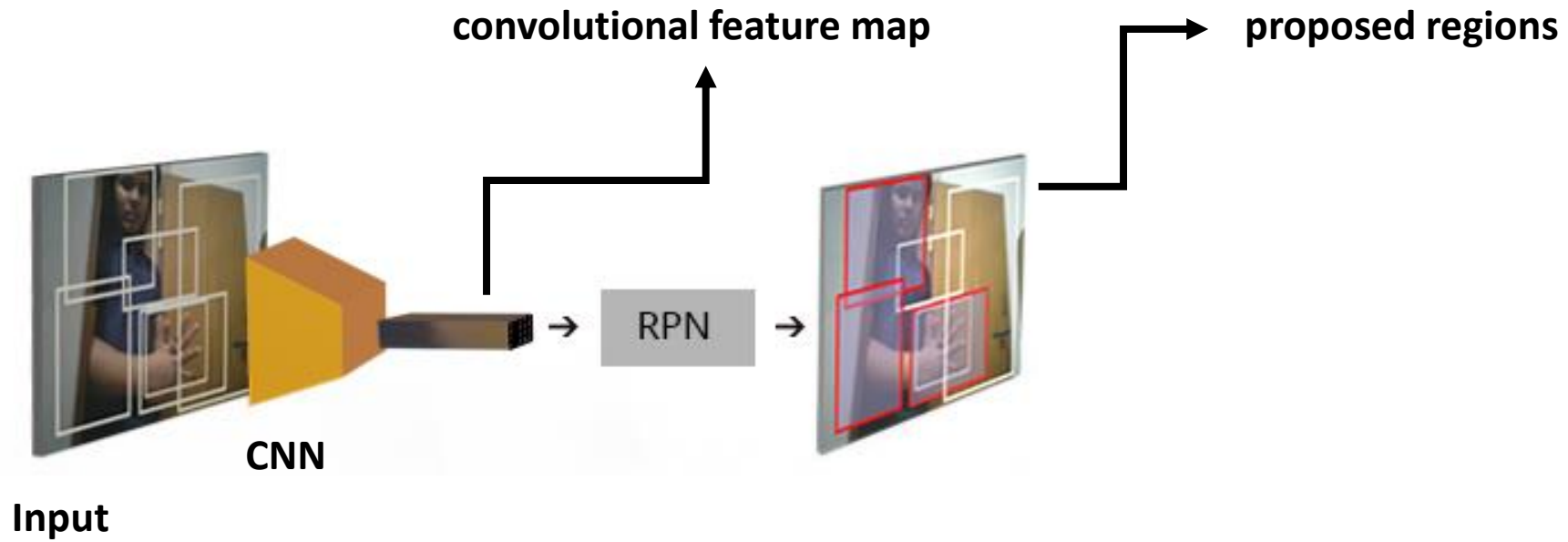
Discarded



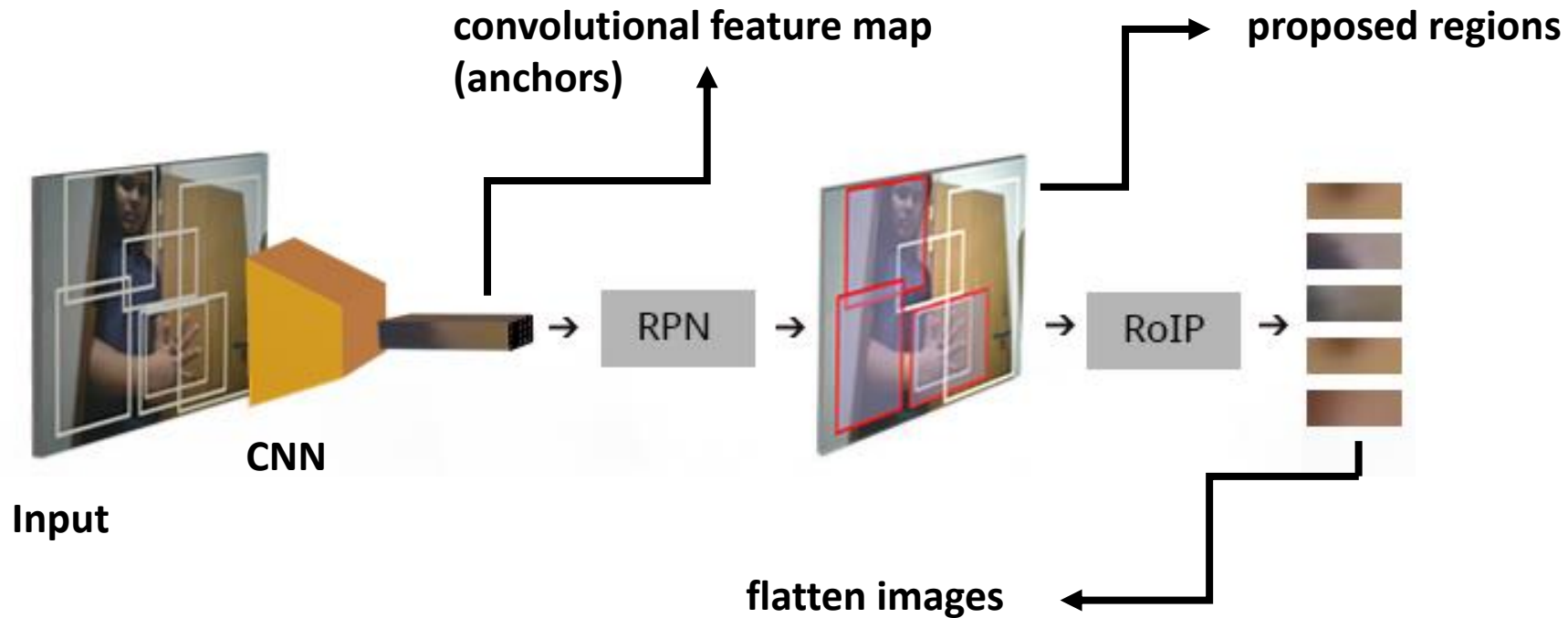
