International Conference on Recent Trends in Data Science and its Applications
DOI: rp-9788770040723.049

A Novel Approach for Live Image Capture and Estimation of Coordinates for Augmented Projection Using Machine Learning

Rahul Chiranjeevi V
Assistant Professor,
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
rahulchiranjeevi777@gmail.com

M V Paavai
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
paavai2203@gmail.com

Aakash PS
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
aakash2492000@gmail.com

M Parthiban
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
parthibanmurali007@gmail.com

Abstract*—In recent times, augmented reality (AR) has emerged as one of the most promising technologies that can provide a high degree of adaptation and independence to facilitate the transfer of knowledge in the majority of different sectors. AR is being used in the process of projecting objects into an augmented environment to demonstrate to customers how those objects will seem in the setting of their choice. Even though there have been and will continue to be significant advances made, the visuals are still being projected in a three-dimensional environment. In order to circumvent this process and project a live image that has been caught in reality, an innovative strategy that makes use of modified machine learning algorithms is currently under consideration. Additionally, a brand-new method known as the "universal clipboard" is suggested, which makes it much simpler to copy and paste photographs between different devices.

Keywords*—Augmented reality, Machine learning, Unsupervised Learning, Deep Learning.

I. INTRODUCTION

As technology keeps getting better, more and more people's daily tasks are being done digitally, which makes life easier in every way. Even if a lot of shortcuts are shown to people, there are still a lot of things that are easy but annoying to do. Different ways that Photoshop can be used to change documents. This got us used to the drag-and-drop feature, which lets us add records to an online site or link them to our emails by dragging them from where they are to where they should be on the web. It is a feature that saves minutes of time that would have been spent browsing through the pop-up window, which doesn't let us switch to a different screen until the file is chosen or the user quits. All in all, Augmented Reality can be very interesting and help you see things from a different point of view. The goal of the project is to change how people interact with and understand the real world by adding virtual information to their immediate surroundings and how they interact with the real world as a whole. Users use the drag-and-drop feature all the time to add records to an online site or link them to our emails. Users do this by dragging the records from where they are to where the web application is. But if users could copy and paste things from the real world onto digital screens, it would save us a lot of time and make our work easier. Expanded reality lets us copy and paste things from the real world into our high-tech screens with just a few clicks.

Through different features on our screens in the digital world, augmented reality has made our experiences better. But the idea of bringing real-world objects into the digital world changes the way users think about and use Augmented Reality Apps in a few big ways. Real-world objects can be pasted by clicking on the object's snap, getting its background, and then sharing it with your desktop. This can be a time-consuming process. With this Augmented Reality app, you just point your camera at the object you want to select. This saves time because the user doesn't have to snap, mask, save, or switch windows. This is a really efficient AR copy-paste Photoshop tweak. AR is the intelligent use of a real-world environment in which items are. AR is often described as a system with three key features: a mix of real and virtual worlds, interaction in real-time, and accurate 3D modeling of both real and virtual objects. The extra sensory information can be helpful, or it can be dangerous. This kind of involvement is always tied to the real world in a way that makes it seem like an immersive part of the vital environment.

Our application has 2 features: Using Augmented reality and machine learning - their project and process images in our application. For processing, Human-computer interaction (HCI) and interface design use commands like "plan," "cut," "copy," and "glue" to transmit information between processes. The cut command takes the information you've chosen out of its original position, while the duplicate command makes a copy of it. Both situations store selected data temporarily (the clipboard). The information from the clipboard is subsequently incorporated whenever a glue command is issued. Machine learning can now reliably locate people and objects, automatically eliminate the backdrop (background removal), and communicate the output to a computer. Considering image projection, AR has so far been utilized to bring enhanced visuals into the actual world. AR Cut & Copy works in reverse, digitally translating tangible objects. This study presents a novel approach to using augmented reality to make life easier by making things happen at the fingertips of a human being. It helps in visualizing objects without the tedious efforts of moving as they use coordinates and SIFT.

II. RELATED WORK

Gang Wang[1] has proposed a system Prior to performing saliency detection, a background estimate is

DOI: rp-9788770040723.049
applied. To gather background information, they use the bounding boxes approach. The model the saliency, there are 3 background works that are primarily addressed, including background connectivity prior, background contrast prior, and spatial distribution prior, enabling the proposed technique to emphasize the salient item as a whole and reduce background clutters.Experiments on two benchmark datasets show that our solution beats 11 leading methods while being more efficient than the majority of them.

