IJARIIT INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 7, Issue 4 - V7I4-1407)
Available online at: https://www.ijariit.com

CARKART: A Machine Learning and Augmented Reality application to ease car retailing

Mrunalini M.
mrunalini.1si17cs063@gmail.com
Siddaganga Institute of Technology, Tumkur, Karnataka

Kasthuri C. M.
kasthuri.1si17cs047@gmail.com
Siddaganga Institute of Technology, Tumkur, Karnataka

Pranjal Sharma
pranjal.1si17cs077@gmail.com
Siddaganga Institute of Technology, Tumkur, Karnataka

Gururaj S. P.
gururajsp@sit.ac.in
Siddaganga Institute of Technology, Tumkur, Karnataka

Abstract—CarKart is a user-friendly futuristic android application with Machine Learning and Augmented reality capabilities. It enables the user to upload the image of the car which can be processed to detect the foremost probable car. The machine learning model is trained on the Stanford car dataset containing 196 classes of cars. Hence, this problem is categorised as a multi-label classification problem. Additionally, the users are going to be able to view various models of cars available within the market alongside their specifications and video feed, Fetch a list of nearby showrooms and repair centres, set service reminders, provide their feedback, survey on used cars, check their availability and costs, be able to contact car dealers, thereby introducing the concept of retail. The app also displays the live-size 3D prototype of certain selected car models using Augmented Reality. So, the user can consider all the essential factors before arriving at an informed decision as per their requirement.

Keywords— *Machine Learning, Augmented Reality, Android Application*

1. INTRODUCTION

We currently live in a Techno-fueled world, where ideas, new and radical, with performing abilities magical, in an effort to make human life a spectacle, are being born out of purpose, and on and on it goes.

The Retail Industry dominates the supply chain. Within the Retailer's perception, it involves a broad spectrum of roles and functions targeting at assembling information about the market trends through robust market intelligence, understanding customer requirements and forecasting the demand, fabrication product-related collections and discovering financing opportunities. Following traditional conventional processes during this fast-paced and ever-changing customer preferences and dynamic economic environment could function as a serious drawback and could cause loss of consumer base.

Shifting the difficulty to buyer's perspective, not only does it consume an excessive amount of energy and time going around a dozen showrooms, trying to find an ideal product but also it requires vivid imagination skills to grasp a product before it's brought then trying out different permutations and combinations to verify if landed on the proper decision. Besides, this complete process is manual, constant involvement of both salesperson and customer is required which results in an exhaustive shopping experience which in turn might cause lower conversion rates of business.

The amalgamation of advanced technologies like Machine Learning(ML), Augmented Reality(AR) with the Retail sectors could make us view things in an exceedingly whole new different perception. Not surprisingly, they will be implemented through one among the foremost vital elements of our lives lately, the mobile phones, which may behave as a portal leading to a completely different dimension forming alliances from learning and decision-making abilities of machines, to adding digital components enhancing the live view of the real-world. it's redefining the way we perceive and experience reality. By employing these technologies to the utmost potential, could result in a competitive advantage to retailers.

ML are often used to boost significant connections between external influences and customer behaviours to form better customer service and personalisation, analysis of historical data

© 2021, www.IJARIIT.com *All Rights Reserved*

might be used to invent alternative methods for cost-saving opportunities hence enhancing brand recognition.

On the other hand, AR is often inclined to Let customers "try" the merchandise before they pip out, by letting them interact with a 3D representation of the product, thus creating a singular immersion experience that blends realistic digital objects with the physical environment. Furthermore, AR can play a semantic role in improving brand loyalty, since a decent buyer experience might be a key factor.

1.1 Motivation

The retail industry is constantly changing and there are always new challenges faced by the players in this competitive industry. Customer preferences will always change, sometimes even faster than you can imagine. Economic circumstances, advertisements and competition in the retail industry have a huge impact on consumer demand. By creating a car retail application we can overcome the problems:

1.2 Save user's time and energy

Since in this application we attempt to provide a virtual platform it can save time for them to go to the showroom to learn about the product.

