16th IEEE International Conference on Industrial and Information Systems (ICIIS) 2021
09th to 11 th of December 2021, Faculty of Engineering, University of Peradeniya, Peradeniya, Sri Lanka.

Hands Off: A Handshake Interaction Detection and Localization Model for COVID-19 Threat Control

A.S. Jameel Hassan[1*], Suren Sritharan [2], Gihan Jayatilaka[1], Vijitha Herath[1], Parakrama B. Ekanayake[1], Roshan I. Godaliyadda[1], Janaka B. Ekanayake[1]

[1]Department of Electrical and Electronic Engineering, University of Peradeniya

[2]School of Computing and IT, Sri Lanka Technological Campus

[*]Presenter (correspondence: jameel.hassan.2014@eng.pdn.ac.lk)

16th (IEEE) International Conference on Industrial and Information Systems (ICIIS) 2021, Peradeniya, Sri Lanka.

Introduction

Problem Statement

- The COVID-19 pandemic
  - Biggest threat to global health having claimed over 5.25M lives → High infectivity.
- Curtailment measures against its spread
  - Need for Social distancing
  - Prevent human-human interactions
- Aim → Design a computer vision model to identify handshake interactions
  - Interaction localization in a multi-person setting

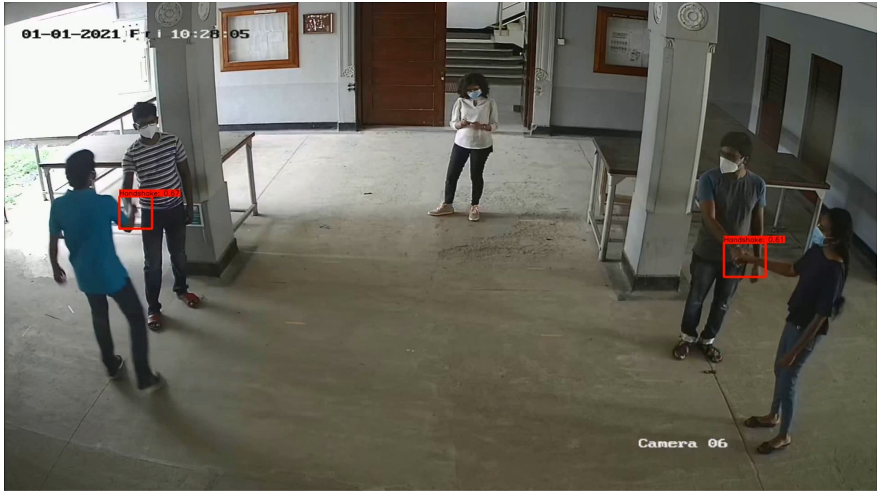

Figure 1: Handshake interaction localization.

Methodology

Overview

Designing a handshake interaction localization model:

- Dataset generation
  - Novel dataset created called the Shakes Dataset.
  - UTI dataset, with new annotations.

- Technique used
  - The YOLOv3 deep learning model was used.
  - Real time inference.

Methodology

YOLOv3 architecture

- Real time inference - 78 FPS
- Key features
    - Three heads
    - Regression of bounding box in end-to-end fashion
- Multi-level predictions to predict objects of different sizes.

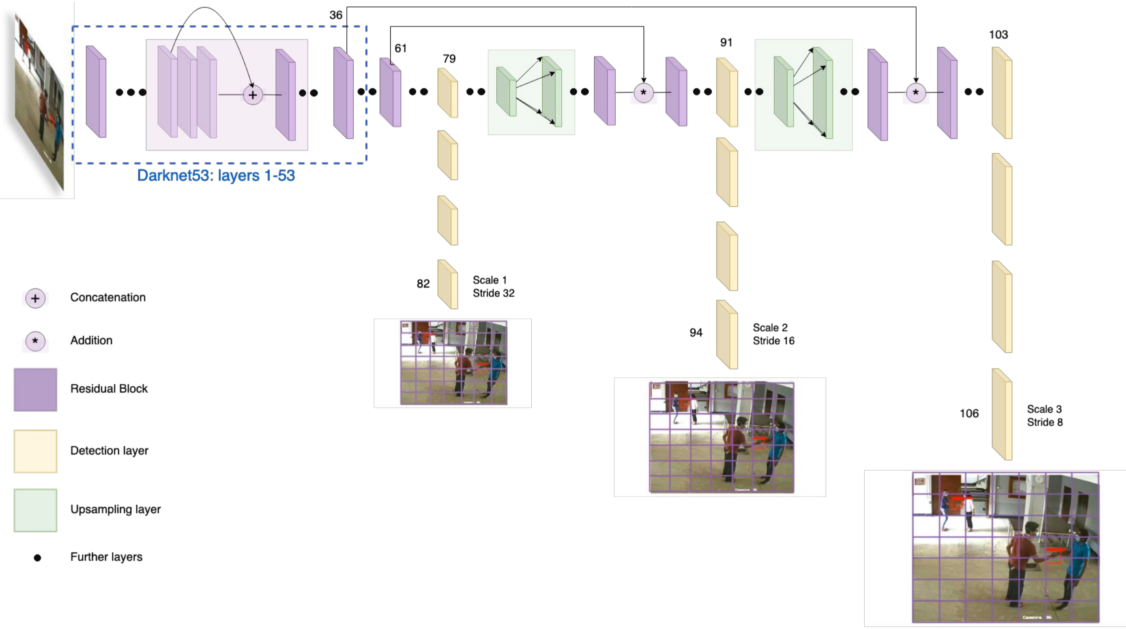

Figure 2: YOLOv3 architecture

Methodology

Datasets

- Lack of datasets for handshake interaction localization.
- New dataset created: Shakes dataset
  - Multi-person setting video sequences.
  - Multiple interactions at the same time.
  - Annotation localizing interaction, not the actors.
- UTI dataset for action recognition.
  - Ground truth was re-annotated for localization task.

