**Mini Project Report**

on

**Background Removal in Video**

**Submitted by**

Ms. Aditi Subhash Nimbarte, 02

Ms. Pranjali Dharmadip Jambhulkar, 18

Mr. Sanket Suhas Tinkhede, 63

**Guided by**

Prof. Amol Gaikwad



**Department of Information Technology**

**Yeshwantrao Chavan College of Engineering, Nagpur**

*(An Autonomous Institute, Affiliated to RTM Nagpur Universiy, Nagpur. )*

**Session: 2021-2022**

Yeshwantrao Chavan College of Engineering, Nagpur

**Department of Information Technology**

**Session: 2021-22**

****

**Certificate**

This is to certify that Ms. Aditi Subhash Nimbarte, Ms. Pranjali Dharmadip Jambhulkar, Mr. Sanket Suhas Tinkhede has completed a **Mini Project** course titled “***Background Removal in Video****”* undertaken at **Zoho Industries** towards the partial fulfillment of requirements for the B.E. seven semester of Information Technology.

Submitted by

**Ms. Aditi Subhash Nimbarte (02)**

**Ms. Pranjali Dharmadip Jambhulkar (18)**

**Mr. Sanket Suhas Tinkhede (63)**

**“Prof. Amol Gaikwad”**

Project Guide

**Prof. B. U. Bawankar Dr. R. C. Dharmik**

Project Co-ordinator(IT) HOD (IT)

**DECLARATION**

We certify that

1. The work contained in this project has been done by me under the guidance of my supervisor(s).
2. The work has not been submitted to any other Institute for any degree or diploma.
3. Where others' ideas or words have been included, I have adequately cited and referenced the original sources.
4. We have followed the guidelines provided by the Institute in preparing the project report.
5. We have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

**ACKNOWLEDGEMENT**

We would like to thank our guide **Prof. Amol Gaikwad** for thorough guidance in the project. We are extremely grateful and indebted to them for their expert, sincere, valuable guidance and encouragement which was of immense help to us.

We would like to express our sincere gratitude to **Dr. R. C. Dharmik,** Head, Department of Information Technology for her constant encouragement towards the successful completion of our work.

We wish to express our sincere thanks to **Dr. U.P. Waghe**, the Principal of our college for providing us with all the necessary facilities and the infrastructure without which we would not have been able to complete our project successfully.

We would also like to thank our Project Coordinator **Prof. B. U. Bawankar** for their continuous guidance owing to which the project could take shape.

We would like to thank all the staff members of our college and technicians for their help in making this project a successful one. Last, but not the least, we would like to thank all the faculty members and non-teaching staff members who helped us despite of their busy schedule.

**TABLE OF CONTENTS**

|  |  |
| --- | --- |
| **Title** | **Page No.** |
| **Title page** | **i** |
| **Certificate of Approval** | **ii** |
| **Declaration** | **iii** |
| **Acknowledgement** | **iv** |
| **List of Tables** | **vi** |
| **List of Figures** | **vi** |
| **Abstract** | **vii** |
| **Chapter 1. Introduction** | **01** |
| **1.1 Advantages………………………** | **02** |
| **1.2 Features………………………….** | **02** |
| **Chapter 2. Literature review** | **03** |
| **2.1 Research Papers………………..** | **04** |
| **2.2 Patents…………………………..** | **08** |
| **Chapter 3. Work Done** | **10** |
| **3.1 Architecture And Design……….** | **11** |
| **3.2 Tools Used……………………….** | **12** |
| **3.3 Process Followed………………..** | **13** |
| **Chapter 4. Result and Discussion** | **15** |
| **Chapter 5. Future Scope and Conclusions** | **20** |
| **Chapter 6. Literature cited/References** | **22** |
|  |  |

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| **Table no.** | **Table Name** | **Page no** |
| 2.2.1 | Table of Patents | 09 |

**LIST OF FIGURES**

|  |  |  |  |
| --- | --- | --- | --- |
| **Figure no.** | **Figure name** | **Page no.** | |
| 3.1 | Activity Diagram of background removal process | 11 |
| 4.1 | User interface | 16 |
| 4.2 | Real-time background removal. | 16 |
| 4.3 | Select Image background change option | 17 |
| 4.4 | Background blur image | 17 |
| 4.5 | Background desaturate image | 18 |
| 4.6 | Background transparent image | 18 |
| 4.7 | Background Change image | 19 |

**ABSTRACT**

We live in the era of video calls. Conducted over the internet and using whatever camera that comes with your laptop or computer, we broadcast our lives to our classmates, coworkers, and families.