(a) Better decision making: The app displays the exterior features of a car like transmission mode, engine specifications, fuel type, available colours and interior features like mileage, top speed, assistant parking, cruise control and so on, which will help the users in arriving at a better conclusion as per their requirements.
(b) Enhanced customer experience: The app enables a user to capture the image of the desired car and to get the make and model and also view a car model in a life-size 3D model allowing them to blend into the virtual environment making the whole process rather a joyous experience.
(c) Automated processes: Since the details are displayed on their own, the need for a salesperson is reduced considerably.

1.3 The features of the application are summarised as follows

The application provides important features like Smart Searching, Smart Demonstration, Cars Information, a List of nearby car showrooms and others.

(a) Smart Searching- Click a picture of a car, upload and acquire its make and model, hence automating the processes.
(b) Smart Demonstration- View live-size 3D-prototype of certain models of a car both exterior and interior thereby enhancing customer experience.
(c) Cars Information- See different classes, both image and video feed of specific cars with its features and specifications resulting in better decision making.
(d) List of nearby showrooms- Display a list of nearby showrooms. This helps the customers to visit the showroom to buy cars. It helps to find the best car deals.
(e) Others- Add contacts of car dealers, set service reminders, and supply feedback, causing it to considerably reduce users' time and energy.

This paper proposes a fully-fledged car retail application with incorporated machine learning and augmented reality facilities which are aimed to improve the customer experience and satisfaction.

2. LITERATURE SURVEY

The development of the fully-fledged automatic working prototype of the application that has both Machine learning and Augmented reality capabilities is a very futuristic concept and its implementation is not an easy task, It requires critical analysis to identify the exact problem-solving method, which ML algorithm is to be used to provide a better and accurate prediction and inspection of feasible methods of rendering the models to life-size representation. For a better understanding of this concept, numerous works of literature and theories have been proposed and are already in public usage.

According to a comprehensive survey - Interactive retail is one of the biggest technology trends in recent years. Machine Learning-based object detection and Augmented Reality have proven as major technological advancements and have a strong foothold in the retail industry as retail in the modern world can be integrated with the features of object detection and AR to provide a better overall experience to the customer. This research work combines retail functionality with the real-world interaction of augmented reality and machine learning-based object detection by providing a novel experience to the customers. Supermarket shopping is provided in an alluring and interactive manner to customers. [1]

From another survey published in - IEEE Transactions on Neural Networks and Learning Systems , three types of knowledge are useful for knowledge transfer : 1) source domain features; 2) source domain features and the corresponding labels; and 3) parameters of the learned source domain models, which indicate instance-based transfer learning, inductive transfer learning and parameter-based transfer learning, respectively. For traditional machine learning approaches, the ideal choice of the training set to predict a testing instance car should contain cars. However, in the case of knowledge transfer, the training set can just contain some relevant categories rather than cars, e.g., wheels, which are similar to the wheels of cars; bicycles, which share the knowledge of wheels with the car wheels, or even some irrelevant objects, e.g laptops and birds, which seem to have no connections with cars, but share certain edges or geometrical layouts with local parts of a car image hence transfer learning comes into the picture. As the age of big data has come, transfer learning can provide more benefits to solve the target problem with more relevant data. Thus, it is believed that more applications on transfer learning will emerge in future research. [2]

On the other hand, research was conducted in image analysis with an AR headset that used a transfer learning algorithm to reuse the prepared model for a new problem. The general idea is to use the knowledge the model has learned from a task that provides multiple labelled training data for a new task which significantly reduces network training time for retraining networks from the beginning of each variant. [3]

