Ahsanullah University of Science and Technology

Real-Time Bangladeshi Sign Language (BdSL) Recognition using Deep Learning

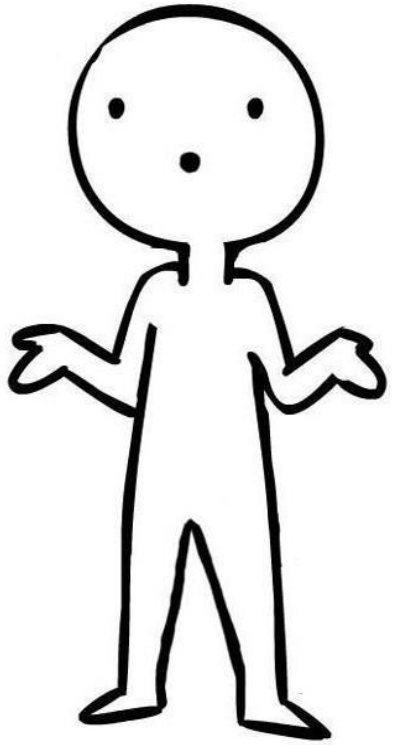
Presented By :

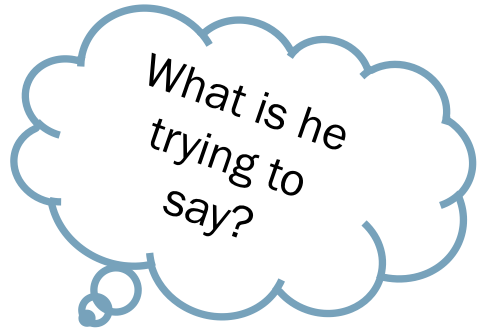
Oishee Bintey Hoque
Al-Farabi Akash
Md. Saiful Islam
Alvin Sachie Paulson

Supervised By :

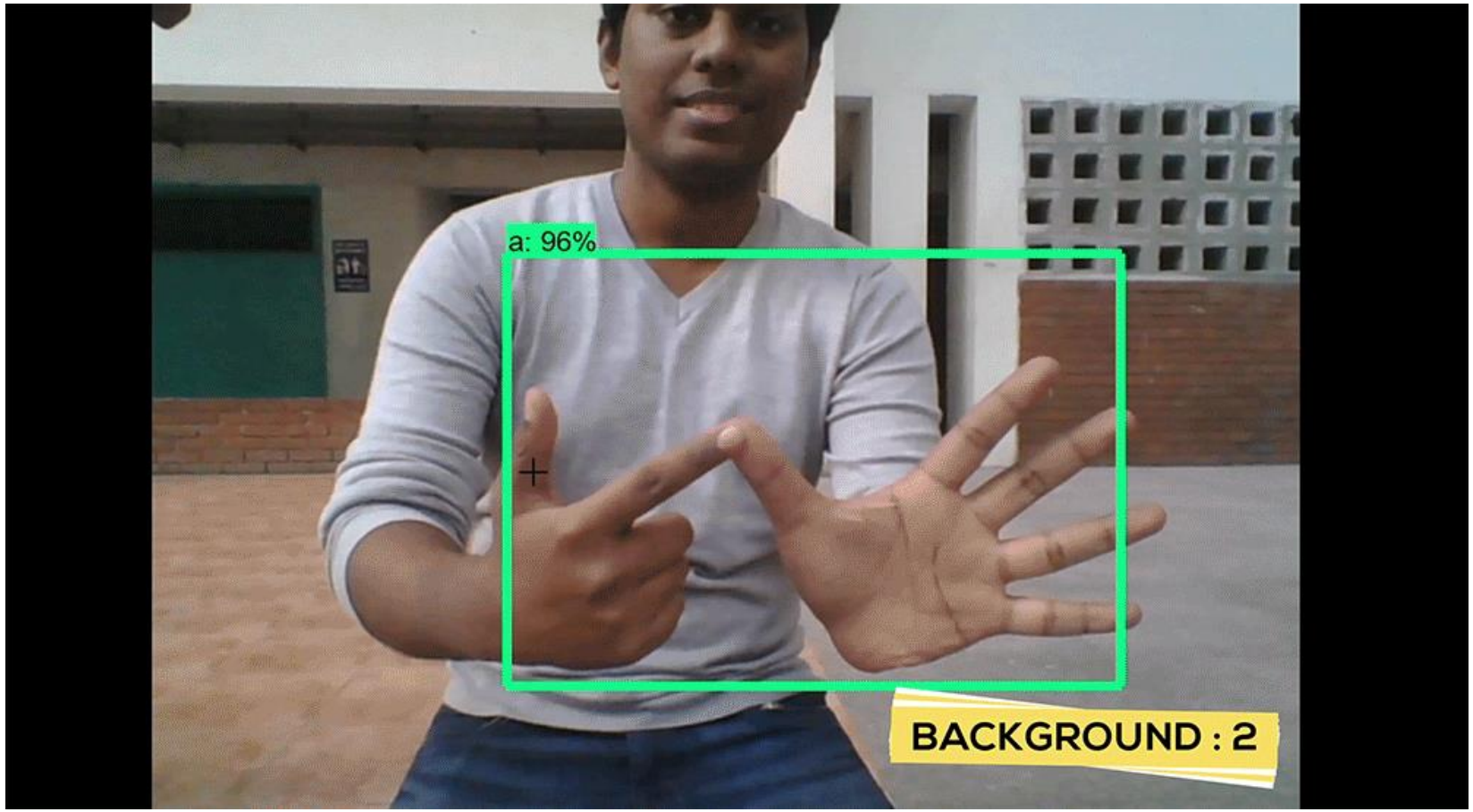
Mr. Mohammad Imrul Jubair
Assistant Professor,
Department of Computer Science and Engineering,
Ahsanullah University of Science and Technology.

HAVING PROBLEM ?