Sometimes, though, we don’t want to broadcast our space. Our office, like many others, has a few perennial pieces of clutter. Sometime we have creative or artistic background on the wall behind us, which doesn’t always scream professionalism.

As a result, video meeting software includes a feature to hide our background, usually behind an image of our choice. While most don’t think of it much, the actual task of determining what determines the foreground and the background in an image is hardly trivial.

**Chapter 1. Introduction**

**Chapter 1: Introduction**

The background of the video needs to be modified as there are so many other interruptions in the background or the background color doesn’t suit the person due to which background or the color needs to be modified. So, we use the real-time background replacement technique to substitute the backgrounds and add replace them with the desired content.

The primary motivation behind this project is to remove unwanted objects from video and fill in the gap in a visually plausible manner. Common approaches require manual identification of the area to be removed, however we aim to be able to automatically detect the target area and create an end-to-end pipeline for background removal.

**1.1 Advantages:**

* Attention.
* Decision speed.
* Brand consistency.
* Compliance.
* Performance.
* Flexibility.

**1.2 Features:**

* Personal and professional use
* Saves time and money

**Chapter 2. Literature Review**

**Chapter 2: Literature Review**

**2.1 Research Papers**

**2.1.1 A background extraction and shadow removal**

Published in: 2014 International Conference on Machine Learning and Cybernetics.

Authors: Yinshi Qin, Shuifa Sun, Xianbing Ma

Shadow removal has always been one of the hot research topics in the fields of computer vision. Recently the ViBe (Visual background extraction) foreground extraction algorithm which is based on probability and statistics gets more and more attention due to its high speed and simplicity. However, its biggest drawback is the poor performance for videos with moving shadows. Because the process of the original ViBe algorithm is carried out in the gray space, it is against the extraction of background as existing shadow removal algorithms typically operate with a colored background image. In this paper, a background extraction and shadow removal algorithm based on clustering for ViBe is proposed. Firstly, the color background is extracted by clustering the values of each pixel in R, G, B components space. Then the background information is used to remove shadow hidden in the foreground. For indoor and outdoor videos with moving cast shadows, ROC (Receiver Operating Characteristic) curve is used to validate the proposed approach. Experimental results show that a good performance has been gained in shadow removal.

**2.1.2 Adaptive Background Subtraction in Images**

Published in: 2018 International Conference on Advances in Communication and Computing Technology (ICACCT).

Authors: Neha S. Sakpal, Manoj Sabnis

Selective background subtraction is the major problem associated with background subtraction technique. For foreground detection, background modeling is used in many different applications to subtract the background and detect foreground object in the image. There are many challenges in elaborating a good background subtraction algorithm and researcher have been appropriated to developing the new innovation and enhancement techniques to overcome all the limitations. In this paper, we present selective background subtraction technique to subtract unnecessary background from the foreground and background scene.

**2.1.3 Go-selfies: A Fast Selfies Background Removal Method Using ResU-Net Deep Learning**

Published in: 2020 28th European Signal Processing Conference (EUSIPCO).

Authors: Yunan Wu

The selfies play an important role in recording meaningful moment in human's daily life. In most cases, before sharing photos, people often synthesis attractive images on some phone applications, such as Photoshop. While these kinds of software have reached good performance nowadays, they are too complex for simple life usage. In this work, we proposed an automatic segmentation model unique to segment human selfies photos. We first constructed a large photo segmentation database and built 8 different models based on resolution, image size and whether or not to use transfer learning and picked the best one among them. We then applied cyclical learning rate method and pre-trained encoder network to fine tune our models. Finally, our best model tested on Google images demonstrated satisfying promising results on both accuracy scores and losses, which will be the precondition in real-time segmentation. We named this lovely web product as "Go Selfies".

**2.1.4 Fully convolutional networks for semantic segmentation**

Published in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Authors: Jonathan Long, Evan Shelhamer, Trevor Darrell

Convolutional networks are powerful visual models that yield hierarchies of features. We show that convolutional networks by themselves, trained end-to-end, pixels-to-pixels, exceed the state-of-the-art in semantic segmentation. Our key insight is to build “fully convolutional” networks that take input of arbitrary size and produce correspondingly-sized output with efficient inference and learning. We define and detail the space of fully convolutional networks, explain their application to spatially dense prediction tasks, and draw connections to prior models. We adapt contemporary classification networks (AlexNet , the VGG net , and GoogLeNet ) into fully convolutional networks and transfer their learned representations by fine-tuning to the segmentation task. We then define a skip architecture that combines semantic information from a deep, coarse layer with appearance information from a shallow, fine layer to produce accurate and detailed segmentations. Our fully convolutional network achieves state-of-the-art segmentation of PASCAL VOC (20% relative improvement to 62.2% mean IU on 2012), NYUDv2, and SIFT Flow, while inference takes less than one fifth of a second for a typical image.