Augmented reality applications merge virtual models and the physical world. These applications are becoming more popular in many verticals, and several SDKs are available to assist in creation. However further development of testing methodologies and tooling is needed. The novel features of an AR user interface is not adequately addressed by traditional testing methods. This research suggests creating test cases that focus on characteristics of perspective, presence, interaction, immersion, persistence, and performance. Automation and machine learning of image detection features also need to be leveraged to assist in the detection of potential defects in AR applications. The revised quality model and the ML-enabled automation framework seek to expand current capabilities and methods to enhance defect detection in AR applications [4]

An important aspect of car classification is feature detection, “Machine learning schemes in augmented reality for features detection” by Ghina Dandachi and colleagues, gives us a clear understanding of it. It also tells us about the method to combine Machine learning and augmented reality. This paper mentions that the Augmented reality domain can be studied from different points of view, but mainly four: features extraction, detection, matching and registration. It discusses and tests several features extraction and classification techniques to recommend the techniques that achieve higher accuracy and real-time results. At the features detection phase, they introduced a simple idea of graph search by introducing the codebook notation and combining it with the machine learning algorithms. At the image registration phase, two methods are proposed, the statistical appearance model and the covariance matrix usage. [5]

In the paper, “An overview of Next Generation Technologies and its applications” published in the 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT). According to this paper, we can implement next-generation technologies to combat 21st-century challenges which have affected citizens globally. This literature has highlighted the working principles and advantages of augmented reality and machine learning along with its implementing technology. Image registration and its method are divided into two stages. Blob detection, Edge detection, corner detection and thresholding are described in the first stage. After that real-world coordinate are restored according to data obtained from the first state. Therefore, we can implement these technologies in public applications.[6]

To augment the virtual image in the real-time environment, we must consider various aspects like the background, an image we want to augment, etc., The IEEE paper on “Mobile Augmented Reality based 3D Snapshots” by Peter Keitler and team gives us the method to augment a 3D image into any environmental background. This paper gives us details about the mobile requirements needed to augment a 3D image in the environment and various technologies used to create AR in mobile applications. It used optical square markers to provide the anchor for reconstructed virtual objects in the scene. A novel approach based on pixel flow highly improves tracking performance. This dual tracking approach also allows for a new single-button user interface metaphor for moving virtual objects in the scene.[7]

To understand the simulation of marker augmented reality using the mobile camera “Augmented Reality Approach for Marker-based Posture Measurement on Smartphones” by Shahin Basiratzadeh and the team gave us a clear insight and method to do it. It discusses the Biomechanics Augmented Reality Tag app detected, tracked, and measured angles and distances between tags as a real-time augmented reality application, and also stored results in a CSV file for more detailed evaluations. This AR smartphone app is viable for posture, range of motion, and distance measurements that are routinely required by clinicians and researchers. Therefore, customizing the mobile application and tag related tools to capture human body posture and range of motion will be the next step in BAR-T development. Using this we can set up the digitalized model of the car in 3D within our mobile camera.[8]

Through all these related works, we opted for a Machine Learning model with a Transfer Learning algorithm to improve the accuracy and also to reduce the duration of training. The related works also gave us great android insights towards augmented reality working. This research paper taking all these insights proposed an android mobile application to improve customer shopping experience concerning cars.

3. EXPERIMENT METHODOLOGY

3.1 Data Description

The dataset selected here is the Stanford car dataset. It consists of 16,185 car images which are divided and stored into 196 different car models classes. The data further is split into 8,144 training images and 8,041 testing images, along with each class split into 50-50 for better training as you can see in Table 1. The classes are represented in the form of make, model and year, eg. the 2012 Tesla Model S or the 2012 BMW M3 Coupe.

The images present in the dataset consist of various car model images. It includes images taken from different angles and orientations’s which helps us add versatility, randomness, and great compatibility for identifying cars. It also includes pictures with different backgrounds which help us to increase the accuracy of prediction. Figure 1 shows the sneak of the dataset used.

