Augmented cut and paste

Jay Arun Bhoir
jaybhoir172@nhitm.ac.in
New Horizon Institute of Technology and Management,
Thane, Maharashtra

Chaitanya Naveen Nawathe
chaitanyanawathe172@nhitm.ac.in
New Horizon Institute of Technology and Management,
Thane, Maharashtra

Shubham Ravi Jambekar
shubhamjambekar172@nhitm.ac.in
New Horizon Institute of Technology and Management,
Thane, Maharashtra

Yogita Chavan
yogitachavan@nhitm.ac.in
New Horizon Institute of Technology and Management,
Thane, Maharashtra

ABSTRACT

Augmented reality (AR) is an intelligently involvement of a real-world environment where the objects that dwell inside the world are upgraded by computer-generated perceptual data, now and then over numerous sensory modalities, including visual, auditory, haptic, somatosensory, and olfactory. AR is regularly characterized as a framework that fulfills three essential highlights: a blend of genuine and virtual universes, real-time interaction, and exact 3D enlistment of virtual and genuine objects. The overlaid sensory information can be constructive (i.e. additive to the natural environment), or destructive (i.e. masking of the natural environment).This involvement is consistently interlaced with the physical world such its seen as an immersive angle of the vital environment.In human-computer interaction and interface plan, cut, duplicate and glue are related commands that give an inter-process communication method for exchanging information through a computer;s interface. The cut command evacuates the chosen information from its unique position, whereas the duplicate command makes a copy; in both cases, the chosen information is kept in transitory capacity (the clipboard). The information from the clipboard is afterward embedded wherever a glue command is issued.The information remains accessible to an application supporting the highlight, in this way permitting simple information exchange between applications.With modern developments in machine learning, it's possible to accurately detect people and objects around us, remove the background automatically and transfer the result to a computer. To date, AR had been utilized to venture advanced pictures into the genuine world.However, the AR Cut & Copy app reverses the process and brings physical things into the digital world.

Keywords— *Augmented Reality, BASNet, Deeplab, SIFT, Screenpoint*

1. INTRODUCTION

As the technology is growing day by day, daily work in people's life is getting digitalized and it's been helping by reducing the load in every some kind of way. Even if there are lots of shortcuts introduced to people, still there are many things that are quite easy but the process can be quite irritating. The editing of documents in different ways by the photoshop.The drag and drop include for including records to an online site or joining them to our emails fair by dragging those from their area to the net location application has become habit to us. It is a feature that saves minutes of tedious browsing through the pop-up window, which doesn't allow us to switch to a special screen until the file is chosen , or the user decides to quit. In all this Augmented Reality can be very interesting to help with a new perspective. An Augmented Reality is a real-time view of a physical environment enhanced by adding computer-generated digital information thereto environment. The purpose of Augmented Reality is to redefine or augment how a private would interact with and interpret the important world by introducing virtual information to his or her direct surroundings and approach to the real-world environment as an entire. The drag and drop feature for including records to an online site or joining them to our emails fair by dragging those from their area to the net location application has become habit to us. But if we were able to copy-paste real-world objects/items into digital screens it will save a lot of time and ease our work.Expanded reality empowers us to copy-paste real-world things into our advanced screens with fair by few clicks.Increased Reality has upgraded our real-world encounter with different highlights inside our screens inside the computerized world.But this concept of bringing real-world objects into the digital world brings a few complete shifts from our perspective and our interaction with Augmented Reality Apps. The process of pasting real-world objects by clicking the snap of an object, getting obviate its background, and sharing it together with your

desktop may be a tedious one. With this Augmented Reality App, it replaces this whole process by pointing your camera to select the desired object, making it efficient by bypassing snapping, masking, saving, and switching windows for the user making it a really efficient AR copy paste photoshop tweak.