**2.1.5 Different Approaches for Semantic Segmentation**

Published in: 2020 5th International Conference on Communication and Electronics Systems (ICCES).

Authors: Greeshma P.G.

Semantic image segmentation is an emerging task in the field of automation. Its application varies from autonomous driving to medical diagnosis. Semantic segmentation of an image means to label each pixel in that image to a particular class. As an example consider an outdoor street image where there are different objects like car, road, sky, trees, pedestrians etc. After applying semantic segmentation each pixel in the image belonging to the car will have the label car and road will have label road and so on. A recent trend in performing semantic segmentation is by using Convolutional Neural Networks, (CNN), which acted as a catalyst for segmentation. In this paper, a detailed discussion of various approaches for segmentation using CNN has been presented. Also, various datasets and their format and evaluations metrics are discussed. All the approaches discussed are diverse and has its pros and cons. Finally, an application-specific semantic segmentation method using a genetic CNN algorithm for classification task has been proposed. The proposed method has shown improvement in the M iou score when tested on the CamVid dataset and on a dataset created by combining two small object classification datasets, MNIST and CIFAR10.

**2.1.6 Real-time background subtraction for video surveillance: From research to reality**

Published in: 2010 6th International Colloquium on Signal Processing & its Applications.

Authors: M. Hedayati, Wan Mimi Diyana Wan Zaki , Aini Hussain.

This paper reviews and evaluates performance of few common background subtraction algorithms which are median-based, Gaussian-based and Kernel density-based approaches. These algorithms are tested using four sets of image sequences contributed by Wallflower datasets. They are the image sequences of different challenging environments that may reflect the real scenario in video surveillances. The performances of these approaches are evaluated in terms of processing speed, memory usage as well as object segmentation accuracy. The results demonstrate that Gaussian-based approach is the best approach for real-time applications, compromising between accuracy and computational time. Besides, this paper may provide a better understanding of algorithm behaviors implemented in different situation for real-time video surveillance applications.

**2.2 Patents**

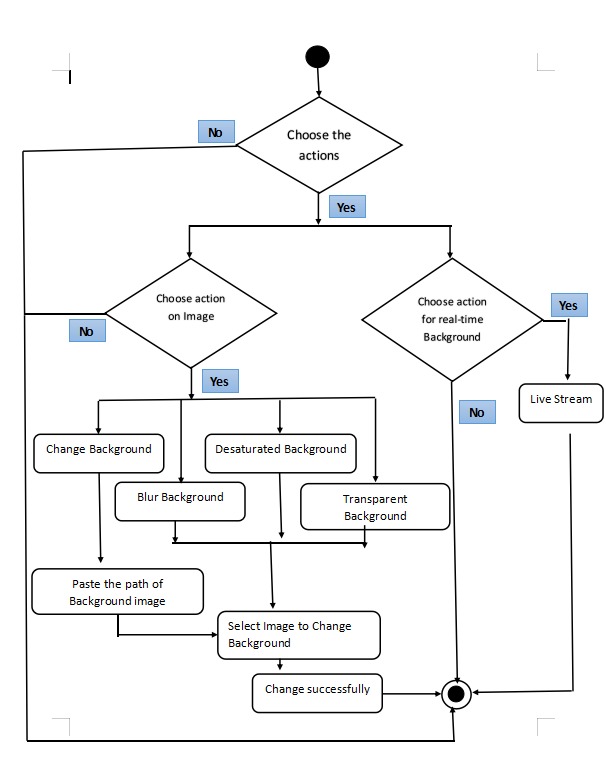
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| PatentApplication No. | | Title of Patent | | Existing Solutions(Abstract of Patent) | |
| US7720283B2 | | Background removal in a live video | | Exemplary systems and methods segment a foreground from a background image in a video sequence. In one implementation, a system refines a segmentation boundary between the foreground and the background image by attenuating background contrast while preserving contrast of the segmentation boundary itself, providing an accurate background cut of live video in real time. A substitute background may then be merged with the segmented foreground within the live video. The system can apply an adaptive background color mixture model to improve segmentation of foreground from background under various background changes, such as camera movement, illumination change, and movement of small objects in the background. | |
| US20100158379A1 | | Image background removal | | Request is received to enter a background removal mode. A marquee is displayed for specifying a portion of an image to be provided to a foreground extraction engine (“FEA”).  The portion of the image defined by the marquee is provided to the FEA. The FEA identifies the foreground and returns it to the application program. The image is displayed with the foreground distinguished from the background. If the marquee is re-sized or re-positioned, the portion of the image defined by the marquee is provided to the FEA. A line segment may be drawn over the background to define a portion of the image for inclusion in the foreground. A line segment may also be drawn over the foreground to define a portion of the image for inclusion in the background. Upon exit of the background removal mode, only the foreground of the image is displayed. | |