Table 1. Dataset Details

Dataset	**Percentage division %**	**Number of images**
Train	50.31	8,144
Test	**49.68**	**8,041**
Total	**100**	**16.185**

3.2 Machine Learning Model - Transfer Learning

Transfer learning (TL) is a research problem in machine learning (ML) that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. For example, the knowledge gained while learning to recognize cars could apply when trying to recognize trucks. This area of research bears some relation to the long history of psychological literature on the transfer of learning, although formal ties between the two fields are limited. From the practical standpoint, reusing or transferring information from previously learned tasks for the learning of new tasks has the potential to significantly improve the sample efficiency of a reinforcement learning agent.

Mathematically, the definition of transfer learning is based on the keywords domains and tasks.

- Consider a domain D which consists of feature space X and the marginal probability distribution of P(X), where X = {x1,x2, xn} ϵ X.
- Given a specific domain, D = {X, P(X)}, a task consists of two components: a label space Y and an objective

predictive function f: X → Y.

- The function f is used to predict the corresponding label f(x) of a new instance x. This task, denoted by T = {Y, f(x)}, is learned from the training data consists of pairs {xi, yi}, where xi ϵ X and yi ϵ Y.
- Given a source domain Ds and learning task Ts, a target domain Dt and learning task Tt, where Ds ≠ Dt, or Ts ≠ Tt, transfer learning aims to help improve the learning of the target predictive function ft(.) in Tt using the knowledge in Ds and Ts.

4. ML MODEL ARCHITECTURE

Transfer learning model architecture is composed of pre-trained models, transferring knowledge of pre-trained models and finally using the top model for training and testing. Here, we start the procedure by loading the input images and pre-training them with resnet-34 which is our pre-trained model. Then later we transfer the processed results and knowledge obtained from this phase to our top model that is the model built using PyTorch which will take less duration as we have the pre-knowledge. Then we can predict the accuracy and get the required prediction.

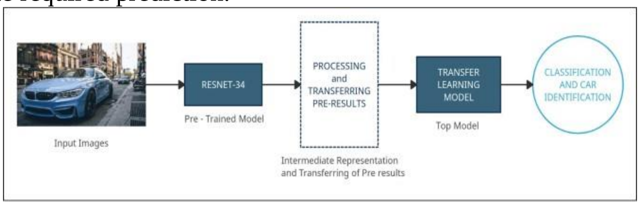

Fig 1. Transfer Learning Model Architecture

5. RESULT ANALYSIS

The proposed transfer learning model is divided into three main categories: loader, training and testing.

- The loader is used to resize the input images so that the memory used to store it will be less.
- In the training phase - the input has been divided into batch sizes of 32 and trained using ten epochs to achieve maximum accuracy.
- During the test phase, the trained model is evaluated to see if the proposed model works alright.

The Table 2 depicts the accuracy, duration and losses obtained after every epoch. The final accuracy obtained is 90%.

Table 2. The epochs and their duration, loss and accuracy. The last column represents the train image accuracy in the network.

Epoch No.	Duration in seconds	Loss	Accuracy	Accuracy % for training images
Epoch 1	1094	3.7831	18.8235	27
Epoch 2	155	1.5685	58.9583	52
Epoch 3	155	0.8333	77.2426	73
Epoch 4	155	0.4969	86.3848	74
Epoch 5	156	0.3360	90.5515	79
Epoch 6	157	0.2314	93.8480	82
Epoch 7	157	0.1074	97.6348	90
Epoch 8	157	0.0688	98.7990	90
Epoch 9	156	0.0608	98.7500	90
Epoch 10	156	0.0553	98.8358	90

5.1 Observation

The model accuracy is analyzed while training and testing to see if the prediction is happening accurately. Let us understand the model precision improvement by plotting three graphs.