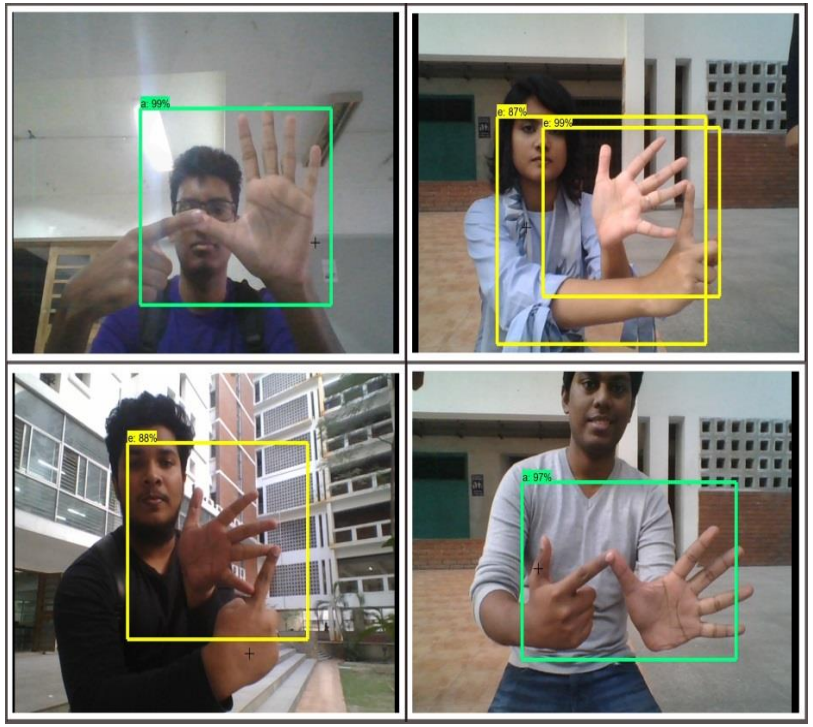


Solution



Research Domain

- Identifying Bangladeshi Sign Letters from Image
- Real-Time Recognition of Sign Letters



Challenges

- Huge & Robust Dataset
- Variation of Background
- Similarities Among Sign Letters
- Real–Time Recognition

Contribution

Domain	Our Contribution
Dataset	• Generated A Robust Dataset • Available For Future Research
Real Time	✓
Background Variation Problem	✗
Similar Gesture Problem	✗

Deep Learning in BdSL

- Most of the BdSL Recognition work has been done using traditional machine learning
- If it is already implemented, what's the problem?

Deep Learning in BdSL (cont.)

- Manual Feature Extraction is difficult in BdSL
- Deep learning methods have overcome manual hardcore feature extraction

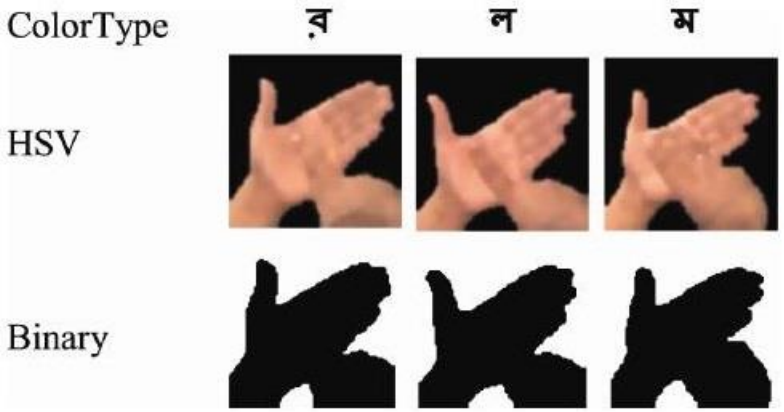


Fig : Rahman et al. converted their dataset to binary image but faced indistinguishable problem while converting HSV to Binary image. As the shown letter has similar feautures, in binary image it is difficult to identify the difference.

Dataset

Dataset Collection Challenges

- BdSL Dataset Not Available
- Deep Learning Method Needs Robust Dataset
- Verification By Signers