**2.2.1 Table of Patents**

**Chapter 3: Work Done**

**Chapter 3: Work Done**

**3.1 Architecture and Design:**



**Figure 3.1:** Activity Diagram of background removal process.

**3.2 TOOLS USED:**

**3.2.1 PYCHARM / VSCODE:**

PyCharm is an [integrated development environment](https://en.wikipedia.org/wiki/Integrated_development_environment" \o "Integrated development environment) (IDE) used in [computer programming](https://en.wikipedia.org/wiki/Computer_programming" \o "Computer programming), specifically for the [Python](https://en.wikipedia.org/wiki/Python_(programming_language)" \o "Python (programming language)) language. It provides code analysis, a graphical debugger, an integrated unit tester, integration with version control systems (VCSes), and supports web development with Django as well as data science with Anaconda.

Visual Studio Code is a source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git.

**3.2.2 MEDIAPIPE:**

MediaPipe is a framework for building multimodal (eg. video, audio, any time series data), cross platform (i.e Android, iOS, web, edge devices) applied ML pipelines. With MediaPipe, a perception pipeline can be built as a graph of modular components, including, for instance, inference models (e.g., TensorFlow, TFLite) and media processing functions.

**3.2.3 CVZONE**

Cvzone is a computer vision package that makes us easy to run like face detection, hand tracking, pose estimation, etc., and AI functions. It contains segmentors which can be used to perform actions on images using various segmentation techniques.

**3.2.4 SELFIE SEGMENTATION**

Selfie Segmentation API allows developers to easily separate the background from users within a scene and focus on what matters. Adding cool effects to selfies or inserting your users into interesting background environments has never been easier.

**3.2.5 FLASK**

Flask is a micro web framework written in Python. It is classified as a micro-framework because it does not require particular tools or libraries. Flask supports extensions that can add application features as if they were implemented in Flask itself.

**3.2.6 MOBILENETV3**

MobileNetV3 is the third version of the architecture powering the image analysis capabilities of many popular mobile applications. The architecture has also been incorporated in popular frameworks such as TensorFlow Lite.MobilenetV3 is needed to apply continuous segmentation of frames in real-time.

**3.3 Processes Followed**

**3.3.1 Project Initiation**

Defining vision of project which is to replace background from real-time camera feed and input images.

**3.3.2 Project planning**

Defining the basic architecture of the project, planning how to proceed with the possible challenges and deciding which modules, tools, libraries, hardware and software’s to use.

The process in which the project starts to come live writing scripts and executing as planned by overcoming the problems arising. Maintaining image dimensions for the used models to work properly, applying proper segmentations on images and displaying those further extending it to options of saving and resizing images.

Adjusting the model threshold for perfect segmentation and background replacement with a frame rate of more than 30 FPS.

**3.3.4 Project Monitoring and controlling**

Testing the working project on various scenarios with multiple objects in focus to check for proper background replacement in real-time. Trying out various complex backgrounds to replace Overcoming bugs and problems.