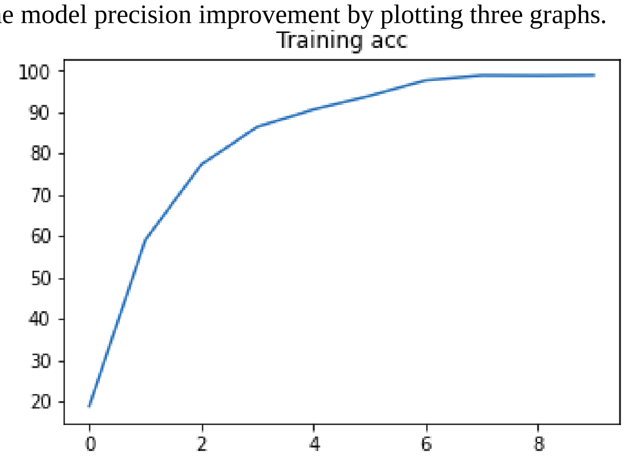

Fig 2. Graph - Training Acc. Vs Iteration

5.1.1 Variation of training accuracy with respect to epochs: Initially, the model accuracy is at 20% and for each epoch, there is a constant improvement. The model accuracy is improving exponentially and it is becoming 100% after 6 epochs. Therefore the average accuracy of the model is 90%. This is depicted in Figure 3. Note: the x-axis represents the iteration.

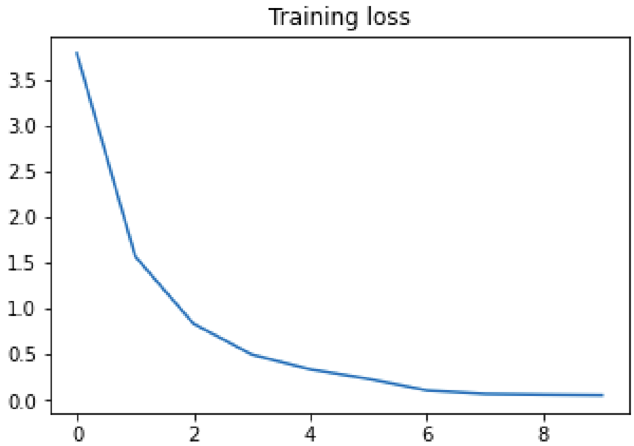

Fig 3. Graph - Training Loss Vs Iteration

5.1.2 Variation of training loss with respect to epochs: Initially, the loss is 35% for epoch 0, and after 6 epochs the loss is becoming zero. This ensures we achieve maximum accuracy. This is depicted in figure 3.

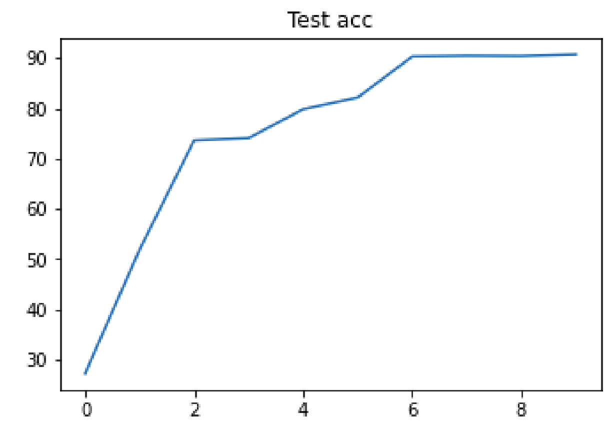

Fig 4. Graph - Test Acc. Vs Iteration

5.1.3 Variation of test accuracy with respect to epochs: Initially, the test accuracy is 0 and after epoch 2 it is reaching the maximum of 70% and it is becoming 100% after 6 epochs. This is depicted in figure 3.

5.2 Android Application - To integrate ML and AR

The application was built using Android Studio, which is an Integrated Development Environment launched by Google used for building applications for the Android operating system. The code was written in Java.