2. LITERATURE SURVEY

R. Silva, J. C. Oliveira, G. A. Giraldiin in the paper titled Introduction to Augmented Reality, Augmented Reality (AR) is a new technology that involves the overlay of computer graphics on the real world. One of the best overviews of the technology is, that defined the field, described many problems, and summarized the developments up to that point. AR is inside a more common setting named Blended Reality (MR), which alludes to a multi-axis range of ranges that cover computer amusement (VR), AR, telepresence, and other related technologies.Virtual Reality may be a term utilized for computer-generated 3D situations that permit the client to enter and associated with manufactured situations.In VR the environment is totally synthetic and in telepresence, it's completely real, in AR the user sees the important world augmented with virtual objects. The problem of removing real objects is quite simply extracting depth information from a scene. The system must even be ready to segment individual objects therein environment. A semi-automatic strategy for recognizing objects and their areas within the scene This empowers the inclusion of virtual objects and cancellation of genuine objects without an express 3D remaking of the environment.[1]

Jonathan J. Hull, Berna Erol, Jamey Graham, Qifa Ke, Hidenobu Kishi, Jorge Moraleda, Daniel G. Van Olst within the paper titled Paper-Based Augmented Reality, proposed a system during which a replacement method for augmenting paper documents with electronic information is described that doesn't modify the format of the document in any way.Appropriate to both commercially printed records additionally as archives that are yield from PC, the method we call Paper-Based Expanded Reality significantly progresses the utility of the paper.We describe the popularity of technology that creates this possibility, also as several applications.Execution on a camera phone is talked about that lets clients recover information and get to joins from paper reports to electronic data.Recognition is performed at a speed of 4 frames per second on a Treo 700w and support is provided for several user applications, including clickable paper printed sites whose appearance is unaltered but which can be explored with a camera phone.A replacement paradigm for augmented reality was described during which electronic data is added to paper documents without changing the looks of the paper document in any way. This approach leverages the essential discovery that singular signatures are often derived from an image of a little patch of text which signatures are often linked to electronic data. [2]

Xuebin Qin, Zichen Zhang, Chenyang Huang, Chao Gao, Masood Dehghan, and Martin Jagersand in the paper titled BASNet: Boundary-Aware Salient Object Detection, proposed a system where Deep Convolutional Neural Networks were received for striking protest discovery and accomplished the state-of-the-art performance. The human vision system, for choosing the most important information from visual scenes, has an effective attention mechanism.Computer vision points at modeling this component in two inquire about branches: eye-fixation location and striking question detection.Their work centers on the moment department and points at precisely portioning the pixels of notable objects in an input image.The

results have immediate applications in e.g. picture segmentation/editing and control, visual following, and client interface optimization. The proposed BASNet consists of two modules. The prediction module may be a U-Net-like densely supervised Encoder-Decoder network, which learns to predict saliency maps from input images. The multi-scale Residual Refinement Module (RRM) refines the coming about saliency outline of the expectation module by learning the residuals between the saliency outline and the ground truth.[3]

Yann LeCun, Yoshua Bengio & Geoffrey Hintonin the paper titled Deep learning proposed an idea where Deep learning allows computational models that are composed of multiple processing layers to find out representations of knowledge with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual perception, object detection, and lots of other domains like drug discovery and genomics. Deep learning discovers detailed structures in large data sets by using the backpropagation algorithm with the aim of indicating how a machine should change its internal parameters that are wont to compute the representation in each layer from the previous layer's representation. Deep convolutional nets have caused breakthroughs in processing images, video, speech, and audio, whereas recurrent nets have shone a light-weight on sequential data like text and speech. [4]

3. PROPOSED SYSTEM AND METHODOLOGY

The user needs to open the application on the mobile phone and then capture an image with the camera on the phone and place it in front of the Computer monitor where his photoshop is Ready and then drop the image on it.

System Requirements:
As we are going to make an application this project is only software-based there will be no hardware components required.