They increase convolutional neural network pooling in order to identify prominent items. On the bottom-up route, they develop a global guidance module (GGM) using a U-shaped architecture to supply layers at various feature levels with salient item location information. To combine coarse-level semantic information with fine-level top-down route characteristics, they offer a module for feature aggregation (FAM). Following top-down fusion processes, FAMs are able to readily fuse coarse-level GGM data with features of several scales. These two pooling-based modules enhance detail-rich saliency maps by refining high-level semantic properties. In comparison with the other models, they find significant objects with sharper characteristics more effectively. Our approach handles 300 *400 photos at above 30 frames per second[3]. They determine how to extend the function of pooling in convolutional neural networks. Some instances, this disregards the crucial spatial information of the whole input picture.

Bi-DAINet[6] This research presents the Bi-directional Discard-Accept-Integrate Network, often known as the Bi-DAINet, for SOD. To begin the process of incorporating multi-scale feature maps into the network, they begin by constructing a Discard-Accept-Incorporate (DAI) module. In order to reduce the amount of congested noises, the DAI module will remove any redundant information from the layer below it, pull useful information from the layer above it, and then integrate the remainder. They construct a CNN that is capable of enforcing message flow in both the low-to-high and high-to-low directions. Foreground and background regions may be inferred using bi-directional networks. A saliency map that contains fine-grained boundary predictions is produced as a consequence of these findings. A supervising role is played by the edge-preserving loss function in object edge prediction. The effectiveness of our method is shown across five important criteria. They examine the contributions made by every Bi-DAINet component.

For this goal[9], the context-aware refining approach of angular embedding is used. The experimental findings demonstrated the effectiveness of the suggested framework. Moreover, a comparison of the framework with several cutting-edge deep learning-based approaches revealed that it may provide similar and intriguing outcomes.

This study[15]FanjieMenga introduces a novel fusion methods pictures by integrating object area recognition with the Non-Subsampled Contourlet Transform (NSCT). Saliency detection creates the IR image's saliency map. Second, a free areas removal approach extracts the IR image object region. Third, NSCT decomposes source pictures and uses distinct fusion algorithms for low- and high-frequency sub bands. Inverse NSCT generates the principal fused picture. Integrating the original fused picture with the object area yields the final fused image. they compare our technique to others using numerous criteria and find that it improves fused picture quality.

TiantianWang[16] proposes a unique Localization-to-Refinement network. A contextual module is used in the Recurrent Localization Network to weight feature mappings at each point. A recurrent technique is also suggested to obtain contextual information for repeatedly enhancing the convolutional features. A refinement module is used in the Boundary Refinement Network to learn local context knowledge via efficient transmission.

HolgerCaesar[18] discusses When they compare the mean entropy of various datasets, they discover that it demonstrates COCO-linguistic Stuff's richness. They are presently investigating how well a contemporary semantic segmentation algorithm works on COCO-Stuff. They evaluate performance across things and thin classes in the hopes of establishing a baseline for future studies on this dataset.

Projected Augmented Reality is a technique that uses projectors to directly overlay digital material on real things. When texts are projected as AR contents, readability is one of the most important considerations. As a result, numerous projector-camera systems for increasing projected content readability have been developed, in which projection settings are modified based on estimated legibility from taken pictures of projected contents. In this study, NaoyukiKazuta[23] introduces a unique approach for determining projected content readability that accounts for the influence of a patterned projection surface by exploiting edge information in collected pictures of projected contents.

IvonaTautkute[24] proposes a multi-modal search engine, which mixes the linguistic inquiries in this study. Our engine's purpose is to retrieve inside items. Their project enables customers to snap one picture of a room and receive products that are similar/aesthetically shot with a high recall. It also enables you to return goods that complement one other visually and artistically. To do this, our technology combines the findings gained via textual and visual modes. We enhance the mean common score from recovered products around 11 percent thanks to our blending method. Our work is being developed into a Web-based application that will be made available to the public.

ShreyashJoshi[27] presented a smartphone application to allow clients who purchase interior items online virtually view how their space would appear after buying the product.For augmented reality, it makes use of Google ARCore. The picture from the device camera will be accepted as input by ARCore. It will detect planes and then digitally overlay 3D models of the product over the picture, displaying it to the viewer in real time.This will provide a nearly genuine experience of putting the product in the user's house. Products such as wall art, furniture, accessories, and electrical equipment may be shown. The application's intelligence will give appropriate recommendations to the user to assist in the purchase of connected things. This will assist a person in decorating even an empty room from the ground up. The Apriori algorithm was used to propose goods to the user during

DOI: rp-9788770040723.049
association mining. This algorithm will discover the most commonly purchased set goods in order to provide recommendations.