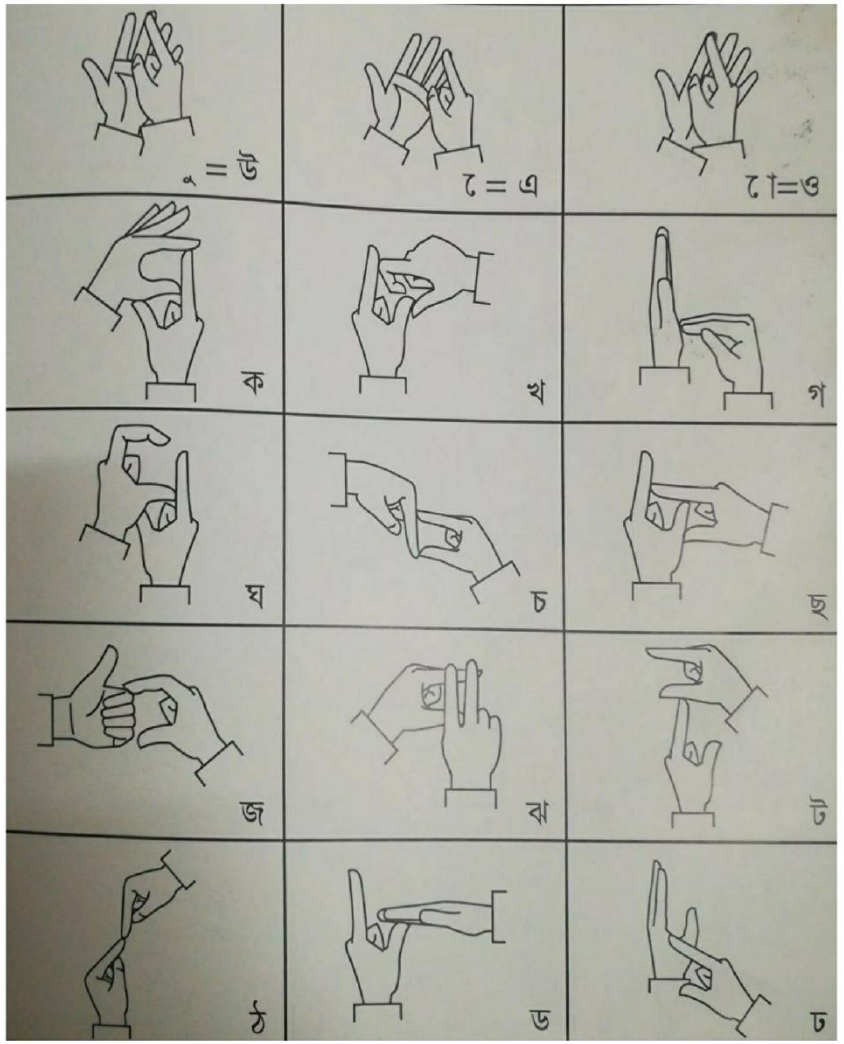


Fig: A set of Bengali Sign Letters, collected from book

Dataset Processing

- Varieties of images per class in terms of background, gender, age, position

BdSLImset

Fig: Examples of some samples of BdSLImset

Dataset Processing

- Varieties of images per class in terms of background, gender, age, position

- Image Labeling for each class

Labled BdSLImset

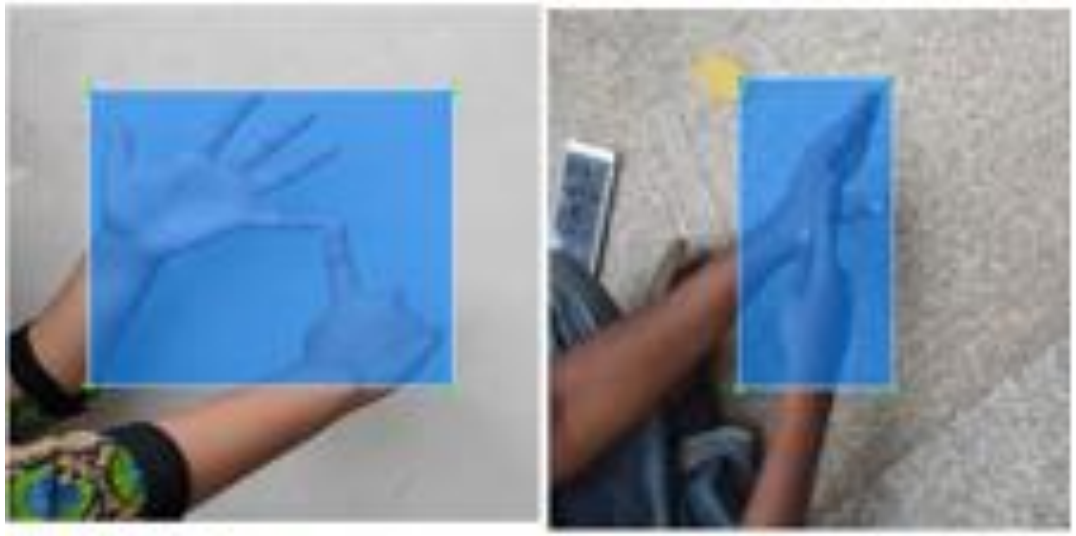


Fig: Examples of some samples of BdSLImset

Dataset Processing

- Varieties of images per class in terms of background, gender, age, position
- Image Labeling for each class
- Dataset is verified by “Dhaka Badhir School”.

BdSLImset (Bangladeshi Sign Language Image) Dataset

Total Images	Total Class	Images/ Class	Image Size	Resolution	Background	Number of Participants	Training Set: Testing Set
1600	10	100	<=200kb	<=700*1280	Dynamic	10	80:20

Other BdSL Datasets & BdSLImSet

Related Work	Background	Image Per Class * Total Classes	No. of Signers	Availability
Rahman et al. 2014	Static	36 * 10	10	×
Rahman et al. 2015	Static	10 * 10	10	×
Ahmed et al.	Static	37 * 14	3	×
Our (BdSLImset)	Randomized	10 * 10	10	✓

Real-Time Recognition With Faster R-CNN

NEURAL NETWORK PROCESS

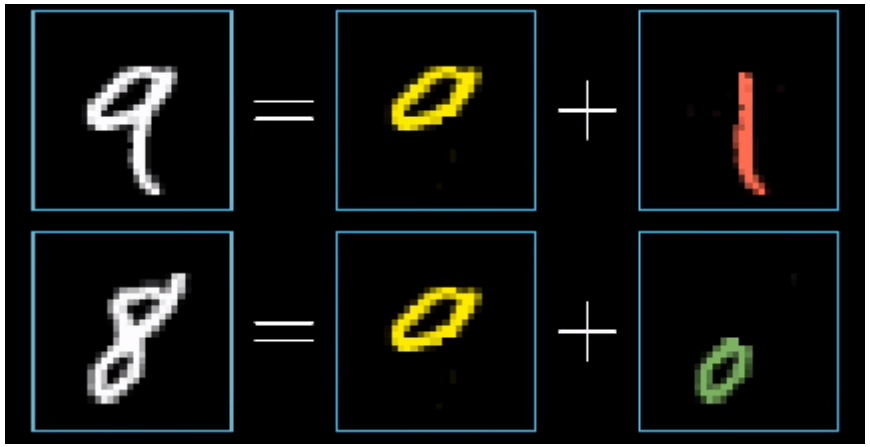


Fig 1: Breaking down numbers in Sub Pattern

Convolutional Neural Network (CNN)