Software Specification:
The Software specification required for the proposed system are as follows:
- Mobile phone with Android version 7+
- Desktop/laptop Operating System – Windows 10
- Programming Language - Python 3.6, ReactNative
- Text editor- Atom
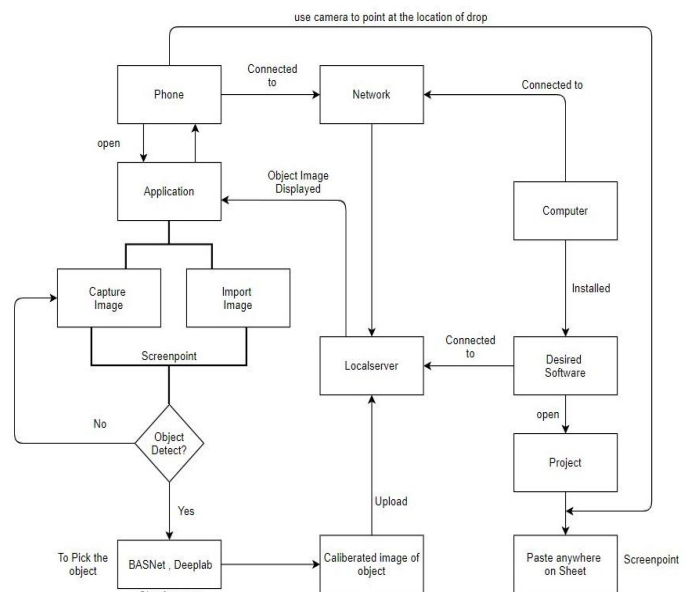- Editing software -Microsoft Word, Powerpoint

Fig. 1: Proposed System Architecture

As it's going to be an application installed on a phone, we open the application first. It will ask permission to use the phone's camera and storage access which you have to allow. It will give you two options to select from, either capture image or import image u want to use. If you choose import image it will direct you to your phone's gallery and you have to choose one image which you need, or if you choose to capture an image it will open your phone's camera and you have to capture an image. After capturing/importing the image the tool uses Scale-invariant feature transform (SIFT).

Scale-invariant feature transform (SIFT) - is a feature detection algorithm in computers to detect and describe local features in images. For any object in an image, interesting points on the object can be extracted to provide a "feature description" of the object. This portrayal, extricated from the picture, can at that point be utilized to recognize the protest when endeavoring to find the question within the picture containing numerous other objects.To perform dependable acknowledgment, it is imperative that the highlights extricated from the picture be distinguishable indeed beneath changes in picture scale, clamor, and illumination.Such focuses ordinarily lie on high-contrast districts of the picture, such as question edges. SIFT can robustly identify objects even among clutter and under partial occlusion because the SIFT feature descriptor is invariant to uniform scaling, orientation, and illumination changes. In detects object in the image. If the captured/imported image is not clear or has low resolution, it may fail to detect the object. If it fails the app will notify you something like *Failed to detect please try again* and redirect you to camera or gallery depends on the option you choose first. Now as the object is detected in the image successfully two algorithms BASNet and Deeplab are used to pick an object out of the image.

Fig. 2: SIFT

Boundary-Aware Salient Object Detection (BASNet) is basically a model used for salient object detection and background removal. BASNet sends back the B&W salience mask which helps in boundary detection and helps in identifying objects perfectly. The accuracy and range of this model are very good.

Deeplab may be a state-of-art profound learning demonstrate for semantic picture division, where the objective is to allot semantic names (e.g., individual, puppy, cat, at that point on) to each pixel.In this case, it will identify and assign labels to each and every pixel of the object to define the object clearly and make sure it outputs perfectly with every detail it had upon capturing.
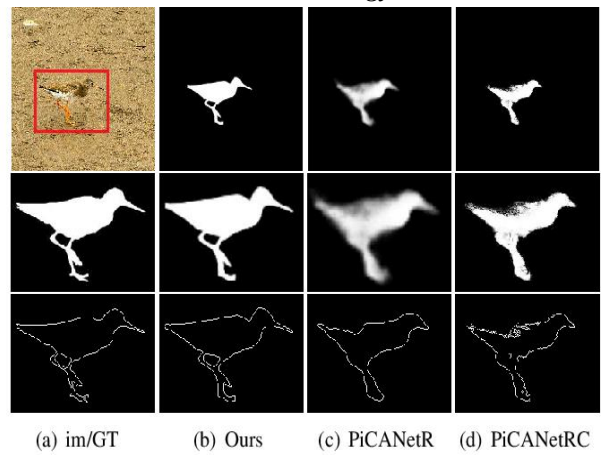