Smartphones may show virtual things into the physical environment at a precise area indicated by a marker thanks to the augmented reality capability. In order to view a 3D home design[33], this study creates a smartphone application for augmented reality. A marker is selected on the floorplan picture that was taken by the smartphone camera, and it is then sent to the server. The server-side application recognises the corners of the picture and generates a vector with 2D coordinates using deep learning and integer programming methods. Utilizing Unity 3D, the client application on the smartphone creates a 3D home model using the 2D coordinates. The model is then overlayed using Vuforia over the floor plan picture marker. According to test findings, it is possible to generate 3D home models that are somewhat accurate and that correspond to the given layout. The augmented reality programme provides a different way to look at a floor plan so that a user may comprehend and interact with the design more fully.

In Ivonatautkute[36] research, they present a multimodal search engine to find objects from a multimedia library that are aesthetically related to the query by combining visual and linguistic clues. Our engine aims to make it possible to quickly retrieve fashion items like furniture or clothing. They merely consider it as an extra detail of the picture in the search and do not account for the situation in which a user searches for "the same blouse but in denim." By modeling contextual connections between aspects of various modalities using a combined neural network architecture, our unique approach, named DeepStyle, addresses these drawbacks. On two distinct, difficult datasets of clothing and furniture, demonstrate the resilience of our strategy. On the evaluated datasets, the engine beats the other approaches by above 20 percent. It is offered via a Web-based application and has been commercially implemented.

These two topics took up the majority of PoonpongBoonbrahma's[38] study. The first involves creating a sizable, stable 3D model or altered by participants who can do collaboration utilizing AR technology and a multiple-markers idea. Creating a smartphone plan for collective augmented reality creation.

Through the use of augmented reality technology, Kaimaris's[40] suggested apps seek to digitally resurrect cultural heritage sites and aid the user in locating locations where areas of interest are partially or completely obscured. Tourists utilize these programmes, which are loaded on mobile devices like smartphones and tablets, to enhance their surfing. The inserted virtual objects into rangu using augmented reality technology are typically 2D or 3D reproduction which serves as the additional objects in space or points of interest that help the user get around more easily and thus revitalize the environment. The most well-known frameworks are also addressed in detail in this article. This report includes data that was experimentally gathered from a number of minor projects that were tried on a variety of expensive and inexpensive equipment.

III. System Architecture

The system architecture of an application refers to the high-level structure of the software, including its components, modules, data stores, and the relationships between them. It provides a blueprint for the design and implementation of the application and helps in determining how the different components will interact with each other.User Interface (UI) component is responsible for presenting the information to the user and receiving input from them.

The UI can be built using react native/Expo user interface elements for mobile applications.Application Server component is responsible for processing requests from the user interface, accessing data from the database server, and returning the results to the UI. The application server can be built using technologies like react native,python.In this the local server act as a database to store the processed imageand then for copy and paste features use predefined python libraries like screenpointer and paste.

For API (ARcore/ARkit) calls our application typically have the components like client,api gateway,API endpoints,And store area.The client component is responsible for making API request to the server and presenting the data to the user.The client can be a mobile application.The API Gateway component serves as the point of entry for API requests and is in charge of forwarding them to the proper API endpoint. Additionally, the API gateway is capable of handling operations like request/response transformations, rate restriction, and authentication. This component, known as the API Endpoint, handles processing API calls and returns the desired data. Building API endpoints is possible with the use of tools like Python.After that our application projection is done use that APIs.

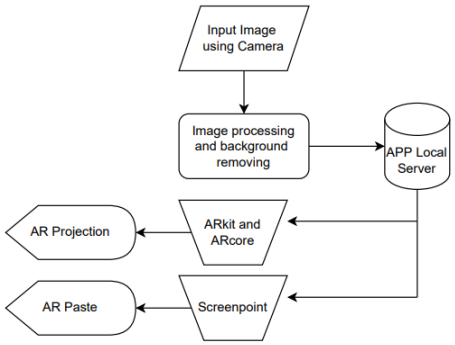

Fig. 1. System Architecture

IV. Methodology

A. *Mobile Application - Expo/React Native*

Expo is an open source framework which comes handy for developers who want to create native universal applications, i.e., android, iOS and the web. The hosted service is known as EAS - Expo Application Services. The applications are coded in javascript and React which has many universal libraries. Expo has a large variety of perks considering one can use it individually or as a team work.