# Faster R-CNN (Architecture)



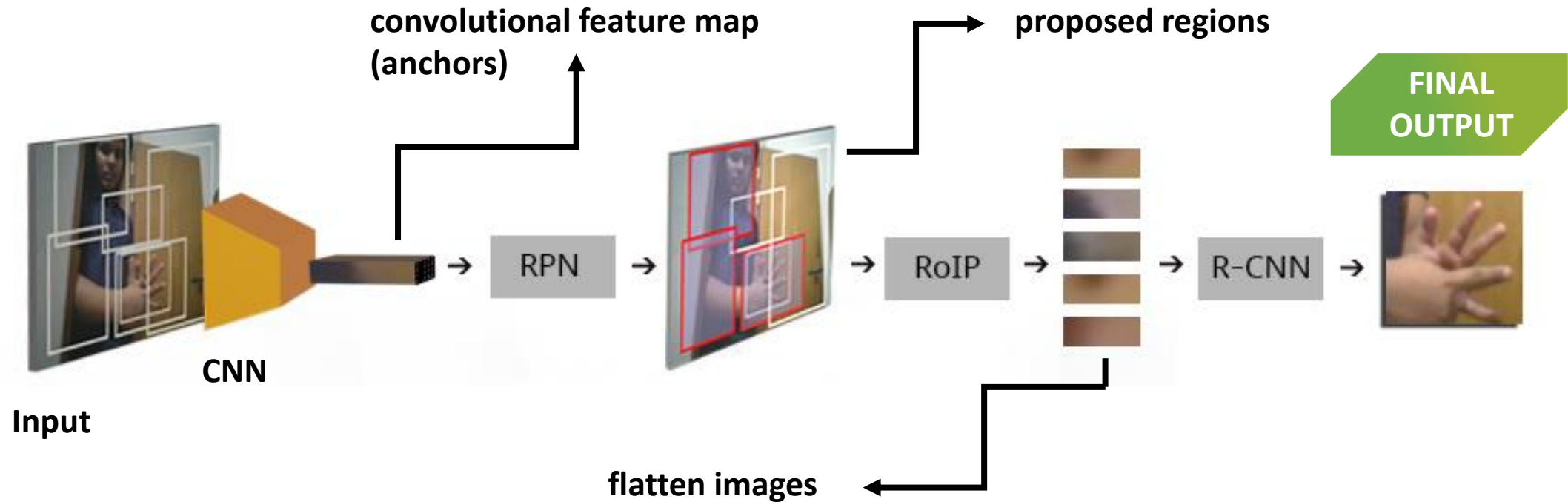
# Faster R-CNN (Architecture)



