



## Hand Gesture Recognition Using Python

Prof Pragati Akre, Abinaya J, Viranch R, Tanmay M and Om W

T.E., Department of Artificial Intelligence and Data Science  
SIES Graduate School of Technology, Nerul, Navi Mumbai.

**Abstract:** This project presents a hand gesture recognition system that uses Python, Mediapipe, Pytesseract, PyAutoGUI, and OpenCV to enable hands-free interaction with digital devices. It has features like mouse control, PowerPoint control, media player, and whiteboard functionalities, the system utilizes real-time hand tracking and gesture recognition for an intuitive user experience. It can also convert written handwriting to text. Additionally, the handwriting model is trained on a dataset of labeled images and uses transfer learning to improve its accuracy. This project enhances human-computer interaction by combining advanced computer vision techniques with versatile Python libraries, making it valuable for presentations, media control, and collaborative whiteboarding. This project could be used by handicapped people, could be used in school or office presentations, medical sectors where hygiene is important or anywhere where a touch free environment is required.

**Keyword:** Hand gesture, Computer vision, Mediapipe, OpenCV

### I. INTRODUCTION

In the present day, as technology is evolving, there is a growing demand for a hands-free interface for hygiene, any disability or speedy work purposes. And gesture control technology has not only proven successful but also provided a possible solution by enabling users to connect themselves to devices via natural hand gestures. Hence, to address the issue this project aims to contribute to a touchless era. It describes the innovative technique of controlling media players, power points, whiteboards etc. with the help of hand signs. It also has numerous other features like whiteboard, handwriting to text and so on. The technology leverages computer vision in combination with depth-sensing cameras to create an immersive environment where users can effortlessly play, pause, adjust the volume, present their work, draw, explain, through simple gestures. This system uses mediapipe, python, opencv and machine learning algorithms for gesture recognition that provides immediate understanding of hand movements with the degree of accuracy and dependability. In addition to that, the interface does not limit itself to media control but allows users to interact with many different other apps and functions so now they benefit from a rich and immersive experience. This project has immense power to turn the paradigm of user

interaction and to make the interface more accessible and convenient by opening new possibilities.

### Literature Review

A literature survey on controlling media players and other functions through hand gestures reveals a growing interest in this innovative technology. Several research papers have delved into this area. Some of the earlier systems are studied as follows:

#### 1. Controlling Media Player with Hand Gestures:

A paper written by Parshav Maloo (2022) looks at the assembling of a hand gesture recognition system for the media player direction and UI element gesture recognition and control. It relies on the sphere of the real communication, including both the limits and chances. This sphere can be implemented in different fields.[1]

#### 2. Hand Gesture Recognition:

In the study published recently by Swetha Kotavenuka, Harshitha Kodakandla, Nimmakayala Sai Krishna, Subba Rao (IJRASET-2023), availability of hand gesture recognition devices is addressed considering machine learning approaches and computer vision for predicting gestures from array of frames, with edge detection having more to do with improved accuracy with well-tun.[2]

#### 3. A systematic review on hand gesture recognition techniques, challenges and applications:

The research paper Lucy Fadiga, Mais Yasen, and Jusoh Shaydah (Peerj-2019) indicates a comprehensive review of hand gesture recognition issues including acquisition techniques, feature extraction, transfer models classification and environment challenges. It presents the use of the sEMG sensors, focusing on the contribution of the ANN as the classifier used mostly, the preference for hand gesture while designing the applications of sign language, the impacts of the background color on classification accuracy and the common existence of overfitting in the datasets.[3]

#### 4. Media Controlling Using Hand Gestures:

The article published by R H Sapat, Shruti Tibhe, Ashwini Joshi, Aishwarya Warulkar, Aishwarya Sonawane, Miss T U Ahirrao (available on IJCRT.org – 2023) is about the development of a hand gestures recognizing system for controlling media player through computer vision technique, which offers beneficial advantages for people.[4]

## 5. Hand Gesture Recognition Using Automatic Feature Extraction and Deep Learning Algorithms with Memory:

The research paper published by Robail Yasrab, Md Mostafa, Kamal Sarker, Rubén E Nogales, Marco E Benalcázar (MDPI-2023) explores hand gesture recognition using manual and automatic feature extraction techniques with a dataset collected from 56 volunteers using the Leap Motion Controller, evaluating different classifiers and proposing a hybrid model for improved recognition accuracy. These research works collectively demonstrate the advancements and potential of using hand gestures to control media players, highlighting the benefits of intuitive and hands-free interactions with technology.[5]

## II. PROPOSED SYSTEM

We have developed a hand recognition system that can control media, power points, whiteboard, and normal mouse movements as well. The system uses Python, MediaPipe, Libraries like Pytesseract, PyAutoGUI and OpenCV etc. Additionally, the system includes a feature that allows to convert drawn text into typed text. The system also includes a frontend by which user can select the desired action.

