EducatAR: Dissemination of Conceptualized Information using Augmented Reality and Image Processing

Hrushikesh Ingale[1], Surhud Khataokar[2], Varad Apte[3], Avani Mehta[4], Prof. T.D. Khadtare[5]

[1,2,3,4,5]*Department of Computer Engineering, Sinhgad Institute of Technology and Science, Narhe, Pune, Maharashtra, India*

Abstract: *In the current date scenario, technology is leveraged in all domains to help humans achieve a better perspective about their day to day needs. EducatAR is a mobile application that facilitates and promotes interactive and better conceptual learning. The application will scan the image from a textbook (as a source for now) and fetch an equivalent 3-D model for it from the cloud. The overall functionality of the application will help the students to understand the topic better. The main aim of this application is to ease the process of learning amongst students. The proposed application provides a user-friendly, interactive experience. An in-built camera in the application will scan the image from the textbook. Vuforia and Unity engine have been used to build the app and Blender to design 3-D models. The 3D models will be kept on cloud. A trained Convolutional Neural Network (CNN) will recognize the image and fetch an equivalent model from the cloud. Once the model is fetched, the user would be able to interact freely with the 3D model and get a crystal clear learning experience of the topic using Augmented Reality. It thus enables the students and the entire educational field to achieve a better understanding.*

Keywords: Augmented Reality, Vuforia, Unity Engine, CNN

I. Introduction

Augmented Reality: An enhanced version of reality where live, direct or indirect views of physical real-world environments are augmented with superimposed computer-generated images over a user's view of the real-world, thus enhancing one's current perception of reality. Augmented reality is the integration of digital information with the user's environment in real time. The origin of the word augmented is augment, which means to add or enhance something. In the case of Augmented Reality (also called AR), graphics, sounds, and touch feedback are added into our natural world to create an enhanced user experience.

Augmented Reality (AR) is an interactive experience of a real-world environment whereby the objects that reside in the real-world are "augmented" by computer-generated perceptual

Figure 1:AR in education

Manuscript received October 9, 2001. (Write the date on which you submitted your paper for review.) This work was supported in part by the U.S. Department of Commerce under Grant BS123456 (sponsor and financial support acknowledgment goes here). Paper titles should be written in uppercase and lowercase letters, not all uppercase. Avoid writing long formulas with subscripts in the title; short formulas that identify the elements are fine (e.g., "Nd–Fe–B"). Do not write "(Invited)" in the title. Full names of authors are preferred in the author field, but are not required. Put a space between authors' initials.

F. A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (corresponding author to provide phone: 303-555-5555; fax: 303-555-5555; e-mail: author@ boulder.nist.gov).

S. B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar.colostate.edu).

T. C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba, Japan (e-mail: author@nrim.go.jp).

information, sometimes across multiple sensory modalities, including visual, auditory, haptic, somatosensory, and olfactory. The primary value of augmented reality is that it brings components of the digital world into a person's perception of the real world, and does so not as a simple display of data, but through the integration of immersive sensations that are perceived as natural parts of an environment. From social media filters, to surgical procedures, AR is rapidly growing in popularity because it brings elements of the virtual world, into our real world, thus enhancing the things we see, hear, and feel.

Several categories of augmented reality technology exist, each with varying differences in their objectives and applications. They are:

1) Marker based Augmented Reality

2) Marker less Augmented Reality

3) Projection based Augmented Reality

4) Superimposition based Augmented Reality

An Android Application will be developed using UNITY engine and Blender 3D, an open source computer graphics software for designing 3D objects. These technologies will help to create realistic simulations of desired topics and concepts. The Android application is trained to detect content of image using Convolutional Neural Networks (CNN). Suppose there is a diagram in a book like a Digestive System, then the app will detect the diagram and would be able to render it into a desired augmented model of the same showing a simulation and realistic working of the digestive system. This can be done for any other subjects as well, since diagrams are involved in almost all subjects.