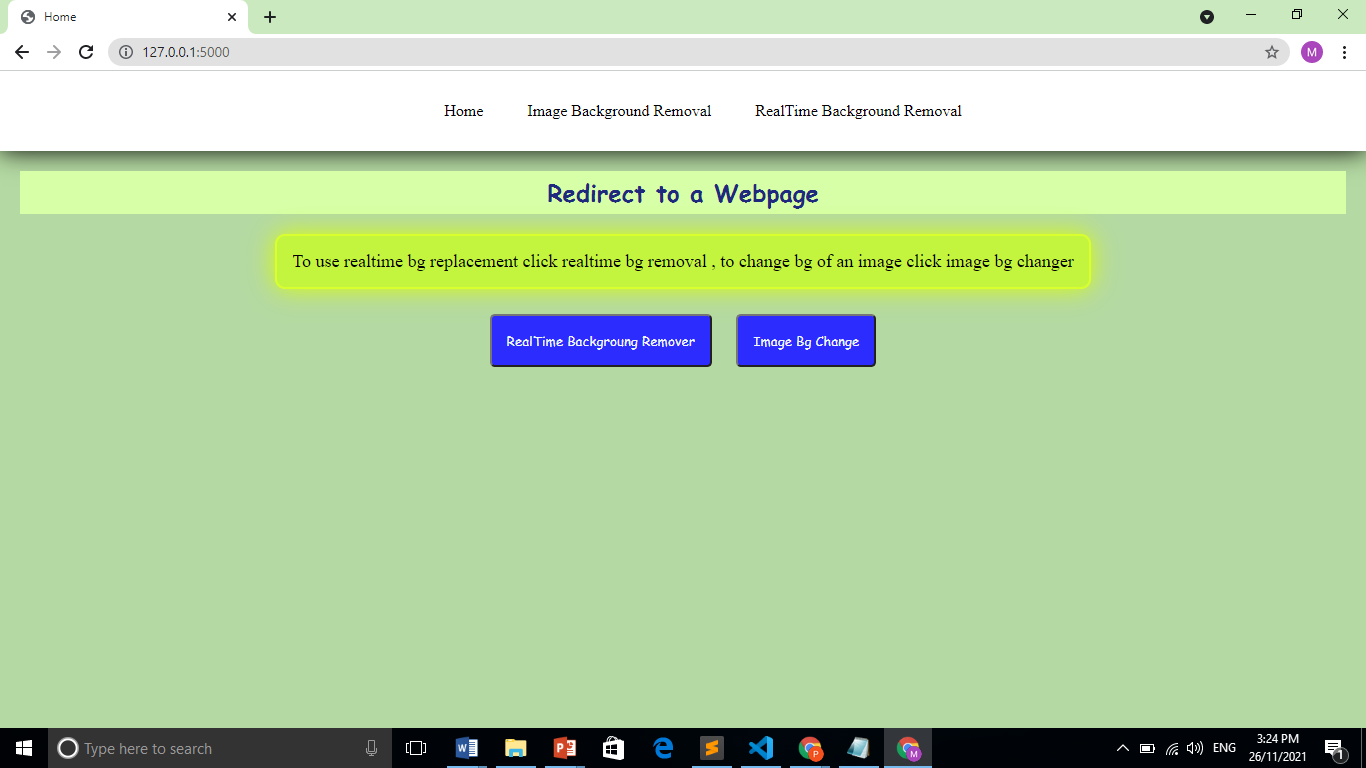
**3.3.5 Project Closing**

Reviewing if the project goal is completely achieved and successfully working without problems.