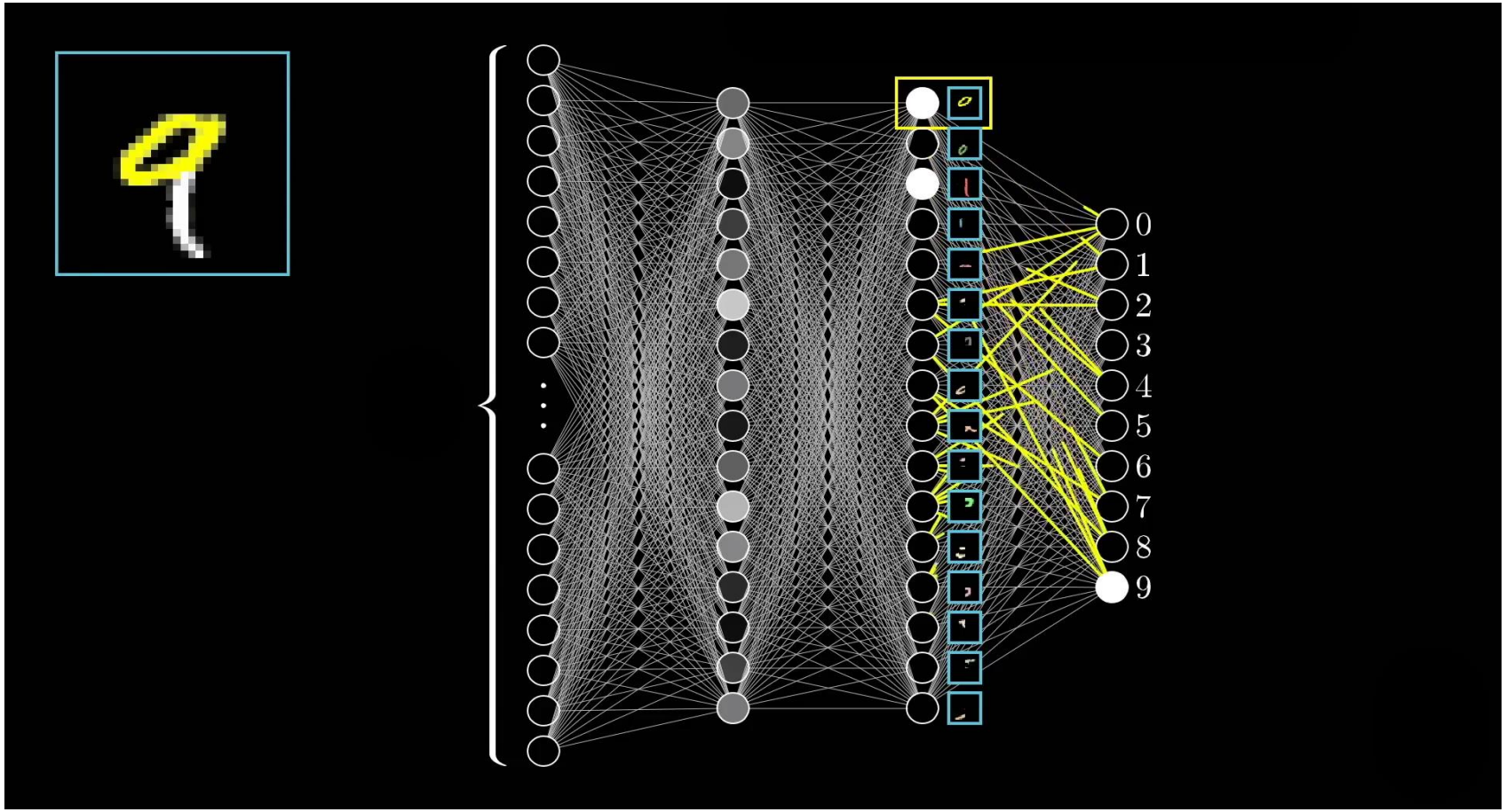


Fig 2: Recognizing Number using Neural Network

Recurrent Convolutional Neural Network (R-CNN)

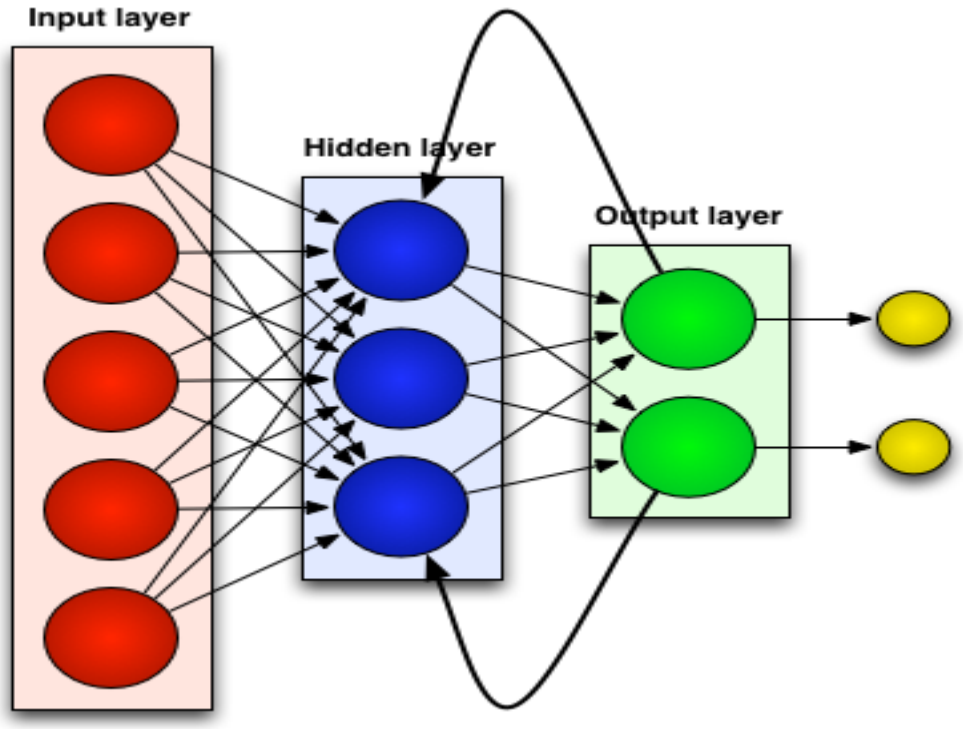
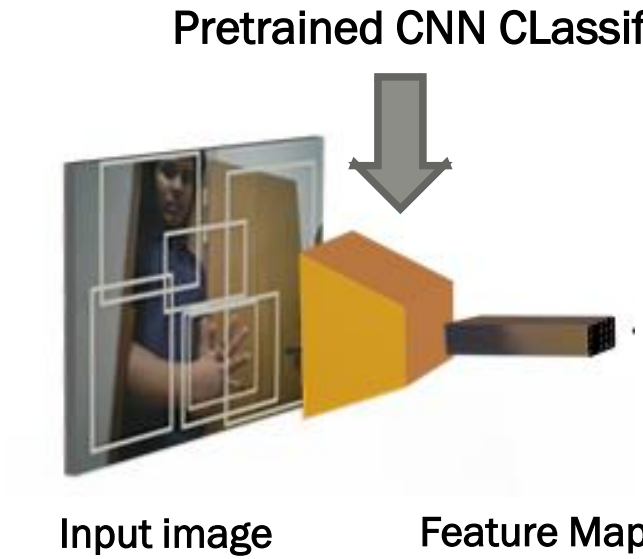