II. Bridging the gap

After extensive research of papers about the existing technologies and inventions(listed at the end of this document) it is clear that there is a significant number of drawbacks in the current system. None of the papers describe an application that has all the three functionalities namely platform independency, image recognition and interactive AR models integrated in a single Android application. Also, no integration of Augmented Reality and Neural Network using Unity Engine was recorded in our survey. In addition to this, there was no official support provided by Unity for integration and use of the above tech-stack in their engine. Coming to the models section, all the systems which use AR models are storing them on the app itself making it heavy and large in size.

III. Overview of our idea and tech used

The app, titled "EducatAR", is an Android application which aims at removing barriers of understanding in e-learning. The app uses the phone's in-built camera to scan images of textbooks, for example a digestive system or a brain. It then recognizes and classifies the image and fetches the appropriate 3D model of the image from the cloud.

The technologies used by the app are listed below:

1) Unity Engine

Unity is a cross-platform real time engine made by Unity Technologies. Unity basically is used to create interactive experiences. Our app is primarily built using Unity Engine.

2) Vuforia

Vuforia is an augmented reality SDK for mobile devices that enables the creation of augmented reality applications. It uses computer vision technology to recognize and track planar images (Image Targets) and simple 3D objects, such as boxes, in real time. Unity Engine offers support for Vuforia for creation of AR apps and we have made use of the same.

3)MobileNet

MobileNet is an architecture which is more suitable for mobile and embedded based vision applications where there is lack of compute power. This architecture was proposed by Google. MobileNets are small, low-latency, low-power models

parameterized to meet the resource constraints of a variety of use cases. They can be built upon for classification, detection, embeddings and segmentation.

4) Blender 3D

Blender is a free and open-source 3D computer graphics software toolset used for creating animated films, visual effects, art, 3D printed models, interactive 3D applications and video games. The 3D models in our app are made using Blender. The models can be exported to Unity in .obj or .fbx format and be used in Android applications.

5) Firebase

Firebase provides a realtime database and backend as a service. The service provides application developers an API that allows application data to be synchronized across clients and stored on Firebase's cloud. The 3D models built in Blender are stored on the Firebase database and fetched at runtime.

6)Tensorflow

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It is used for both research and production at Google.TensorFlow was developed by the Google Brain team for internal Google use. It was released under the Apache 2.0 open-source license on November 9, 2015.TensorFlow offers multiple levels of abstraction so you can choose the right one for your needs. Build and train models by using the high-level Keras API, which makes getting started with TensorFlow and machine learning easy.

7) Tensorflowsharp

TensorFlowSharp provides APIs for use in .NET programs, including C# and F#. These APIs are particularly well-suited to loading models created in Python and executing them within a .NET application.This library is not official and not supported by Google.