### Dataset:

To train our handwriting model, we used a dataset of images labeled with the correct class of handwriting. The dataset contains images of several handwriting. We used approximately 3000 images for training and 750 images for validation.

### Model Architecture:

We developed a Hand Recognition System using Python, integrating libraries such as Mediapipe, Pytesseract, PyAutoGUI, and OpenCV. The architecture includes:

1. **Hand Tracking Module:** Utilizes the Mediapipe library for real-time hand tracking and landmark detection. Employs OpenCV for image processing to enhance hand gesture recognition accuracy.
2. **Gesture Recognition Module:** Incorporates machine learning techniques for hand gesture classification. Custom-trained model using a labeled dataset of hand gesture images.
3. **Text Extraction Module:** Integrates Pytesseract for optical character recognition (OCR) to convert handwritten text to machine-readable text.
4. **Control Module:** Uses PyAutoGUI for mouse control and specific algorithms for PowerPoint, media player, and whiteboard functionalities.
5. **Transfer Learning for Handwriting Recognition:** Trains a handwriting recognition model using transfer learning on a labeled dataset, improving accuracy. Adapts pre-trained models and fine-tunes them on the specific handwriting dataset.
6. **Application Scenarios:** Designed for scenarios such as presentations, medical sectors, and environments requiring landmarks using the Mediapipe library, and interprets specific hand gestures to trigger keyboard events. The system then processes the hand gestures, counting the number of raised fingers, and simulates corresponding keyboard inputs for play/pause, volume up/down, and forward commands.

touch-free interaction. Enhances human-computer interaction through a combination of computer vision and versatile Python libraries.

Software used:- HTML, CSS for frontend .VS studio is used to make webpage. Flask is used for the integration of frontend and python scripts. Machine Learning for handwriting recognition

## III. SYSTEM IMPLEMENTATION

This integrated system incorporates multiple features for diverse interactive applications, each utilizing a combination of Python libraries to enhance user interaction. Firstly, the system employs the cvzone library, along with OpenCV functionalities, to facilitate hands-free control over presentations. Hand gestures, such as swiping left or right and lifting the index finger to annotate, are detected and translated into commands, providing an intuitive method for managing slides. Secondly, the system leverages the cv2, numpy, mediapipe, and deque libraries to create an interactive drawing application. Hand gestures are interpreted to draw lines in various colors on a virtual canvas, offering users a dynamic and engaging drawing experience. It also has a model trained for detecting handwriting and converting into text. The third feature utilizes OpenCV, Mediapipe, and PyAutoGUI for gesture-based mouse control. Hand landmarks are detected, and specific gestures trigger predefined keyboard commands, providing hands-free computer control. The count\_fingers function calculates the raised fingers, ensuring accurate gesture recognition. Finally, the fourth feature integrates OpenCV, Mediapipe, and PyAutoGUI to create a gesture-controlled media interface, enabling users to navigate videos and applications intuitively. It can also be used for controlling volume and brightness as well. Currently we have 4 features:

**Feature 1(PPT Control):** It utilizes the cvzone library, OpenCV, and Mediapipe for hands-free control of presentations. It employs hand tracking and gesture recognition to interpret specific movements, enabling functionalities such as swiping to navigate slides and lifting the index finger for annotations.

**Feature 2(White board):** It leverages cv2, numpy, and deque to create an interactive drawing application. The script captures webcam frames, detects hand landmarks using Mediapipe, and interprets gestures to draw lines in various colors on a virtual canvas. The dynamic color-changing and clearing functionalities enhance the drawing experience.

**Feature 3(Mouse Control):** It utilizes OpenCV, Mediapipe, and PyAutoGUI to enable gesture-based mouse control. It captures webcam frames, detects hand landmarks, and interprets gestures to trigger keyboard events, offering hands-free computer control. The introduction of a time-based mechanism ensures accurate and deliberate gesture recognition.