Fig 2: Recognizing Number using Neural Network

Faster R-CNN Architecture



- CNN Creates Feature Map
- Feature Map Includes Anchors

Fig : Process Of Faster R-CNN

Anchors

Terms of anchors

- 9 anchors at each position of an image
- Three colors represent three scales
  128*128, 256*256, 512*512
  ratios 1:1, 1:2 and 2:1 respectively.
- Final image with a bunch of anchors separated by pixel.

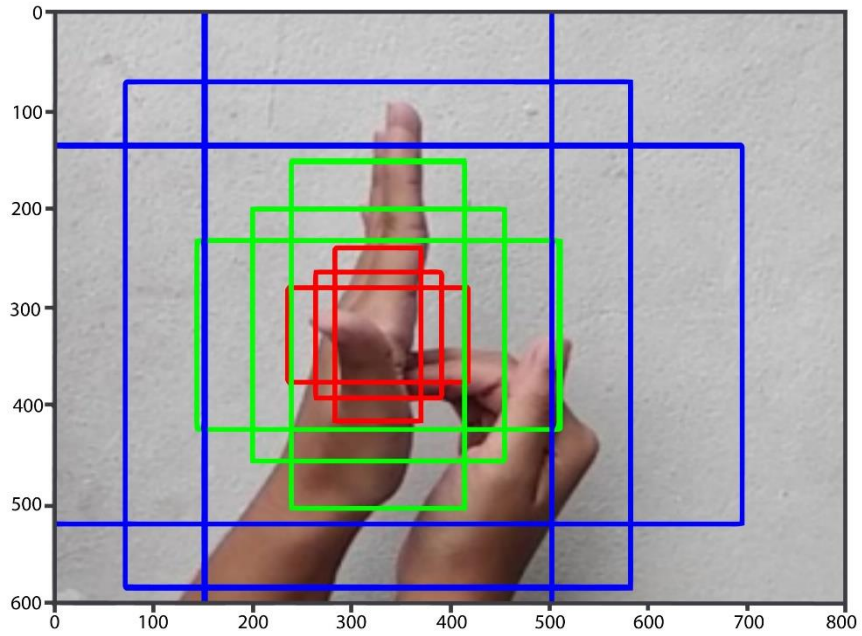
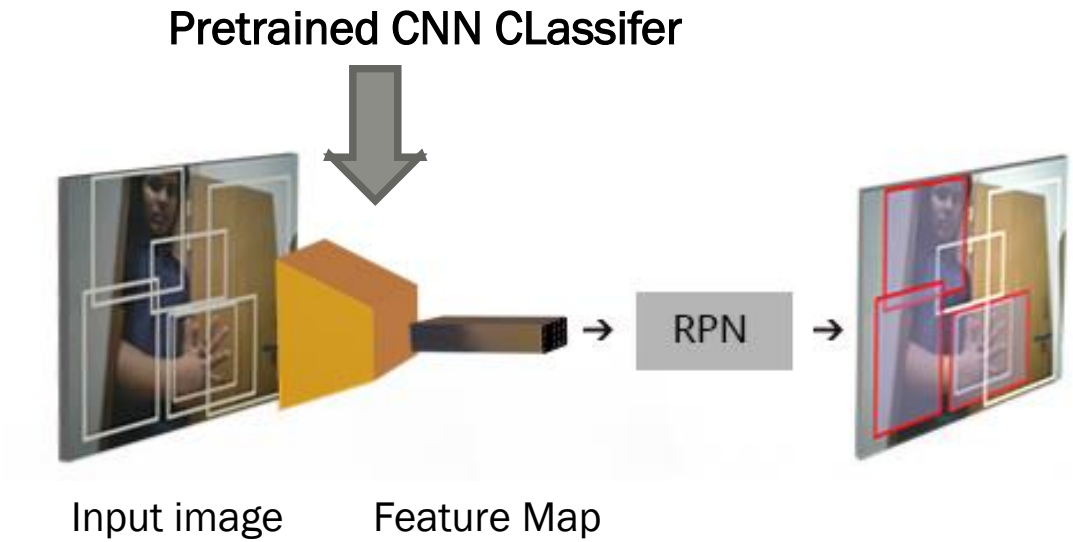


Fig : Anchor positions

Faster R-CNN Architecture



- RPN takes all the anchors.
- Mainly RPN does two modification for output
  - Finds the probability of objectness score of each anchor.
  - Bounding box regression for adjusting the anchors to better fit.