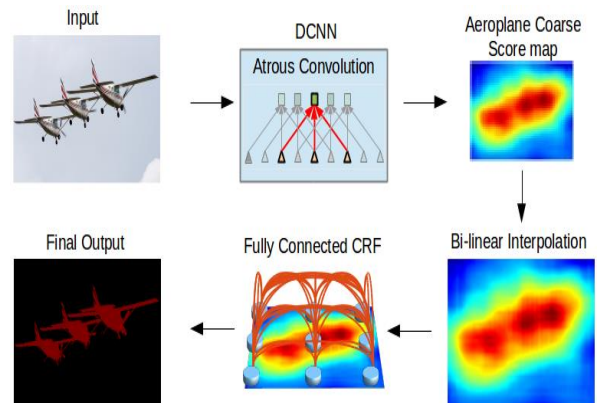(a) im/GT   (b) Ours   (c) PiCANetR   (d) PiCANetRC
Fig. 3: BASNet

Fig. 4: DeepLab

As soon as we get the image of the object which will be saved in the app itself so it can upload it to the local server. It is the interface between the mobile app and Photoshop. The phone and the system on which photoshop is installed should be connected to the same network for the local server to work. The local server will be already created by the mobile application. We only have to open the photoshop software and open a new project and then go to preferences and open the remote host connection option. In there we have to enter the IP of the local server that the application has created to make sure it is connected to the server and will be ready to accept the drop i.e., the image of the object.

After you set up your software which is photoshop, in this case, the image of the object should be on your phone screen, and on the background camera should be opened. You now have to put the phone in front of the screen and point the camera anywhere on the screen. Now it uses a module called Screenpoint.

Screenpoint finds the (x,y) facilitates of the centroid of a picture (e.g.: a portable camera picture) indicating at another picture (e.g.: a computer screen) utilizing OpenCV SIFT.Then click on the image within a couple of seconds the object will appear on your computer screen.
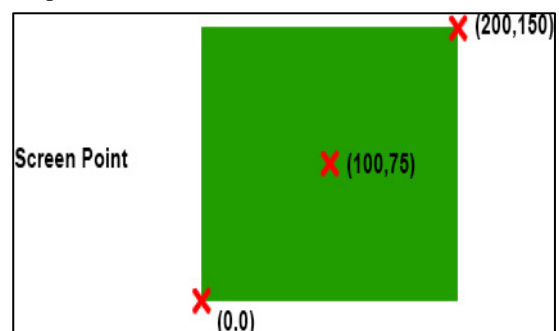

Fig. 5: Screenpoint

4. RESULT AND ANALYSIS

The android app is made using reactnative language. The camera will open by default at the start. We have to place the object inside the four edges and capture the image.
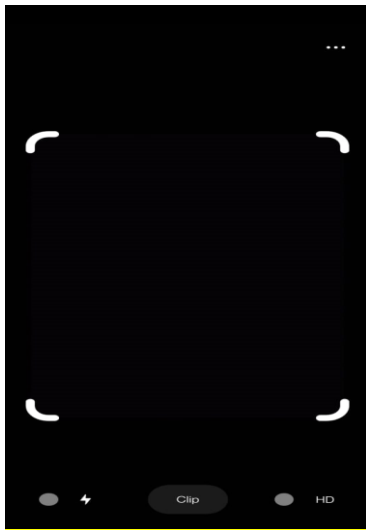

Fig. 6: Android app GUI

Once the image is captured it will show the object and its ready to export or share to your system.
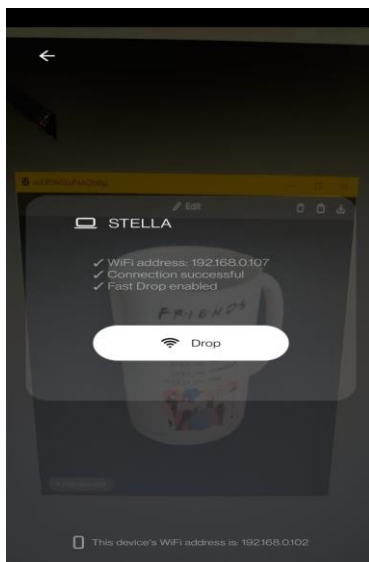