**Feature 4(Media Control):** It utilizes OpenCV and PyAutoGUI to create a gesture-controlled media interface. It captures video frames from the camera, detects hand

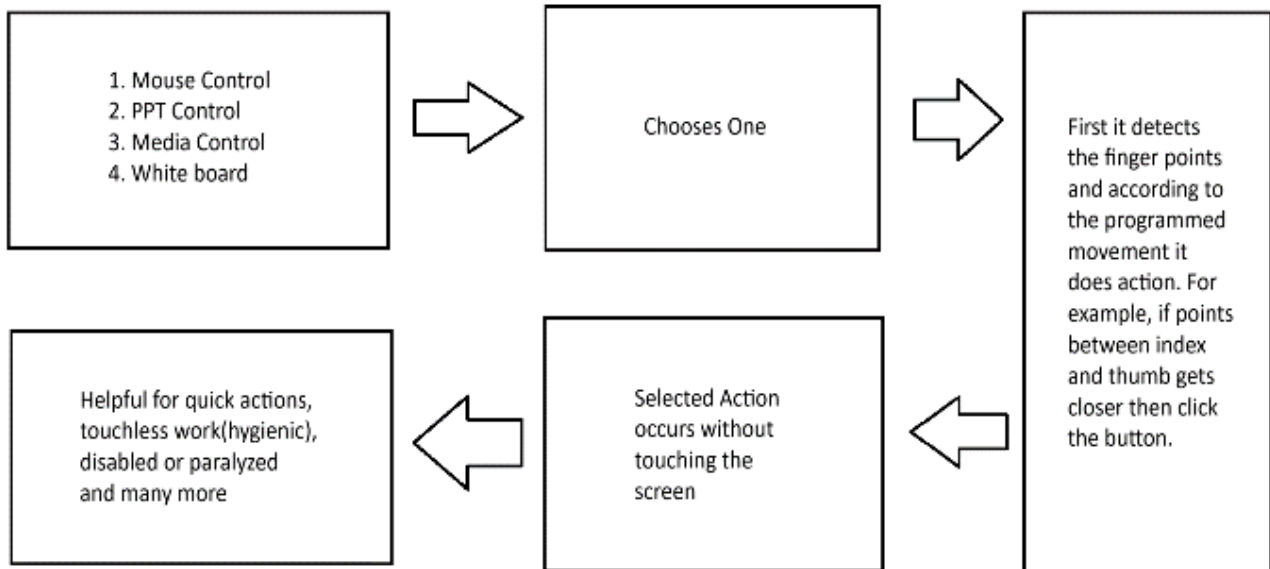


Fig 1: Block diagram

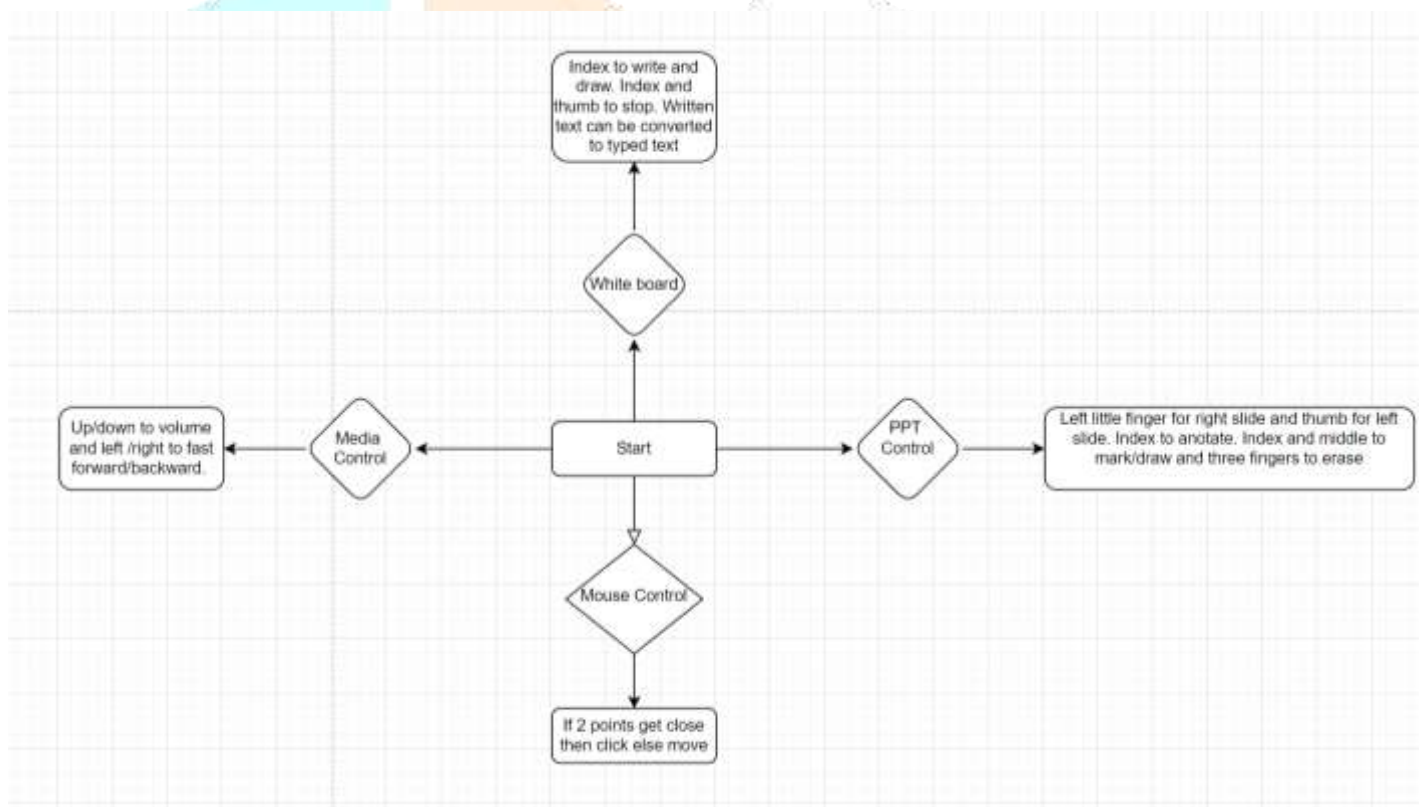


Fig 2: Architecture

### IV. ANALYSIS AND RESULTS

Proposed system is implemented with the various technologies, the screen shots exhibit the layouts of its components.

```

120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

Fig 3: Snapshots of code in Python

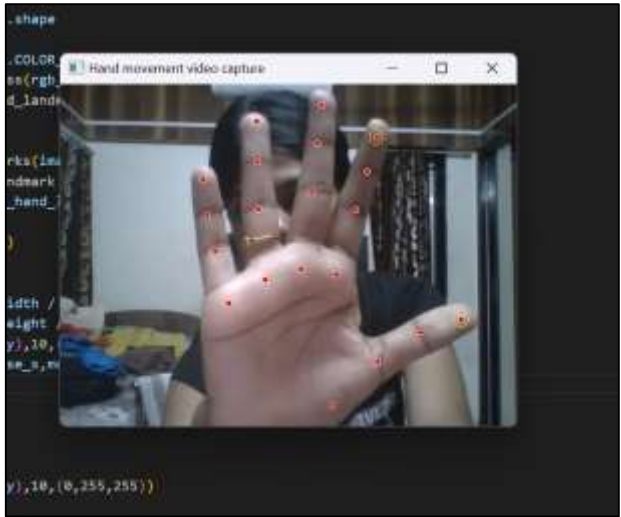


Fig 5: Mouse Control



Fig 6: PPT Control- Going to left Slide

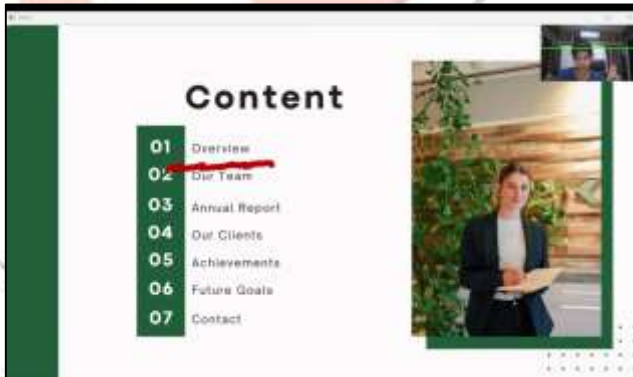


Fig 7: PPT Control- Annotating and Marking

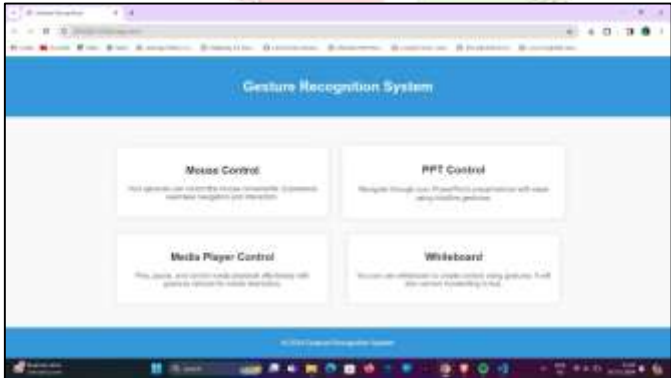


Fig 4: Software

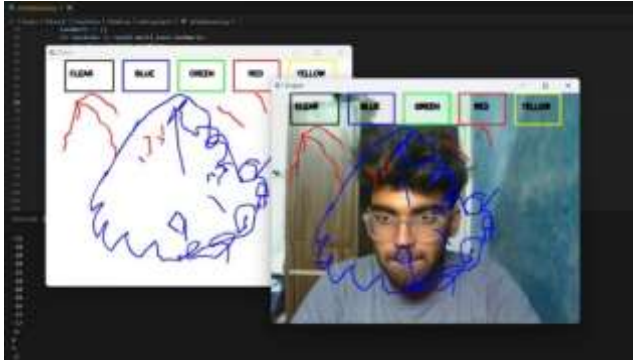


Fig 8: White board



Fig 9: White board (Detects Text)

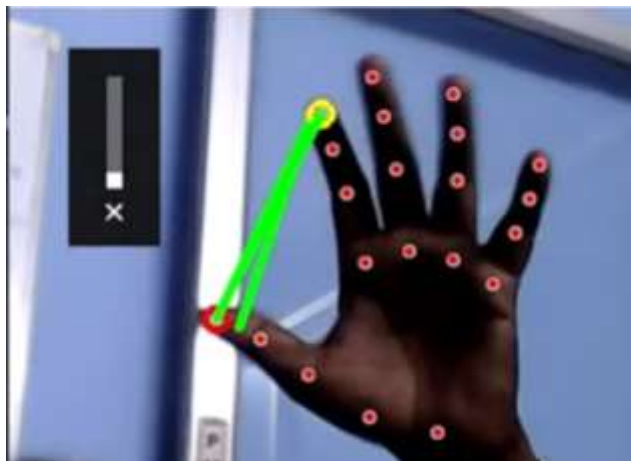


Fig 10: Media Player (Volume Control)



Fig 11: Media Player (Pause the video)

## V.CONCLUSION AND FUTURE WORK

In conclusion, our project on Hand Gesture Recognition has successfully demonstrated its practical applications in enhancing user interaction with various systems. The implementation of