Deep Learning Based Framework for Hand Keypoints Detection from a Monocular RGB Images

Thesis submitted in partial fulfillment
for the Award of Degree

Doctor of Philosophy

by

Purnendu Mishra

DEPARTMENT OF ELECTRONICS ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY
(BANARAS HINDU UNIVERSITY)
VARANASI - 221005

Roll No. 16091004

Year 2022

CERTIFICATE

It is certified that the work contained in the thesis titled ***"A Deep Learning Based Framework for Hand Keypoints Detection from a Monocular RGB Image"*** by ***Purnendu Mishra*** has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all requirements of Comprehensive Examination, Candidacy, and SOTA for the award of Ph.D. Degree.

Supervisor

Dr. Kishor Sarawadekar
Assistant Professor,
Department of Electronics Engineering,
Indian Institute of Technology (BHU) Varanasi,
Uttar Pradesh, INDIA 221005.

DECLARATION BY THE CANDIDATE

I, ***Purnendu Mishta***, certify that the work embodied in this Ph.D. thesis is my own bonafide work carried out by me under the supervision of ***Dr. Kishor Sarawadekar*** from ***July 2016*** to ***Dec 2022*** at ***Department of Electronics Engineering***, Indian Institute of Technology (BHU) Varanasi. The matter embodied in this thesis has not been submitted for the award of any other degree/diploma. I declare that I have faithfully acknowledged and given credits to the research workers wherever their works have been cited in my work in this thesis. I further declare that I have not willfully copied any other's work, paragraphs, text, data, results, *etc.* reported in journals, books, magazines, reports, dissertations, theses, *etc.*, or available at websites and have not included them in this thesis and have not cited as my own work.

Date: 22/12/2022

Place: Varanasi

Purnendu Mishra

(Purnendu Mishra)

CERTIFICATE BY THE SUPERVISOR

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

(Dr. Kishor Sarawadekar)
Assistant Professor,
Dept. of Electronics Engineering,
Indian Institute of Technology (BHU) Varanasi

Signature of Head of Department

विभागाध्यक्ष/Head
इलेक्ट्रॉनिकी अभियांत्रिकी विभाग
Department of Electronics Engineering
भारतीय प्रौद्योगिकी संस्थान
Indian Institute of Technology
(बनारस हिन्दू यूनिवर्सिटी)
(Banaras Hindu University)
वाराणसी/Varanasi-221005

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis: A Deep Learning Based Framework for Hand Keypoints Detection from a Monocular RGB Image

Name of the Student: Purnendu Mishra

Copyright Transfer

The undersigned hereby assigns to the Institute of Technology (Banaras Hindu University) Varanasi all rights under copyright that may exist in and for the above thesis submitted for the award of the *Doctor of Philosophy*.

Date: 22/12/2022

Place: Varanasi

Purnendu Mishra

(Purnendu Mishra)

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provided that the source and the Institute's copyright notice are indicated.

Dedicated to

Almighty Mahadev,

Shri Kashi Vishwanath

ACKNOWLEDGEMENT

With the completion of the thesis, I wish to express my extreme gratitude to my research supervisor, Dr. Kishor Sarawadekar, for his continuous guidance, encouragement, and critical evaluation throughout this research work. His constant untiring supervision at the vital stages of this work has factored in the smooth completion of the thesis in the most appropriate period of time.

Moving forward, I wish to extend my sincere gratitude to the entire Indian Institute of Technology (BHU), Varanasi, especially the Head of the Department(HoD), Department of Electronics and Engineering, for their encouragement and insightful comments throughout this research work.

This research work would have been intense labor without the constant support of my friends, Dr. Mumtaz Ali Ansari, Shyam Gopal Yadav, Tanushree Meena, and seniors Dr. Gourav Modanwal, Dr. Subiman Chatterjee, who was always ready to guide me throughout my journey.

During this research, my colleagues, Sumit Kumar Yadav, and Bharat Bhushan extended their unconditional support, which contributed to the effortless completion of this work.