# Faster R-CNN (Architecture)



# Faster R-CNN (Architecture)



# Bangladeshi Sign Language Image Dataset (*BdSLImSet*)

# BdSLImSet

- Background Variation
- Different Signers



Fig : A bit samples from our Dataset

# BdSLImSet: Labelling

- Each image of per class is labeled
- Converted into **XML** files
- Available in **Github** (<https://github.com/imruljubair/bdslimset>)

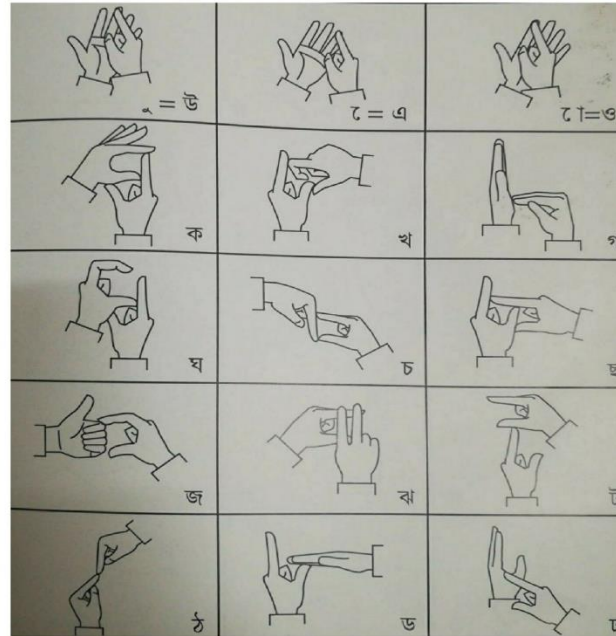




# BdSLImSet: Verification

➤ This dataset has been verified By –

⊙ **DHAKA BADHIR HIGH SCHOOL ( ঢাকা বধির হাই স্কুল)**



Source: A book of Bangla Sign Letters

# BdSLImSet: Current specifics

| Total Images | Total Class | Images/ Class | Image Size | Resolution | Number of Participants | Training Set: Testing Set |
|--------------|-------------|---------------|------------|------------|------------------------|---------------------------|
| 1600         | 10          | 100           | ≤200kb     | ≤700*1280  | 10                     | 80:20                     |

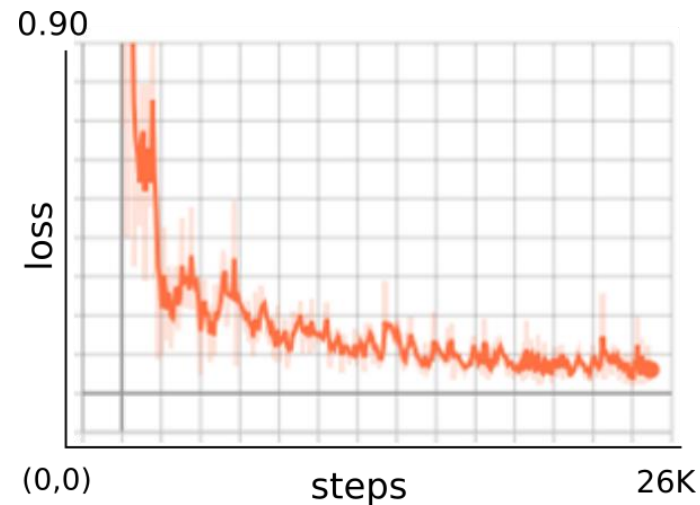
# Experiment & Result

# Training Phase

- Implementation platform –
  - ⊙ Tensorflow GPU-V1.5
  - ⊙ CUDA V9.0
  - ⊙ CPU from Intel<sup>®</sup>. CoreTM i7-7500U of 2.7 GHz upto 3.5Ghz
  - ⊙ GPU Nvidia 940mx with 4.00GB
- 10 classes with 100 images for each letter.