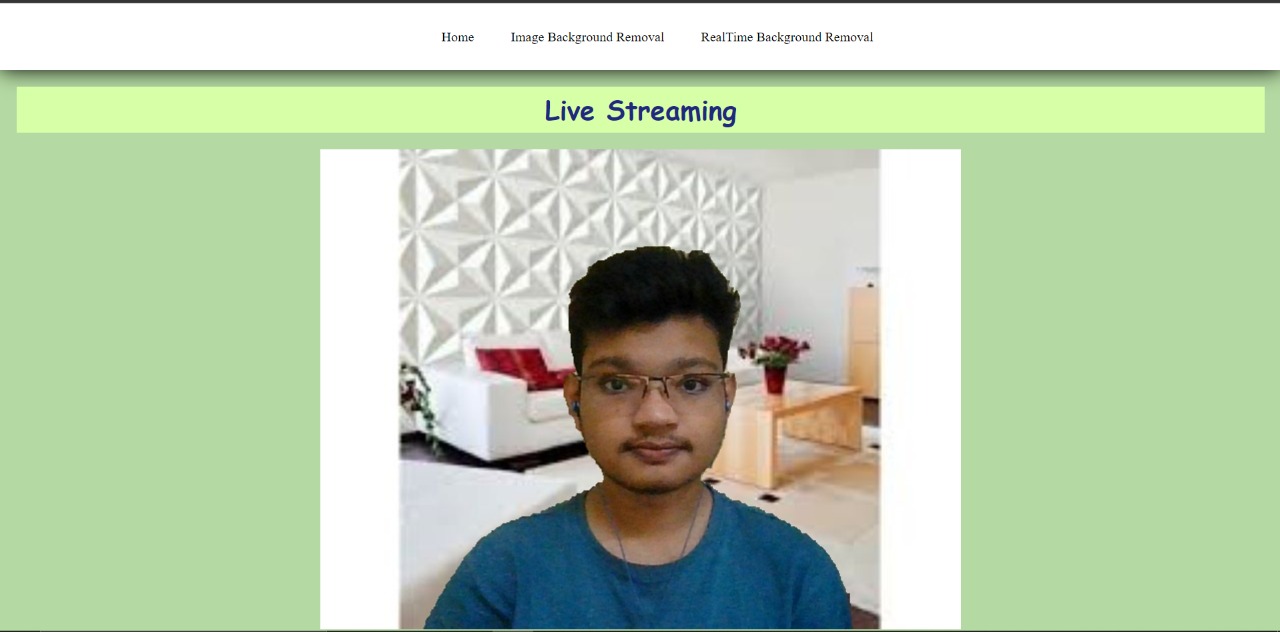
**Chapter 4: Results and Discussions**

**Chapter 4: Results and Discussions**

When the user visits the website, the given interface appears

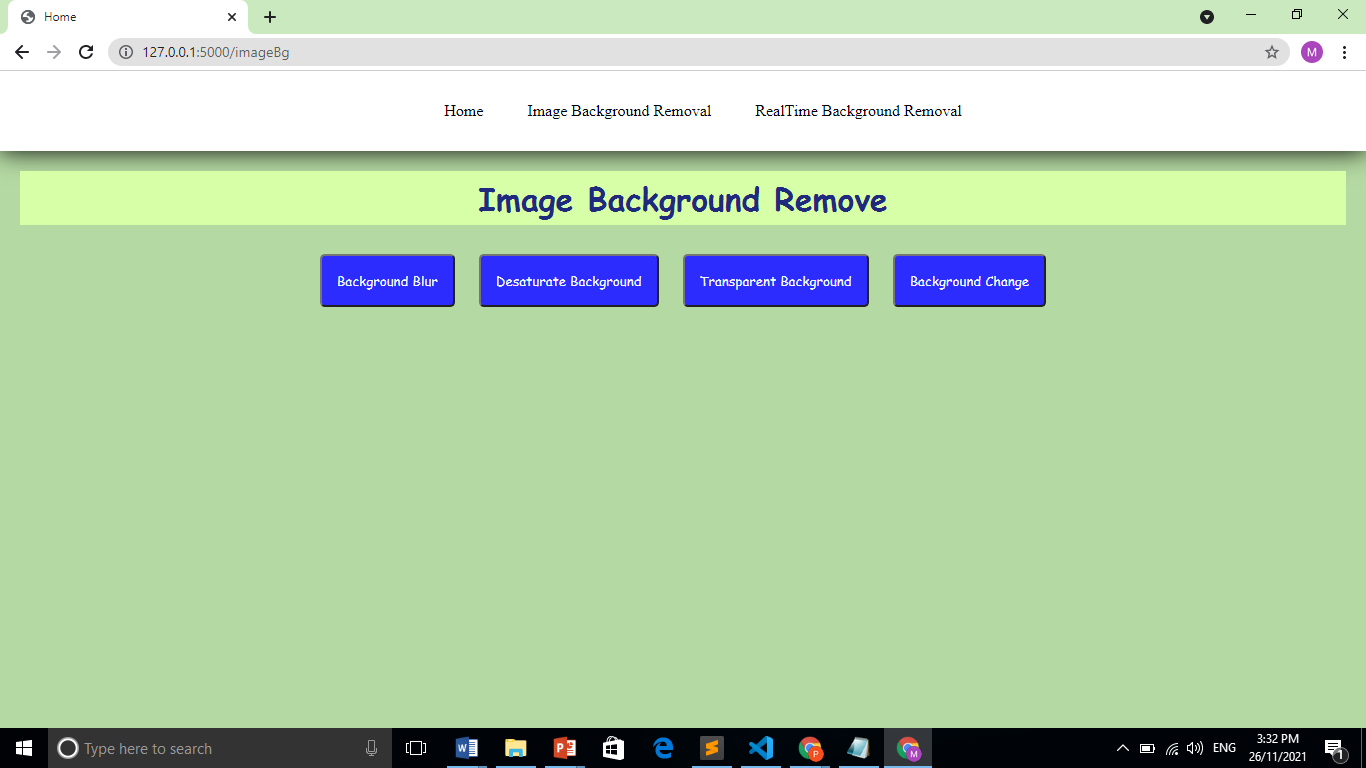


**Fig 4.1 User Interface (Select Real-Time background Remover option)**



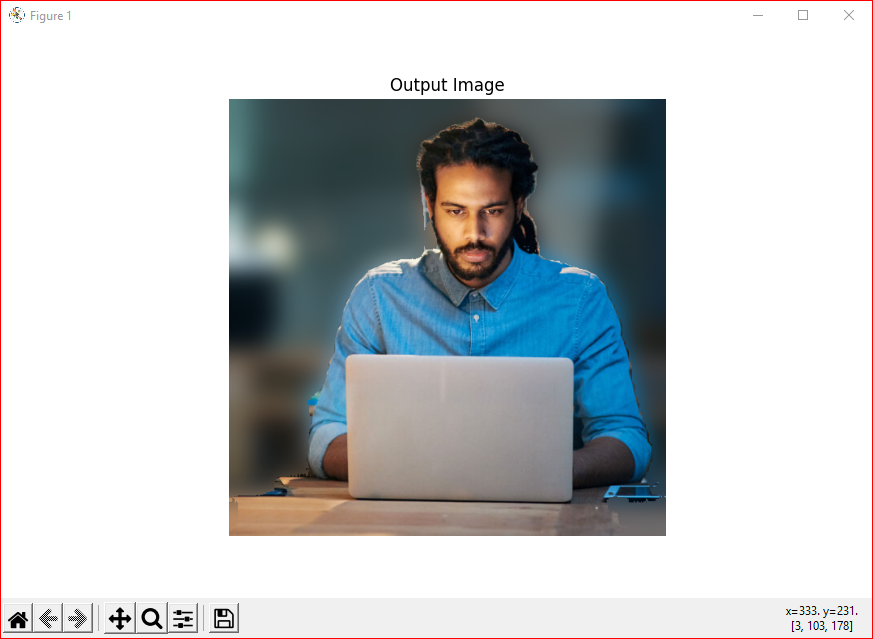