Fig : Process Of Faster R-CNN

Faster R-CNN Architecture

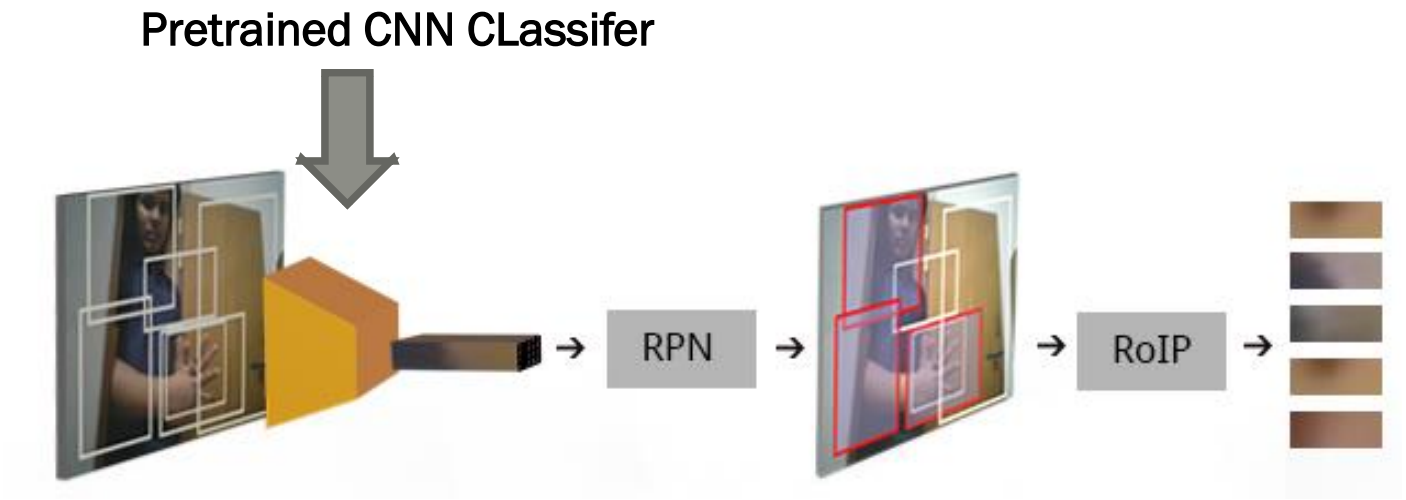


Fig : Process Of Faster R-CNN

- RoIP takes the proposals of RPN along with feature maps.
- Splits the input feature maps into a fixed number of roughly equal regions.
- Then apply the max pooling on every region
- Comes with a fix sized output

Faster R-CNN Architecture

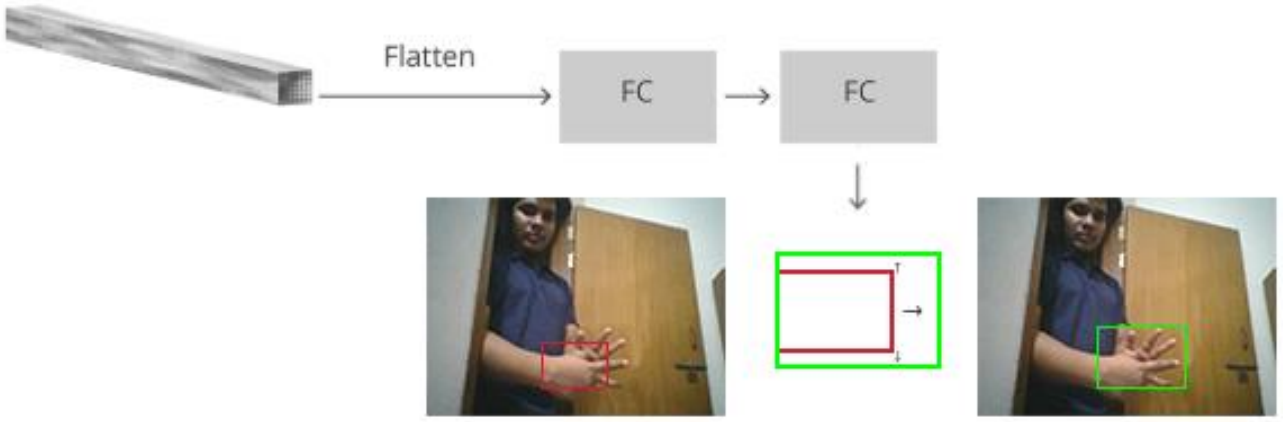


Fig : R-CNN, shaping the final output

- Proposals of RoIP (Flatten image) passes through Fully-connected layers.
- Classify the proposals into one class with better adjustment of bounding box

Faster R-CNN Architecture

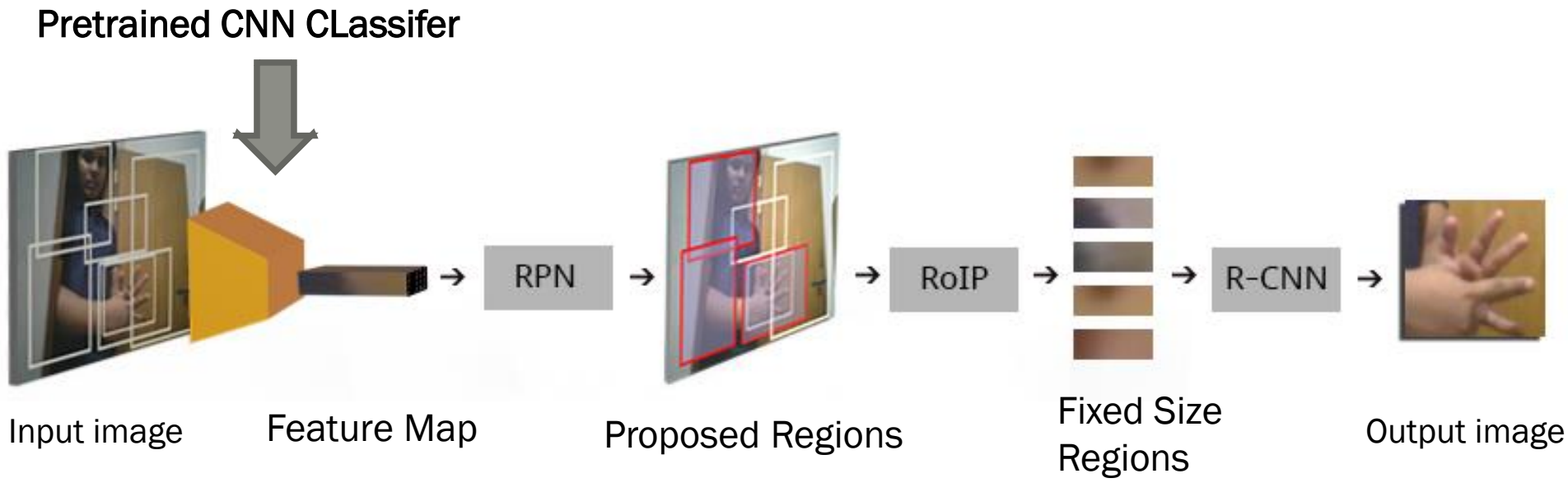
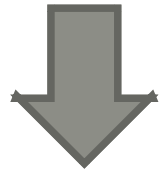


Fig : Architecture Of Faster R-CNN

Why Faster R-CNN ?