System Architecture

The following diagram shows the proposed architecture of EducatAR-

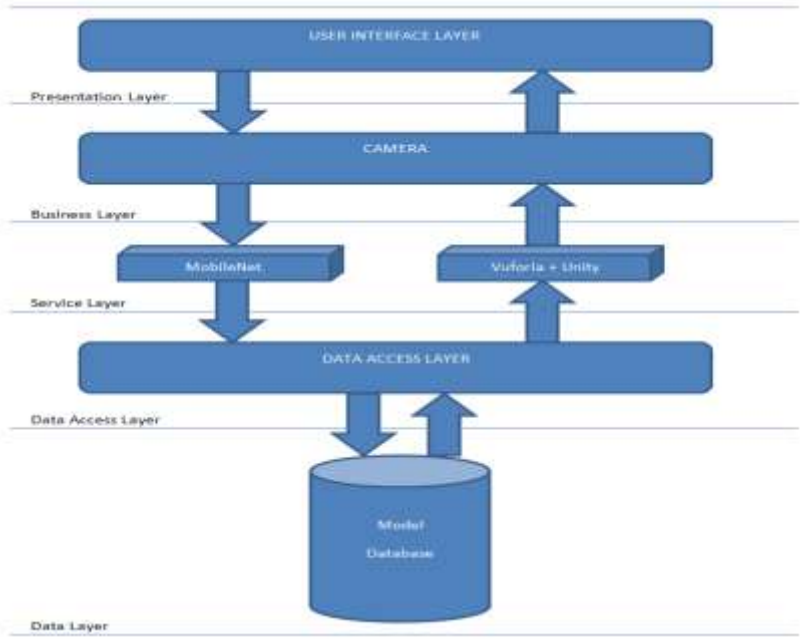

Figure 2: System Architecture

IV. Methodology and working

1) Image recognition and classification

This is the first phase of the working mechanism. In order to identify the image which is scanned using the camera, the system must be trained with appropriate data. The focus of the app is currently on the subject of biology, so the models built are mainly human organs like heart and brain. To train the system to recognize such images, MobileNet v2 has been used. The training dataset was the first 100 images that pop up on Google images after you search for a particular image eg: The first 100 images of a human brain. After the image is recognized, it is classified into the particular category that fits the image. The only restriction is that the scanned image must be undistorted and should have a clear background.

2) Model Building

Model building is an important aspect of this system. The 3D models are built using Blender 3D, a computer software for making 3D figures. Depending on the classified image, the appropriate 3D model is fetched from the database. The models themselves have been constructed from the ground up and are strictly not auto-generated. Currently 5 models related to biology have been made, and will be expanded to other subjects in the future.

3) Model and video fetching

The 3D models which were built in the previous phase are stored on the Firebase database online. In addition to this, a video has also been uploaded which is related to the model. Like if a brain image is scanned, a video related to the functioning of the brain will also be fetched along with the 3D model of the brain. The videos are also stored on the firebase database.

4) Model interaction

Model interaction is basically the final stage of the app working. Once the model is fetched and is visible on the screen, the user can interact with the same. Basic functions like zooming in and out, and rotating the model have been added to make it interactive and more readily visible. Further, live tracking is enabled on the model so that it does not leave the location of its projection, even if the user changes the angle of the camera. Besides this, the video which was stored on the database is fetched and played in a plane adjacent to the model, hence making the learning experience even richer and exciting for the user.

Below are some screenshots of the application that show the working of the same:

Figure 3: Image is scanned and identified

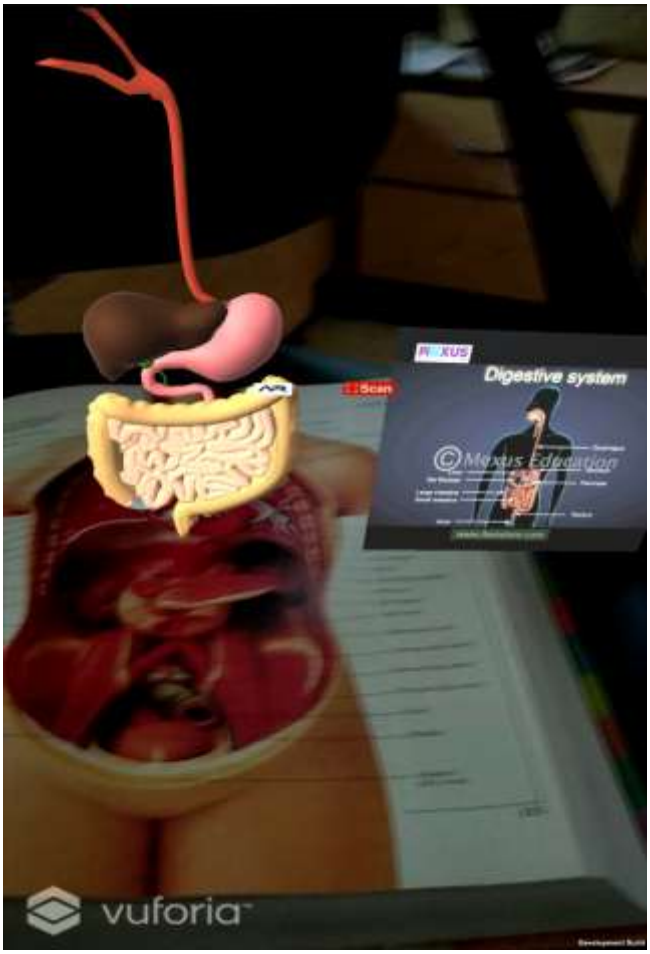

Figure 4: Model and video is fetched from database

As you can see from the above photos, first a digestive system is identified by the app and then the corresponding model as well as video is fetched from the database. The fetched model can be interacted with by the user.

V. Conclusions

In general, the use of Augmented Reality in the field of education and training has somewhat been limited. That being said, there are quite a few applications or systems who have tried to make the learning process interactive. The use of Augmented Reality along with Virtual Reality, termed as "Mixed Reality" by experts surely could be a valuable asset for years to come.

The application designed in this project has the capacity to overcome most of the technological as well as accuracy related drawbacks found in the current systems aiming at using AR in education. The fact that it uses a smartphone powered camera makes it even better in the portability and usability aspect. In addition to that, the application is powered by image recognition APIs, thus eliminating the need to place guidance markers or codes for the computer to understand. Further, by using CNN to train the app to understand and recognize images, it boasts upto 95% success rate of accurate image recognition. In addition to recognizing any image of the reference textbook, the app will show a 3D model of the same, built using Blender 3D. These models will be stored in a administrator-access-only database, strictly not available offline. This ensures that the app data is secure and reliable. This is the first time an attempt has been made to integrate neural network with Augmented Reality, hence making it the pinnacle of cutting edge technology.

Future work on the application will focus on animating the models for better visual aid, and also on making the application interactive in Virtual Reality, a task which is already in development and will be available pretty soon in the next version of the app.

VI. References

[1] Na Liu, Lihong Wan , Yu Zhang , Tao Zhou , Hong Huo ,Tao Fang, "Exploiting Convolutional Neural Networks with Deeply Local Description for Remote Sensing Image Classification", 10.1109/ACCESS.2018.2798799, VOLUME 4, 2016.

[2]Haifa Alhumaidan,KathyPui Ying Lo, Andrew Selby, "Co-designing with Children a Collaborative Augmented Reality Book Based on a Primary School Textbook", 10.1016/j.ijcci.2017.11.005, November 2017.

[3]Baoguang Shi, Xiang Bai, Senior Member, IEEE, and Cong Yao, "An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition", 0162-8828, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 39, NO. 11, NOVEMBER 2017.

[4]Luis Chamba-Eras , Member, IEEE, and Jose Aguilar, Member, IEEE, "Augmented Reality in a Smart Classroom—Case Study: SaCI", 1932-8540, IEEE REVISTA IBEROAMERICANA DE TECNOLOGIAS DEL APRENDIZAJE, VOL. 12, NO. 4, NOVEMBER 2017.

[5]HeenChen,KaipingFeng,ChunliuMo,Siyuan Cheng, ZhongningGuo,Yizhu Huang, "Application of Augmented Reality in Engineering Graphics Education".

[6]ShardulGurjar,HinalSomani, "A Survey on Use of Augmented Reality in Education", 2016 IJEDR,Volume 4, Issue 4, ISSN: 2321-9939.

[7]Haibin Ling, "Augmented Reality in Reality".

[8]Jun He, Peng Han, Huan Liu, Shiying Men, Lu Ju, Pu Zhen, Ting Wang, "The Research and Application of the Augmented Reality Technology", 978-1-5090-6414-4/2017 IEEE.

[9]Lucas Farias, RummeniggeDantas, AquilesBulamaqui, "Educ-AR: A tool for assist the creation of augmented reality content for education", 978-1-61284-890-7/11/2011 IEEE.