Apart from all the people directly contributing to the research, I also want to draw attention to the ones who have been a pillar of strength, even without being directly associated with the investigation. I would like to acknowledge and express my sincere gratitude to the lab Senior Technician, Mr Vinod Kumar Verma, for being constantly around and being just a call away in case of any discrepancy.

This research work has been an incredible journey where mentors like Dr. Kishor Sarawadekar have installed his belief and have motivated me through the various research paper submissions. He relentlessly devoted himself to my betterment and the entire department.

Apart from the academic guidance, he has been a GUIDE in a true sense who dilated every situation and has always advised the most appropriate solutions to every

obstacle.

I also wish to acknowledge the incredible guidance and support of the entire institute, especially during the pandemic years. I, in particular, am grateful for all the opportunities that came my way.

Along with all the academic guidance, I wish to express my sincere regards to my parents, Nirbhaya Kumar Mishra, Ranjana Mishra, Bijay Kumar Mishra and Anju Mishra. My siblings Navendu Mishra and Kalindi Mishra, along with my Brother-in Laws Dhananjay Jha and Alok Mishra. They all have been instrumental in keeping my mental strength intact and always ensure positivity and zeal throughout this journey.

I want to give special thanks to my wife, Minakshi Mishra, who has stood by me through all my travails, my absences, my fits of pique and impatience. She gave her support and help, discussed ideas, and prevented several wrong turns. She also supported the family during much of my research phase.

Last but not least, I would like to thank God Almighty for bestowing his power and positivity which was key to this entire research process.

(Purnendu Mishra)

Contents

List of Figures

List of Tables

List of Symbols

Symbol	Description
σ	Distance threshold in PCK calculation
χ	Chi function
$\mathbb{N}$	Number of samples
$\mathbb{R}$	Real number
w	Width of the bounding box
h	Height of the bounding box
S	RGB image dimension
C	Probability of a grid cell
M	Number of grids
U	Numbers of grid rows
V	Number of grid columns
I	RGB image
δ_p	Grid cell probability threshold
δ_m	Multi-label classification probability threshold

Abbreviations

Abbreviation	Description
ADAM	Adaptive Moment Estimation
AI	Artificial Intelligence
AR	Augmented Reality
ASL	American Sign Language
ANN	Artificial Neural Network
CAGR	Compound Annual Growth Rate
CDF	Cumulative Distribution Function
CIFAR	Canadian Institute of Advanced Research
CLR	Cylic Learning Rate
CMU	Carnegie Mellon University
CNN	Convolutional Neural Network
DL	Deep Learning
DNN	Deep Neural Network
DOF	Degree of Freedom
F1Score	Harmonics Precision-Recall Mean
FCN	Fully Connected Network
FFNN	Feed-Forward Neural Network
FN	False Negative
FP	False Positive
FPHA	First Person Hand Action
FPN	Feature Pyramid Network
GAP	Global Average Pooling
GPU	Graphics Processing Unit
GUI	Graphical User Interface
HCI	Human Computer Interaction
HGR	Hand Gesture Recognition
IOU	Intersection Over Union

Abbreviation	**Description**
LR	Learning Rate
MHP	Multi-view Hand Pose
ML	Machine Learning
MLP	Multi-Layered Perceptron
MR	Mixed Reality
NN	Neural Network
NUI	Natural User Interface
PC	Personal Computer
PCK	Probability of Correct Keypoint
POV	Point of View
POLY	Polynomial Learning Rate
ReLU	Rectified Linear Unit
RHD	Rendered Hand Pose
RNN	Residual Neural Network
ROC	Receiver Operating Characteristic
ROI	Region of Interest
SGD	Stochastic Gradient Descent
SGDR	Stochastic Gradient Descent with warm restart
SLR	Sign Language Recognition
SSD	Single Shot Detector
STB	Stereo Hand Pose Tracking Benchmark
TN	True Negative
ToF	Time of Flight
TP	True Positive
VGG	Visual Geometry Group
VR	Virtual Reality
WHO	World Health Organization
YOLO	You Only Look Once
YOLSE	You Only Look what You Should See