# Training Phase

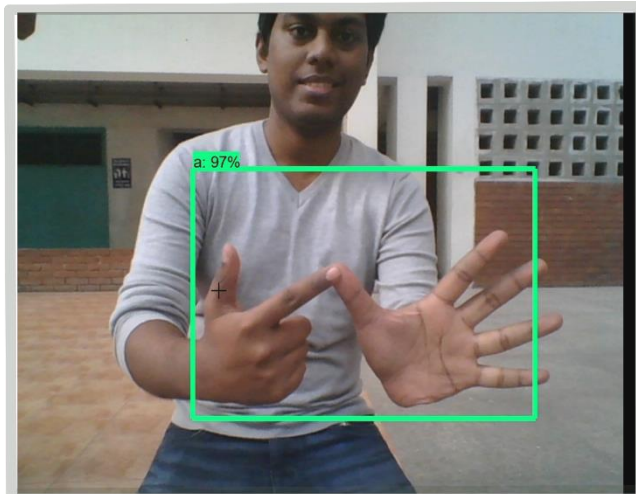
- ④ Took about 12 hours
- ④ 28000 iterations to train the model.
- ④ Started with loss of 3.00, quickly dropped to 0.8. Stopped at 0.07



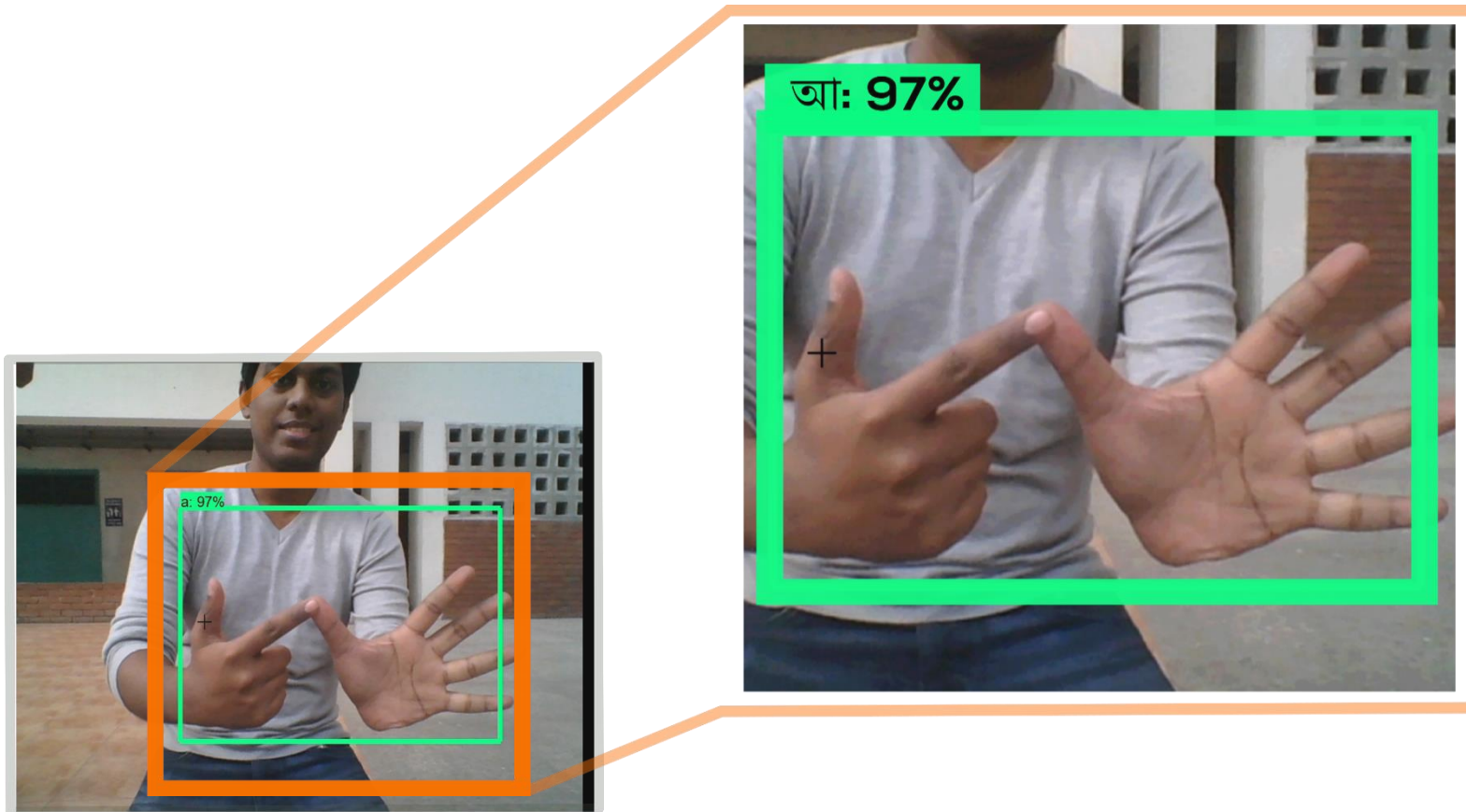
# Experimental Result

- Confidence Rate on average **98%**.
- Detection Time **90.03 ms**.
- Accuracy Rate **98.20%**.