Work	R-CNN	Fast R-CNN	Faster R-CNN
Test time per image	50 seconds	2 seconds	0.2
Speed-up	1x	25x	250
mAP	66.0%	66.9%	66.9%

Table : Shows works of different Neural Network

Experiments and Results

Setup

- Faster-RCNN model based training in our module
- Training Set contains 80 percent of the images
- Test set contains 20 percent.
- Tensorflow-GPU V1.5 and cuda V9.0.
- Faster RCNN Inception V2 model.
- CPU from Intel R. CoreTM i7-7500U of 2.7 GHz
- GPU Nvidia 940mx with 4.00 GB and with 8.00GB memory on Windows 10 operating system.

Experiments and Results

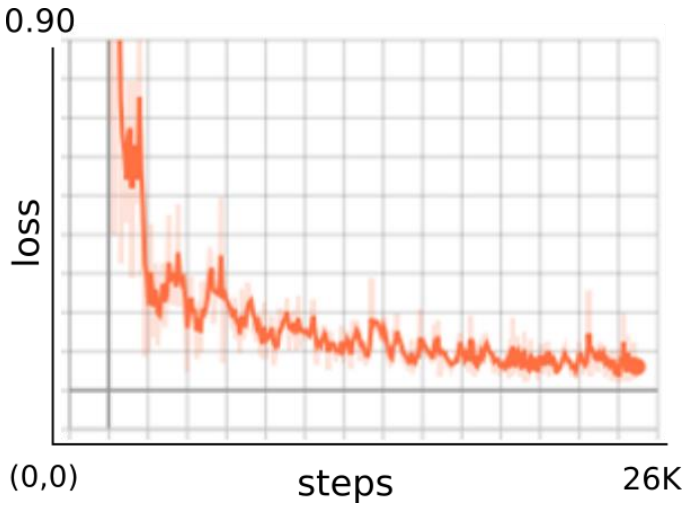
Training

- Initial training was for 10 classes with 50 images for each letter.
- Took about 12 hours
- 28000 iterations to train the model.
- Started with loss of 3.00, quickly dropped to 0.8. Stopped at 0.07