Figure 3: Shakes dataset(top) and UTI dataset(bottom) with new annotation in red.

Methodology

Model training

- Backbone weights initialized using trained weights from ImageNet.

- Transfer Learning.
    - Network was initially trained on 3000 images of hands from the Open Images dataset.
    - Transfer learning was then done by training on the two handshake interaction localization datasets – UTI and Shakes.

- Train/Test split.
    - Out of the 20 video sequences, 17 videos from UTI dataset was used for training.
    - Out of 10 videos, 5 videos were used from the Shakes dataset.

Results

Analysis and comparison of implemented models:

- The popular Average Precision (AP) metric used in detection challenges.
    - AP is the area under the curve for the Precision vs Recall curve.

Dataset	**Average Precision**
Shakes Dataset	88.47%
UT-Interaction dataset	95.29%

Results

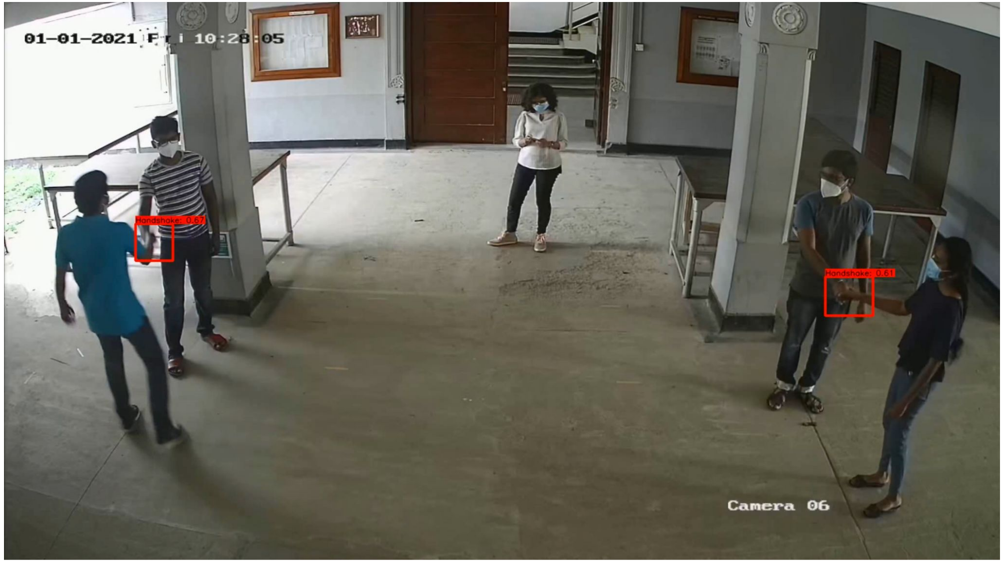

Figure 4: Shakes dataset example

Figure 5: UTI dataset example

Results: Notable Edge Cases

Figure 6: Fake handshake simulated by occlusion. Model does not predict as a handshake.

Figure 7: Model fails to identify both handshakes.

Results: Notable Edge Cases

Figure 8: Model mis-identifies an instance as a handshake

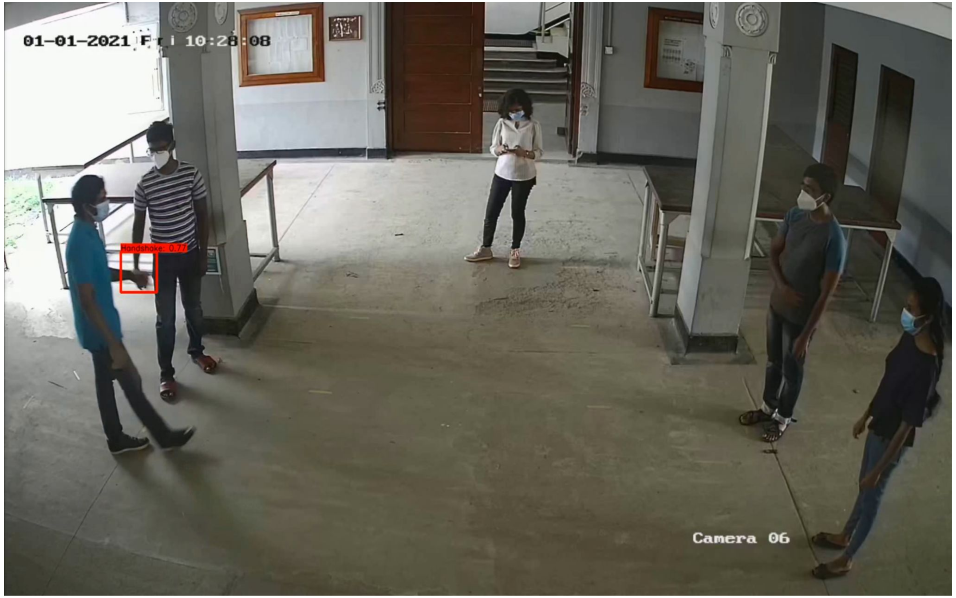

Figure 9: Model mis-identifies an occlusion as a handshake

Conclusion

- A computer vision based model to identify handshake interactions towards combating the spread of COVID-19 is proposed.

- The model is able to identify multiple interactions in a multi-person setting in realistic scenarios, in real time.

- The model deployed in public settings can help mitigate the spread of COVID-19 and can be enhanced by incorporating more interactions such as kisses and hugs.

Acknowledgements

- Funding

IDRC · CRDI
International Development Research Centre
Centre de recherches pour le développement international

Lewis Power, Singapore

- Volunteers who helped with creating the dataset.