# Experimental Result

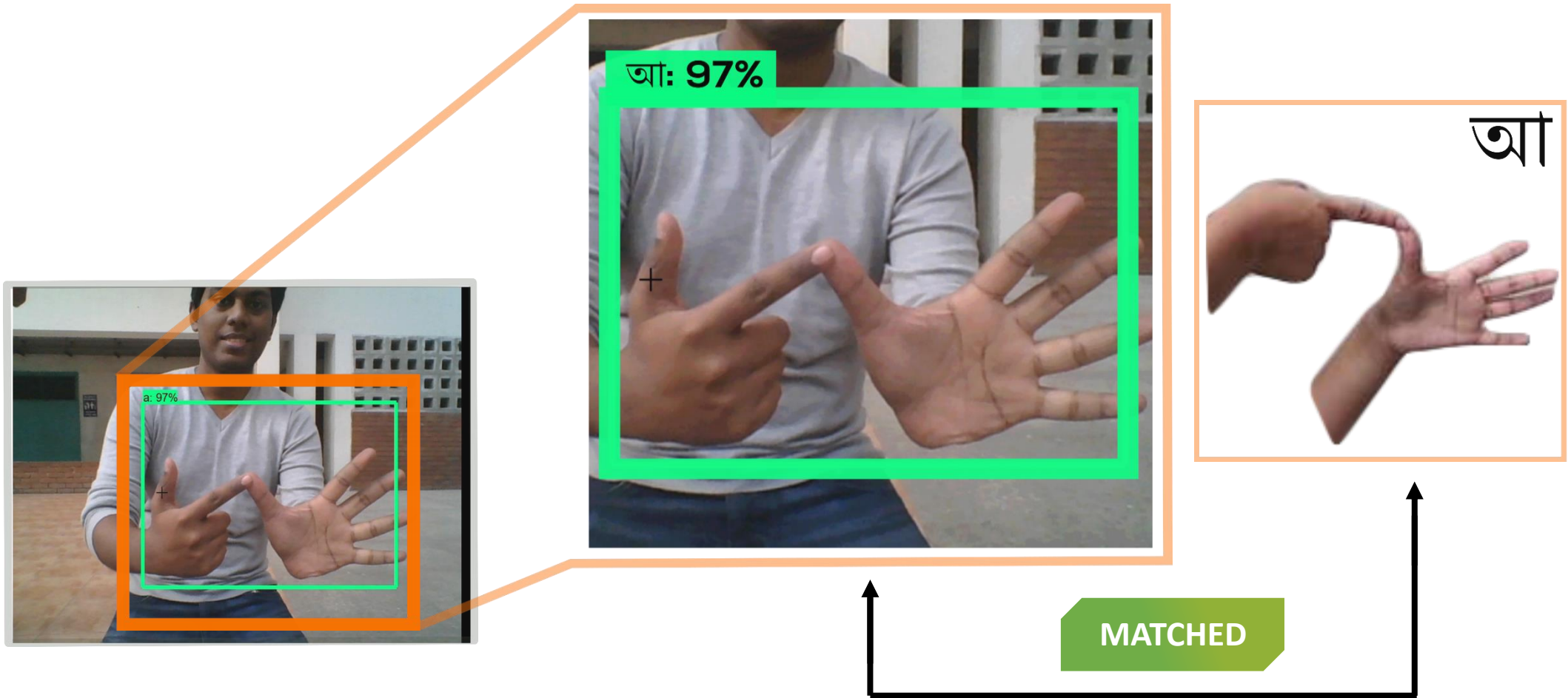


# Experimental Result

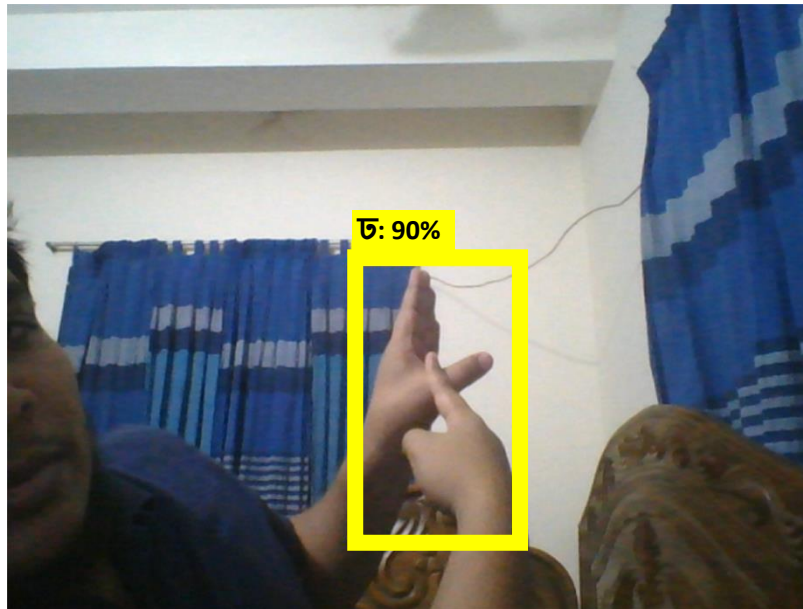




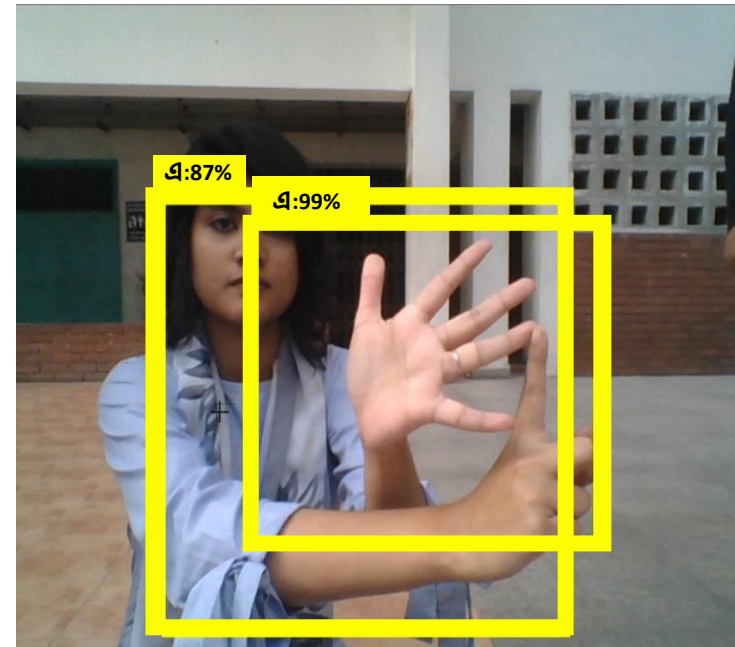
# Experimental Result



# Experimental Result

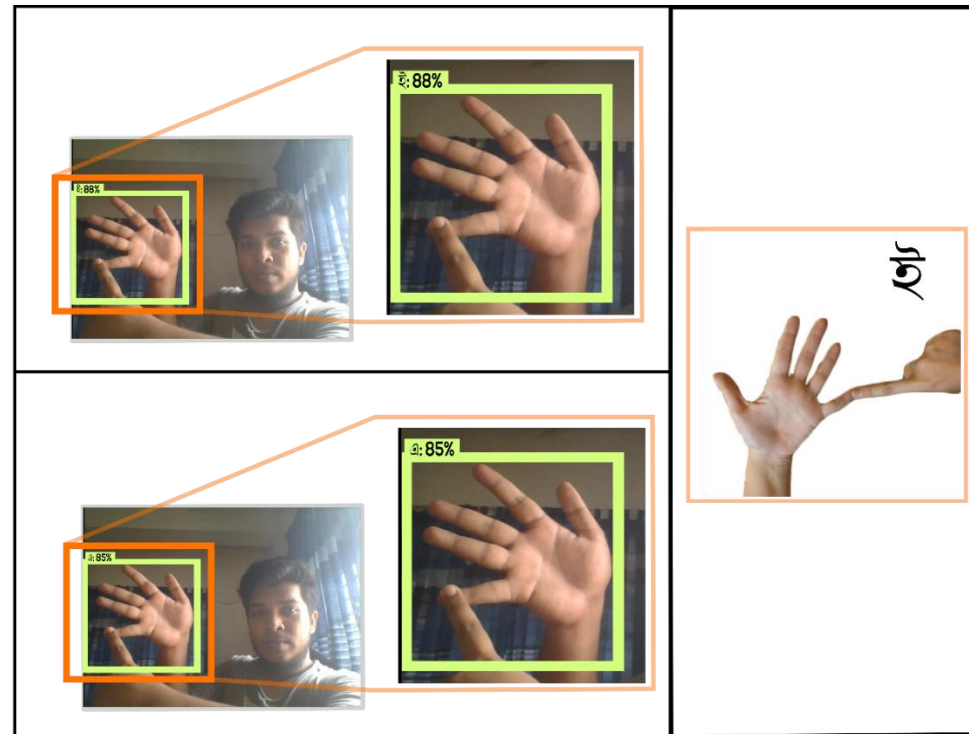