Experiments and Results

Training

- Started with loss of 3.00, quickly dropped to 0.9. Stopped at 0.07



Experiments and Results

Recognition

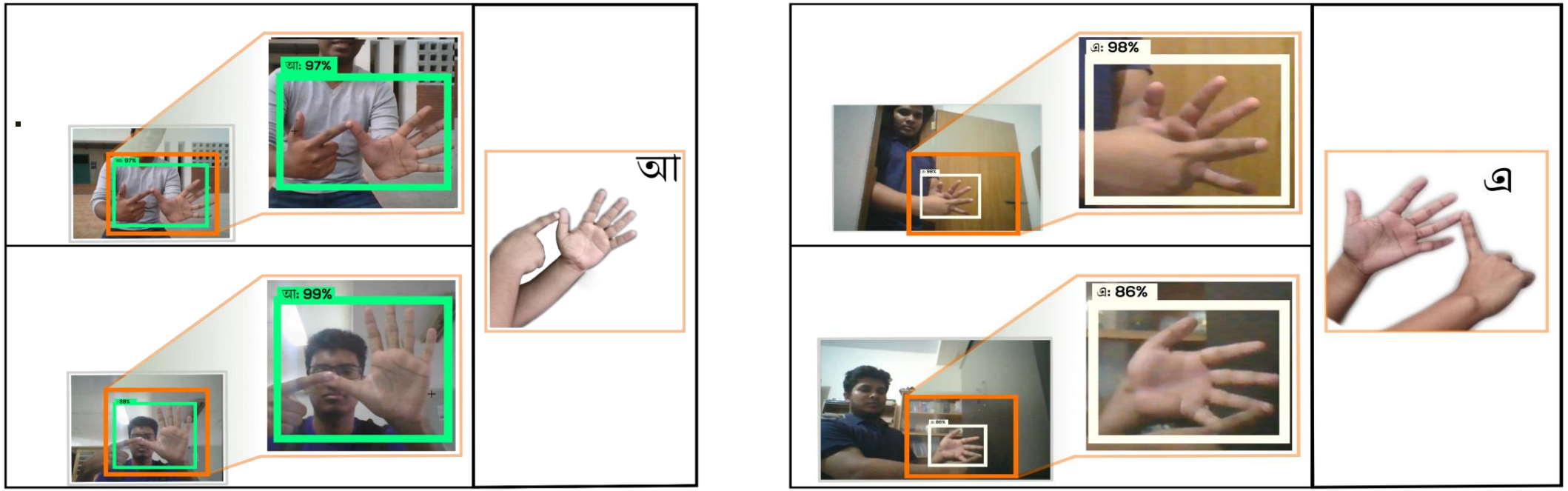


Fig : Detection with Faster R-CNN

Experiments and Results

Recognition

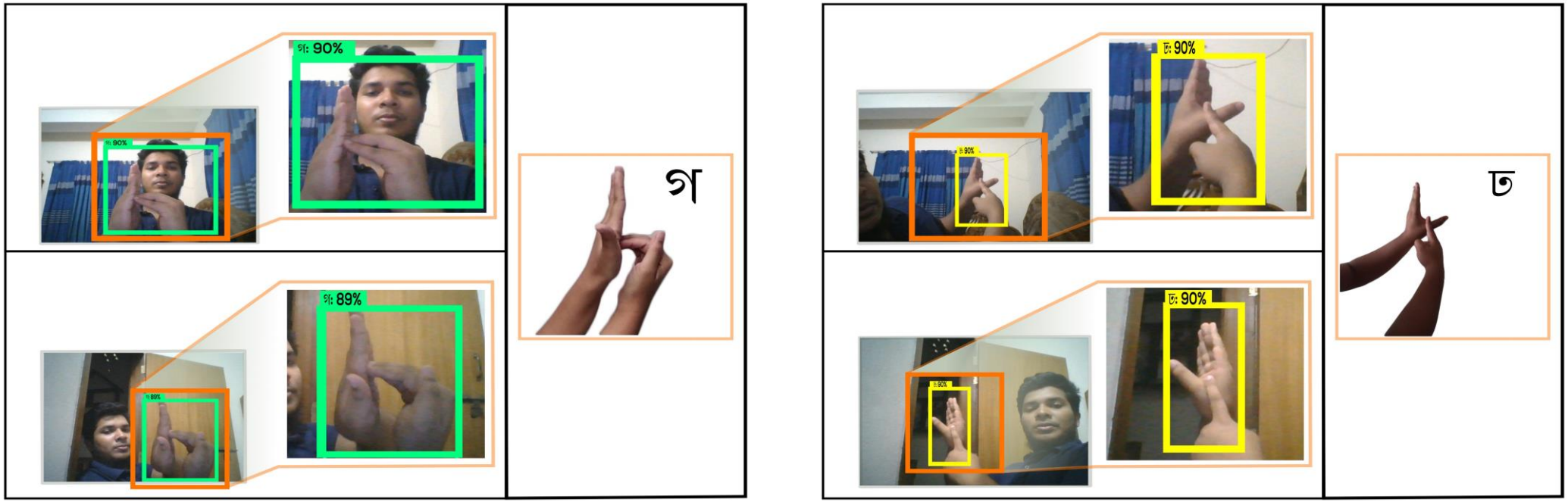


Fig : Detection with Faster R-CNN

Experiments and Results

Result

- Result was satisfactory.
- Though detections were not accurate for all letters, especially, for the letters with similar hand gestures.
- Later training for the other two classes with increased dataset gave extraordinary result.
- Accuracy was higher and it was detecting almost perfectly in various environment and different person.
- The confidence rate was 98% on average.

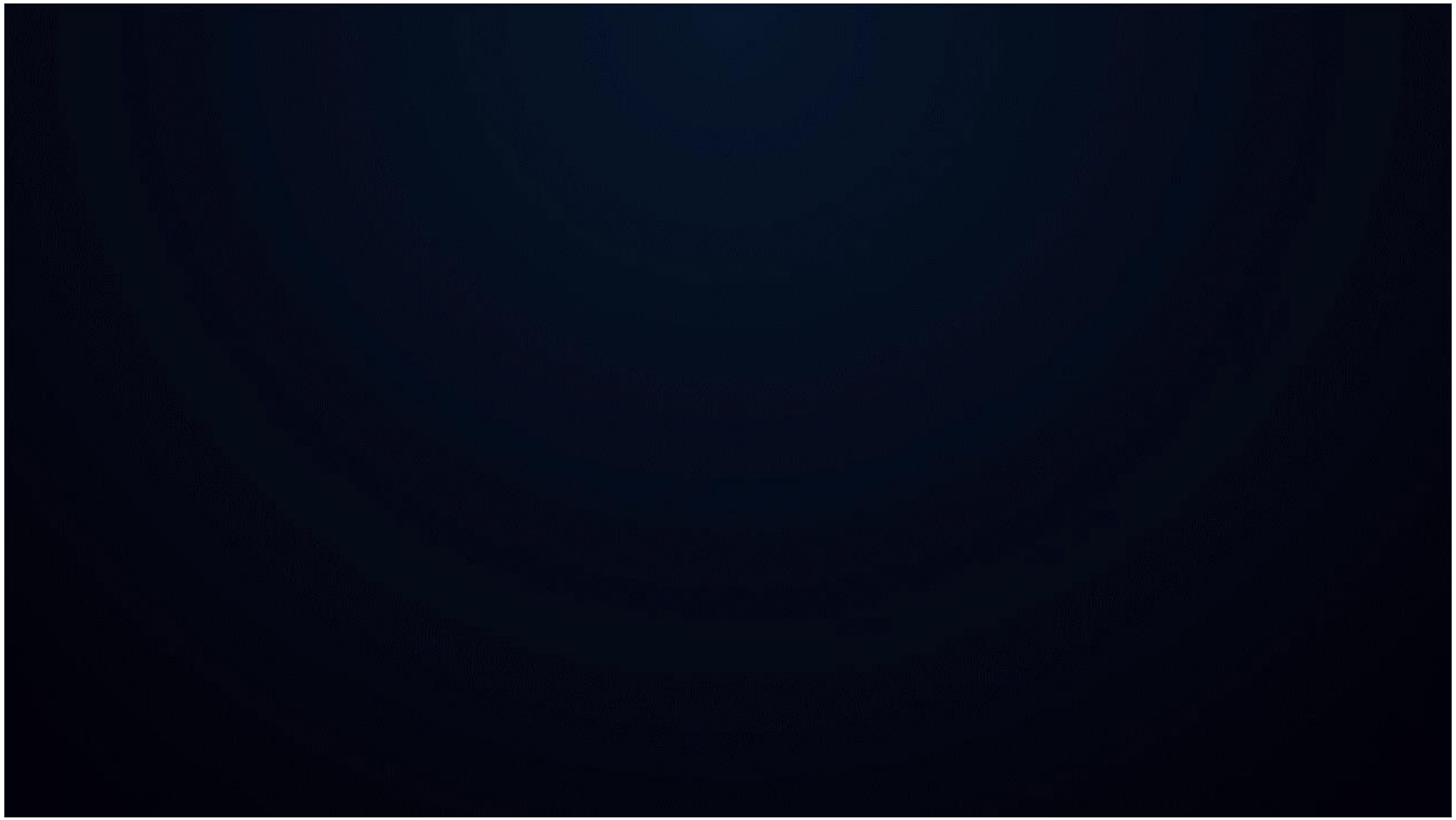
FUTURE PLAN

- Android Based Recognition System
- Real-Life Implementation For Signers
- Word and Sentence Recognition In Real Time

FUTURE PLAN

- Workshops on Deaf School For more Data Collection.
- Meeting with the officials of Centre for Disability & Development (CDD) to help us with our idea.
- Communicate with the signers to maintain the user friendliness of our system.

BdSL Recognition



Thank You

Any Questions?