this technology as a Media Controller, ppt Controller, and Mouse Controller using the Hand Gesture System, along with real-time detection through Python and OpenCV, has proven its effectiveness in making system interaction more seamless.

Looking ahead, the potential applications of this technology extend beyond the realms of conventional user interfaces. The integration of hand gesture recognition could play a pivotal role in diverse fields such as medical rehabilitation and virtual reality, promising advancements in how individuals interact with technology for improved healthcare and immersive experiences.

The versatility of this technology opens up possibilities for integration into Home Control systems, healthcare systems, gaming zones, automobiles, and robotics. The prospect of using hand gesture recognition to facilitate communication

between users and devices holds great promise for creating intuitive and efficient interfaces across various domains.

As we move forward, there is ample room for improvement in the system. Enhancements in recognition accuracy and the integration of advanced controllers can further refine the user experience and broaden the scope of applications. Our project serves as a foundation for future developments in the field, paving the way for more sophisticated and widespread use of hand gesture recognition technology.

## REFERENCES

- [1] Parshav Maloo, 2022, Controlling Media Player with Hand Gestures, INTERNATIONAL JOURNAL OF RESEARCH & TECHNOLOGY (IJERT) Volume 11, Issue 11 (November 2022), <https://www.ijert.org/controlling-media-player-with-hand-gestures>
- [2] Kotavenuka, S., Kodakandla, H., Krishna, N. S., & Rao, D. S. P. V. S. (2023). Hand Gesture Recognition. International Journal for Research in Applied Science and Engineering Technology, 11(1), 331–335. <https://doi.org/10.22214/ijraset.2023.48557>
- [3] Yasen, M., & Jusoh, S. (2019). A systematic review on hand gesture recognition techniques, challenges and applications. PeerJ Computer Science, 5, Article e218. <https://doi.org/10.7717/peerj-cs.218>
- [4] Shruti Tivhe ,Media Controlling Using Hand Gestures (2023), [https://ijcrt.org/viewfull.php?&p\\_id=IJCRT2312800](https://ijcrt.org/viewfull.php?&p_id=IJCRT2312800)
- [5] Nogales, R. E., & Benalcázar, M. E. (2023). Hand Gesture Recognition Using Automatic Feature Extraction and Deep Learning Algorithms with Memory. Big Data and Cognitive Computing, 7(2), 102. <https://doi.org/10.3390/bdcc702102>