Firebase was used to provide the application with backend infrastructure services like authentication and a real-time database. In firebase, the data is stored in JSON format which makes it easy to store and retrieve the data.

To be able to use the features of the app, the user has to signup and log in to the application using an email id and a password or directly from a google account. By doing so, the user will be assigned a unique-Id and the user details will be stored under it.

The home page is divided into 2 sections

- Image feed - Consists of images of various classes of cars like SUV, Sedan, Hatchback and Convertible which upon clicking fetches the data stored in Firebase dynamically loads a webpage and displays in a separate activity where the user will be able to view the varied selection of used-cars, price, and be able to contact the concerned person. The information stored is obtained as List and displayed as ListView.
- Video feed - Consists of a brief description of the car and upon clicking directs to an activity which plays the video and the detailed description regarding the features and specifications of that car. This is achieved by integrating the SHA-512 key, available in the Gradle script which is unique for an app with Youtube API.
- Profile Page - on clicking the icon the user can view the details through which they have logged in, once provided access to manage the location the address will be automatically picked and displayed.
- Navigation Drawer - The features of the app can be accessed through the navigation drawer.

The application contains the following features:

(a) Carlens - The image of the car will be captured or uploaded through the user and then the prediction regarding the car-make will be made by allowing the machine learning model to work on the image. The machine learning model must be first converted into a.tflite model to make it type-compatible to integrate with android studio.

(b) AR Room - This section directs to a new intent where the user will be able to view the life-size 3D interior and exterior of the car. This is done by integrating the model developed using unity with the android application.

(c) Search Cars - Few cars make features and image URLs have been statically stored in the real-time database, which upon searching, the matched contents are fetched through a java program as a list and are displayed as a result. Picasso library is used to view the image in a circular format.

(d) Nearby Showrooms - Users can get a list of showrooms that are in proximity in a location with the user. This is realized by implementing a google map URL with keywords ‘car’ and ‘showroom’ used in a java program that automatically fetches latitude and longitude coordinates and displays it in a webview. Location permission should be granted here.

(e) Save Contacts - This feature allows to directly save the contact of the car-dealers in the in-built application of the user’s contact application. Permission to access contacts should be granted.

(f) Add Service Reminder - Users can add reminders for their vehicle service by creating an intent that leads to an in-built application in their mobiles where they can manually set the reminder.

(g) Feedback via mail, call or message - Feedbacks can be provided by the user either via call, mail or text message.

(h) About App - A brief description of our application is displayed in this section.

(i) In addition, the user can even reset the password, once forgotten. The password reset link will be directly sent to the registered mail-id

5.3 Augmented Reality - Introduction

The word augment is derived from the Latin word aug(ēre), which means adding or increasing. So as the meaning depicts augmented reality is the enhancement of the real world where we mix the real world with virtual objects. Hence, it is also sometimes referred to as mixed reality. Widely famous game ‘Pokemon Go’ and face filters have incorporated this technology.

Rendering objects start with an AR-enabled phone, it uses a system of camera and in this case 3D depth sensor to understand the physical world. Simultaneous Localization and Mapping (SLAM) is what our phone does many times a second. They map the physical world around us and also keeps information on where we are with real-world the AR device and also senses other characteristics such as intensity of light, colour temperature. In this way, virtual content is placed in the real world that allows us to believe that it is there.

6. BLENDER 3D - MODELLING 3D CAR

Blender is an open-source 3D creation platform. The graphics software tools provided by the software are used in building the replica of various car models. The major feature used here is modelling. Modelling consists of several geometric primitives used in designing various parts of the car. Some commonly used primitives in our design are polygon meshes, circles, subdivision surfaces etc.

Another feature called digital sculpting is used to maintain symmetry and give a fine finishing touch. Finally, the created model is saved in .blend format which stores all objects, materials, scenes and textures in a single file. For our application, we created two complete 3D car blender models which are shown in figure 4.