# Experimental Result



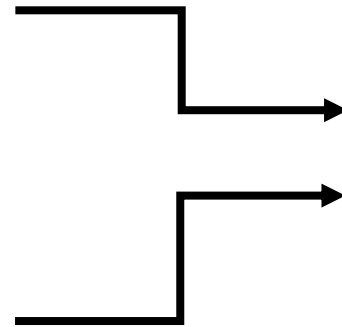
# Limitations

- Two letter with similar gesture had faulty recognition sometimes.



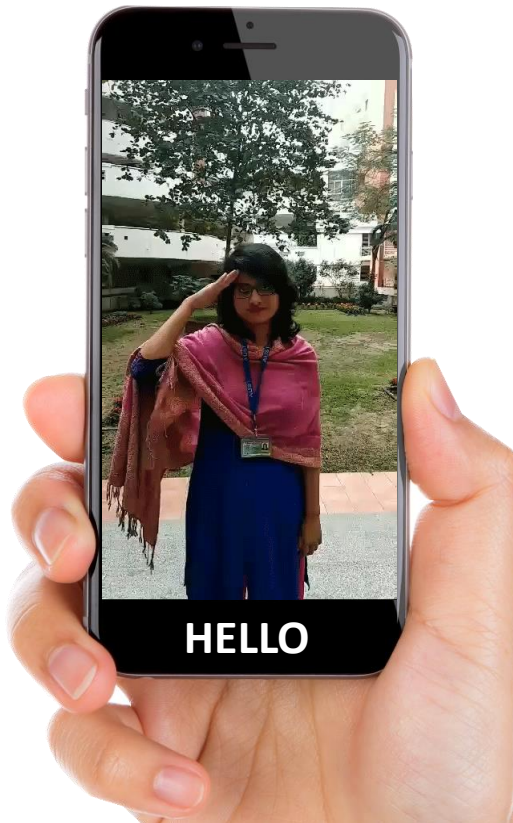
# Future Plan

## ➤ Android Based Recognition System



# Future Plan

- Word and Sentence Recognition In Real Time



# Future Plan

- User friendly system
- Increasing dataset
- Evaluating the performance of our system by genuine users  
(i.e. hearing-impaired people or sign language teacher)

# Demo Screenplay

➤ Go to the youtube link:

<https://youtu.be/8NLwOpQCmW0>



*Any Questions?*