DOI: rp-9788770040723.049

There are many functionalities in expo like camera, location, gyroscope and other useful features in the form of packages.

To install packages, the commands used are :

1) Expo SDK package : npx expo install
2) Camera : npx expo install expo-camera
3) To import into code : import {Camera} from 'expo-camera'

B. Object detection / Background removal

In order to identify unique areas or items that individuals are likely to focus their eyes on in visual situations, saliency detection models the human visual attention mechanism. For this visual job, contextual information is crucial[2].Traditional approaches use specifics concurrently or divided, but their capacity to identify the essential prominent objects in complex scenarios is constrained by the absence of high-level semantic information. As the extraction is in a higher level of systematic news, convolutional neural networks (CNNs) fortunately considerably encourage the models which can detect objects using salient object detection (SOD)[3].From a high level perspective, there are three basic methodologies for saliency detection: bottom-up, top-down, and mixed[9].The main challenge is accuracy of object localization has always been impacted by the complexity of the settings and the variety of the targets themselves. CNNs' deep feature layers are more semantically rich than their shallow counterparts, which is better for object location. The retrieved deep features require a clear separation between the salient regions and the background in order to correctly discover whole objects[14]

- *Deeplab:*Deeplab is an independent service that is in charge of managing the salience detection as well as the background removal tasks. Deeplab makes available an application programming interface (API), which allows you to run our Visual Relationship Detection (VRD) demo on your own photos by using an environment that is controlled by programming instructions. The VRD Application Programming Interface (API) is built using a REST-based architecture and makes use of the standard HTTP response codes. In addition to that, it accepts requests with form-encoded contents, gives results in JSON format, and supports the format.

C. Augmented Reality

The term "augmented reality" refers to techniques and technology that realistically incorporate virtual items into a series of actual photographs taken with a device's camera In contrast to VR. The customer is completely submerged in a digital world. It often takes place in a 3D environment and seeks to add some virtual content to the real world while still providing them with more choices. Virtual objects are designed to provide the user of an augmented reality system with relevant and context - dependent information. Utilizing technological devices with camera sensors, such as glasses, smartphones, tablets, etc., makes this combination conceivable[26].Language and visuals are the most common forms of human information transmission in everyday life, followed by scientific texts.75% of the information that people acquire visually are got by images[31]

- *AR Paste:*Photoshop is the software to be used in this project. Given that the user should transfer the captured image from the phone to the software, user have certain procedures to follow : Firstly, the image is captured by the camera of the device. When the device is pointed towards the screen, the centroid of the picture (x,y) is found by the software using OpenCV SIFT. This module can also be called Screenpoint.
- *AR Projection* Augmented Reality is always portrayed as a headset being worn with a small field of view. But our approach is to use an external camera and project it on the device. The implementation is done using technologies such as ARkit and ARcore. This helps in seamless projection of the image in the desired location or spot. Our project aims to project with adjustable dimensions of the image as needed.

D. Local server

In contrast to a local server, which operates on a different system, a local server in an application is a server that runs on the same physical machine as the programme. Data for an application may be managed and stored on a local server.The local server component handles storing and retrieving data for the application while the application component is in charge of sending requests to the local server and displaying the data to the user. Python-based technologies may be used to create the local server.The main purpose of the server is to store the data and provide to the application when thers is a requested for the process is needed.This system architecture gives a local server in an application a simple layout. The unique specifications of the application being created will determine the components' exact locations and the nature of their connections. On the other hand, this architecture offers a solid foundation for creating a local server that can efficiently handle and store data for an application.

V. CONCLUSION

The surveyed papers have certain downfalls and some of them are Occlusion - two or more objects are too close and seemingly merge or combine with each other.Lighting difficulties - contrast amplification in image-processing applications. Time complexity is high. The suggested systems in the examined publications come from a variety of fields. It is challenging to maintain correct spatial registration of actual and virtual objects. Systems with high accuracy are mostly experimental results.Our proposed system will be an ensemble model which uses both image processing and Augmented reality.

REFERENCES

[1] G.Wang, Y.Zhang,and J. Li, "High-level background prior based salient object detection," Journal of Visual Communication and Image Representation, vol. 48, pp. 432–441, 2017.