Fig. 7: Object captured
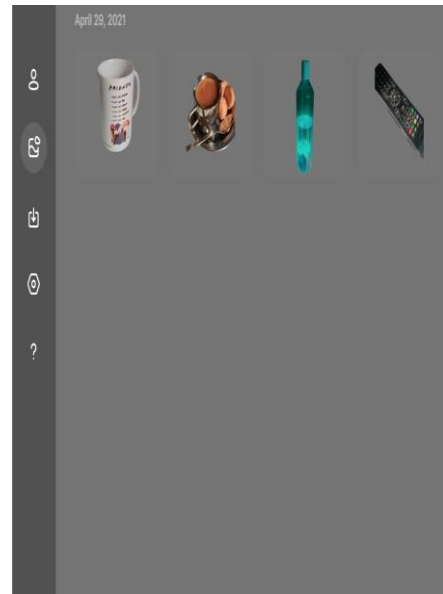

Fig. 8: Connection successful

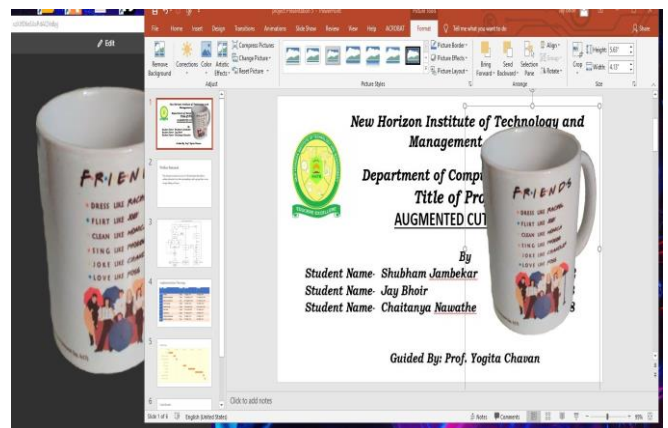**Fig. 9: Application console on the system**

Fig. 10: Drag and Drop

By using the pointer drag and drop the object image in powerpoint presentations or Microsoft word documents.

5. CONCLUSION AND FUTURE SCOPE

This project will be very useful in terms of making presentations, editing, and documents that require images. Most of the time while making a presentation we go to google for images and spend a lot of time finding the image, sometimes we found it but with some problems with pixels or quality so we then edit those images. This application will save a lot of time as u directly pick the image from your surroundings and pass it on to the computer.

Our project is only a prototype, i.e., this project is made for only the cut and paste of images. We can further integrate scanning the movement of a specific object and then adding to the desired content. Availability of the tool can be increased to more software and system.

6. REFERENCES

[1] "IEEE Standard for Augmented Reality Learning Experience Model," in IEEE Std 1589-2020, vol., no., pp.1-48, 17 April 2020.

[2] G. Lindstrom, "Programming with Python," in IT Professional, vol. 7, no. 5, pp. 10-16, Sept.-Oct. 2005.

[3] X. Qin, Z. Zhang, C. Huang, C. Gao, M. Dehghan and M. Jagersand, "BASNet: Boundary-Aware Salient Object Detection," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 7471-7481.

[4] Z. Xu, J. Wang. and L. Wang, "Infrared Image Semantic Segmentation Based on Improved DeepLab and Residual Network," 2018 10th International Conference on Modelling, Identification and Control (ICMIC), Guiyang, 2018, pp. 1-9.

[5] J. J. Hull et al., "Paper-Based Augmented Reality," 17th International Conference on Artificial Reality and Telexistence (ICAT 2007), Esbjerg, Jylland, 2007, pp. 205-209.