**Fig 4.2: Real-time background removal.**

Selecting image background change option this interface will appear it as four option as shown below,



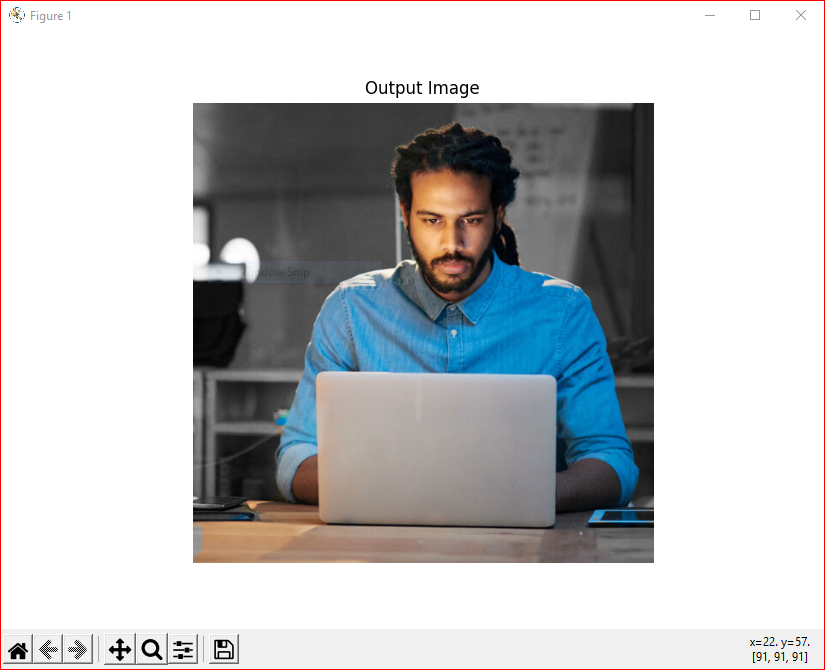
**Fig 4.3: Select Image background change option**

Selecting background blur image background will be blurred.



**Fig 4.4:** **Background blur image**

Selecting background desaturate image background will be desaturate.