[2] N.Liu, J.Han,and M. H. Yang, "PiCANet: Learning Pixel-Wise Contextual Attention for Saliency Detection," In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2018.

[3] J.J.Liu, Q.Hou, M.M.Cheng, J.Feng,and J. Jiang, (n.d.), "A Simple Pooling-Based Design for Real-Time Salient Object Detection".

[4] W.Wang, S.Zhao, J.Shen, S. C. H.Hoi,and A. Borji, “Salient object detection with pyramid attention and salient edges,” Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1448–1457, 2019.

[5] Q.Zhang, Y.Shi,and X. Zhang, “Attention and boundary guided salient object detection,” Pattern Recognition, p. 107, 2020.

[6] C.Yao, L.Feng, Y.Kong, B.Jin, Y.Liu,and L. Li, “Bi-DAINet: Bi-Directional Discard-Accept-Integrate Network for salient object detection,”Neurocomputing, vol. 428, pp. 142–152, 2021.

[7] K. H. Rhee, “Detection of Spliced Image Forensics Using Texture Analysis of Median Filter Residual,” IEEE Access, vol. 8, pp. 103374–103384, 2020.

[8] Z.Yang, S.Soltanian-Zadeh,and S.Farsiu,“BiconNet: An edge-preserved connectivity-based approach for salient object detection. Pattern Recognition, p. 121, 2022.

[9] M. Moradi, F.Bayat, andM. Charmi, “A salient object detection framework using linear quadratic regulator controller,”Journal of Visual Communication and Image Representation, 2021.

[10] X.Qin, Z.Zhang, C.Huang, M.Dehghan, O. R.Zaiane, and M. Jagersand, “U2-Net: Going deeper with nested U-structure for salient object detection,”Pattern Recognition, 2020.

[11] R. Udendhran, M. Balamurugan, A. Suresh, andR. Varatharajan, “Enhancing image processing architecture using deep learning for embedded vision systems,”Microprocessors and Microsystems, 2020.

[12] T. Wang, A. Borji, L. Zhang, P. Zhang, andH. Lu, “A Stagewise Refinement Model for Detecting Salient Objects in Images,” Proceedings of the IEEE International Conference on Computer Vision, October 2017.

[13] Dhanabalan, S. S., Sitharthan, R., Madurakavi, K., Thirumurugan, A., Rajesh, M., Avaninathan, S. R., & Carrasco, M. F. (2022). Flexible compact system for wearable health monitoring applications. Computers and Electrical Engineering, 102, 108130.

[14] Z.Yao,and L. Wang, “ERBANet: Enhancing region and boundary awareness for salient object detection,”Neurocomputing, vol. 448, pp. 152–167, 2021, https://doi.org/10.1016/j.neucom.2021.03.094.

[15] F.Meng, M.Song, B.Guo, R.Shi,and D. Shan, “Image fusion based on object region detection and Non-Subsampled Contourlet Transform”, Computers and Electrical Engineering, 2017.

[16] T.Wang, L.Zhang, S.Wang, H.Lu, G.Yang, X.Ruan,and A.Borji,“Detect Globally, Refine Locally: A Novel Approach to Saliency Detection,” Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018.

[17] G. Xu, Z. Wang, Y. Cheng, Y. Tian, andC. Zhang“A man-made object detection algorithm based on contour complexity evaluation,” Chinese Journal of Aeronautics, 2017.

[18] H.Caesar, J.Uijlings,and V. Ferrari, (n.d.). COCO-Stuff: Thing and Stuff Classes in Context.

[19] Chinese Association of Automation. Technical Committee on Data Driven Control, L. and O., Beijing jiao tong da xue, Institute of Electrical and Electronics Engineers. Beijing Section, IEEE Industrial Electronics Society, and Institute of Electrical and Electronics Engineers. (n.d.). Proceedings of 2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS’18) : Enshi, Hubei Province, China, May 25 -27, 2018.

[20] M.Feng, H.Lu,and E. Ding, “Attentive feedback network for boundary-aware salient object detection,” Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1623–1632, 2019.

[21] Y.Li, L.Cao, J.Zhu,and J. Luo, “Mining fashion outfit composition using an end-to-end deep learning approach on set data,” IEEE Transactions on Multimedia, vol. 19, no. 8, pp. 1946–1955, 2017.

[22] M. F.Hashmi, B. K. K.Ashish, A. G.Keskar, N. D.Bokde,and Z. W. Geem, “FashionFit: Analysis of Mapping 3D Pose and Neural Body Fit for Custom Virtual Try-On,” IEEE Access, vol. 8, pp. 91603–91615, 2020.

[23] N.Kazuta, T.Abe,and T. Suganuma, “Legibility estimation for projected AR contents,” 2020 IEEE 9th Global Conference on Consumer Electronics, GCCE 2020, pp. 615–616, 2020.

[24] I.Tautkute, A.Możejko, W.Stokowiec, T.Trzciński, Ł.Brocki,and K. Marasek, “What Looks Good with my Sofa: Multimodal Search Engine for Interior Design,” 2017.

[25] B.Šumak, M.Špindler, M.Debeljak, M.Heričko,and M. Pušnik, “An empirical evaluation of a hands-free computer interaction for users with motor disabilities,” Journal of Biomedical Informatics, p. 96, 2019.

[26] M. A.Sabri,and Institute of Electrical and Electronics Engineers. (n.d.). 2020 International Conference on Intelligent Systems and Computer Vision (ISCV) : June 09-11, 2020, Faculty of Sciences Dhar El Mahraz (FSDM), Fez, Morocco.

[27] S.Joshi, P.Walavalkar, P.Shetkar, S.Joshi, S.Mahale, S.Bhogan,and K. Sawant,“Product Visualization Using Augmented Reality,”International Research Journal of Engineering and Technology, 2020.

[28] Y.Zhang, P.Pan, Y.Zheng, K.Zhao, Y.Zhang, X.Ren,and R. Jin, “Visual search at alibaba,”Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2018.

[29] Y. Song, R. Koeck, andS. Luo, “Review and analysis of augmented reality (AR) literature for digital fabrication in architecture,” In Automation in Construction, Elsevier B.V., vol. 128, 2021.

[30] D. Rajaratnam, D.M.L.P.Weerasinghe, M. Abeynayake, B. A. K. S. Perera, andJ. J. Ochoa, “Potential use of Augmented Reality in pre-contract design communication in construction projects,”Intelligent Buildings International, 2021.

[31] T. Sheng, “Real-Time AR Technology Assisted High-Resolution Image Processing and its Graphic Design Application,”IEEE Access,p. 8, 2020.

[32] O. V. Kanivets, I. Kanivets,and E. Shmeltser, “Development of mobile applications of augmented reality for projects with projection drawings Machine vision in agriculture View project,” 2020.

[33] M. Auliaramadani, N. Suciati, S. C.Hidayati, H. Fabroyir, andR.R. Hariadi, “Augmented Reality for 3D House Design Visualization from Floorplan,” Image. Proceedings of the International Conference on Electrical Engineering and Informatics, October 2020.

[34] Gomathy, V., Janarthanan, K., Al-Turjman, F., Sitharthan, R., Rajesh, M., Vengatesan, K., &Reshma, T. P. (2021). Investigating the spread of coronavirus disease via edge-AI and air pollution correlation. ACM Transactions on Internet Technology, 21(4), 1-10.

[35] J.D.Lee, H.K.Wu,and C.T. Wu, (n.d.). A Projection-based AR system to display brain angiography via Stereo Vision.

[36] I.Tautkute, T.Trzcinski, A. P.Skorupa, K.Lukasz,and K. Marasek, “DeepStyle: Multimodal Search Engine for Fashion and Interior Design,” IEEE Access, vol. 7, pp. 84613–84628, 2019.

[37] H.Hu, Y.Wang, L.Yang, P.Komlev, L.Huang, X.Chen, J.Huang, Y.Wu, M.Merchant,and A. Sacheti, “Web-scale responsive visual search at Bing,” Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 359–367, 2018.

[38] P.Boonbrahm, C.Kaewrat,and S. Boonbrahm, “Effective collaborative design of large virtual 3D model using multiple AR markers,”Procedia Manufacturing, vol. 42, pp. 387–392, 2020.

[39] EcoleNationaled’Ingénieurs de Sfax, Institutfrançais des sciences et technologies des transports, de l’aménagement et des réseaux, INRIA Sophia-Antipolis, and Institute of Electrical and Electronics Engineers. (n.d.). IEEE Third International Image Processing, Applications and Systems Conference: booklet of abstracts, Inria Sophia Antipolis, France, December 12-14, 2018.

[40] D.Kaimaris, T.Roustanis, K.Klimantakis, I. A.Karolos,and P. Patias, “Possibilities of spatial correlation of 3d models in an archaeological augmented reality application,” International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, vol. 46(M-1–2021), pp. 355–359, 2021.