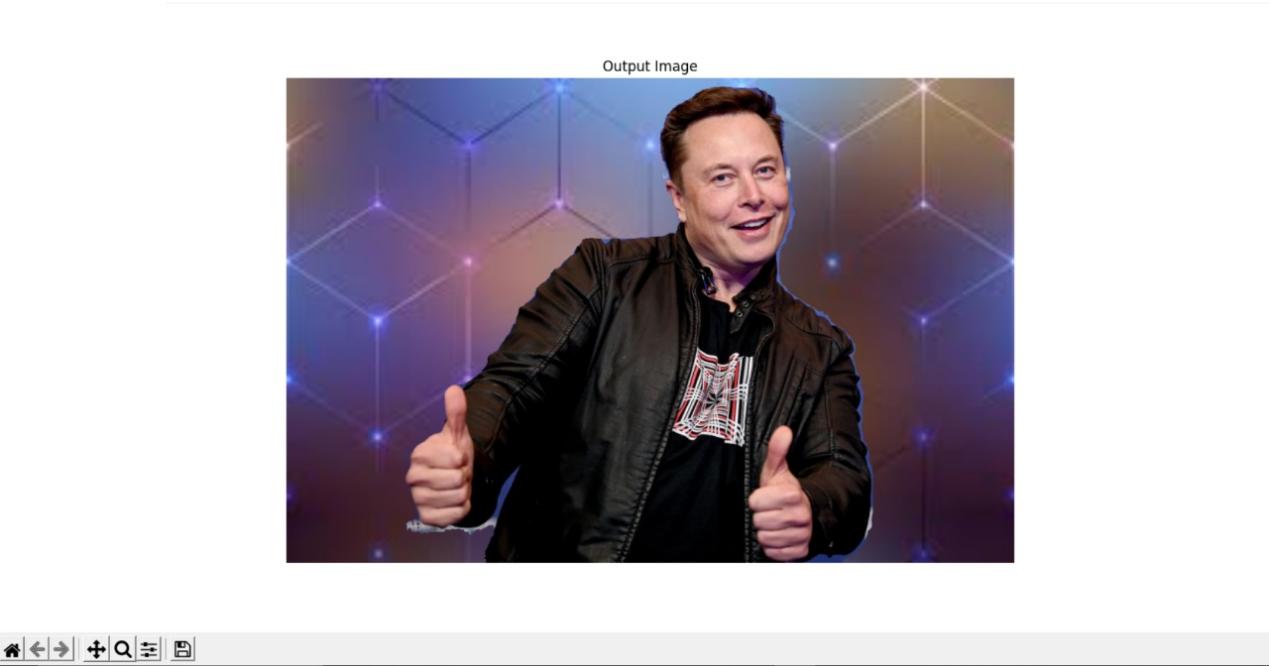
**Fig 4.5:** **Background desaturated image**

Selecting background transparent image background will be transparent.

****

**Fig 4.6:** **Background transparent image**

Selecting background Change image background will be change

****

**Fig 4.7:** **Background Change image**

**Chapter 5: Future Scope and Conclusions**

**Chapter 5: Future Scope and Conclusions**

**5.1 Future Scope**

* The largest marketplaces in the eCommerce world require neutral or white backgrounds behind the product as the high-quality images on the neutral backgrounds pop up the features of the products and thus making it convenient for the buyer to assess the color and the elements.
* The image background removal can help in creating consistent imagery that you can later use for your website or catalog designing.

**5.2 Conclusion**

In the current technically advanced world, virtual meets are becoming fairly important and faster way of communication. The background of the user might be distracting and sometimes improper for an urgent virtual meet. We have implemented a program which overcomes this problem and is efficient enough. The background removal in real time would help one to focus without worrying about the background. The higher fps provided makes it a better experience for the user.

**Chapter 6: Literature cited/ References**

**Chapter 6: Literature cited/ References**

[1] “ MediaPipe Selfie Segmentation”, https://google.github.io/mediapipe/solutions/selfie\_segmentation.html

[2] “Background removal with deep learning”,https://towardsdatascience.com/background-removal-with-deep-learning-c4f2104b3157

[3] “Background Removal with Python”, https://towardsdatascience.com/background-removal-with-python-b61671d1508a

[4] Yong Xu, Jixiang Dong, Bob Zhang, Daoyun Xu, “Background modeling methods in video analysis: A review and comparative evaluation”, https://doi.org/10.1016/j.trit.2016.03.005, Volume 1, Issue 1, January 2016.

[5] Siying Qian; Chenran Ning; Yuepeng Hu, “MobileNetV3 for Image Classification”, https://ieeexplore.ieee.org/document/9389905, 26-28 March 2021

[6] Liubov Zatolokina, “Deep Learning-based Background Removal And Blur In A Real-Time Video” https://mobidev.biz/blog/background-removal-and-blur-in-a-real-time-video, Feb 22, 2021

[7] Lorena de los Angeles Guachi Guachi,"Background Subtraction for moving object detection", http://dspace.unical.it:8080/jspui/bitstream/10955/1308/1/Thesis-Background%20Subtraction%20for%20moving%20object%20detection.pdf

[8] Ahmed Mohammed Ahmed Madani, “Moving objects detection using background subtraction technique", http://api.uofk.edu:8080/api/core/bitstreams/c9e0b3de-c36a-41a2-b2e6-2acccffcb9b6/content,July 2011