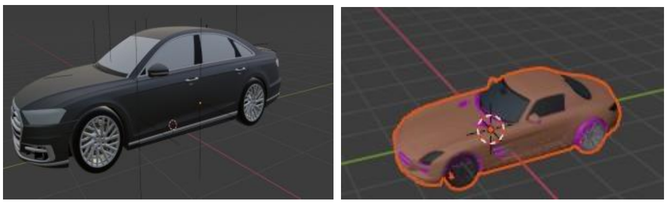

Fig 4: Audi and Mercedes-Benz models created using the blender

6.1 Unity - For building AR Interface

This open-source, cross-platform software allows us to create rich 3D experiences and deploy them across our mobile and wearable AR devices. The created 3d blender models will be exported to unity in .fbx format (film box) which provides interoperability between digital content creation applications. Along with this many features to our application such as the sliding of the door, interior view, car colour, animation are

© 2021, www.IJARIIT.com All Rights Reserved

rendered using unity.

The Unity platform is used to create the Augmented Reality Interface and render the car models. The User Interface consists of the following:

1. Plane Identifier: It is one of the most important aspects of AR. It is used to identify the plane surface to place the 3D model in the real-world environment.
2. Colour palette: It is represented by the three colours circled icon. This is used to change the colour of the car. With this feature, the customers can see the various colour options available for the car.
3. Key icon: This feature is used to represent the opening of car doors to see the interior of cars like the steering wheel, seats and other aspects.
4. Car Menu: This feature is marked as an upward arrow in the down-mid of the screen. By using this feature the customers can select the various range of cars to be viewed in the real world.

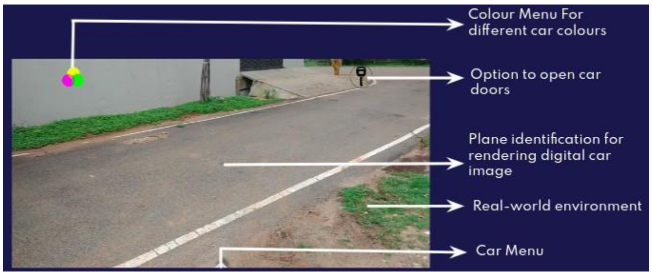

Fig 5. The user interface in a real-world scene

After the user interface is being created the blender models are exported into the unity platform and placed in the car selection menu. The models are then sized and also added animation to enable the above-mentioned features.

Fig 6. Positioning of the car on the plane, car doors opened after the key icon is tapped

Fig 7. Interior view of the car, colour palette selection

7. CONCLUSION

In this paper, we are combining modern technologies like Machine Learning and Augmented reality to achieve a better customer experience. We are using the transfer machine learning model which is exclusively developed for classification and identification, for detecting the car model. Using this, the user will simply capture the image of a particular car and get the details of its specifications within seconds. Augmented reality is used for augmenting the digital model of the car into a real-time environment. Through this, the user can see the life-size model of the car and make better decisions. It also helps the user to see how the particular model of the car looks in their front yard or any other place and also various colours models. All these technologies are combined in an android application which helps the user to easily work through without having to use much effort and thinking. The user interface is smooth to understand and handle. In this application the user just needs to sign up for registering into the app and log in using valid credentials, later he can use the navigation panel to navigate through the application. After logging in he can use the above features to get the car model specifications. The logout button helps the user to log off.

Carkart aims to create a prototype of a car retail mobile application that will have smart searching where the customers just need to capture the picture of a car and get all its specifications at their fingerprints. With Augmented reality, they get to view the live size 3D model of the car in the real world without actually visiting the showroom. By this, it reduces the effort to visit the showroom and look at the model of the car. Through all this customer experience is improved and also satisfied.

8. REFERENCES

[1] Upadhyay, G. K., Aggarwal, D., Bansal, A., & Bhola, G. (2020). Augmented Reality and Machine Learning based Product Identification in Retail using Vuforia and MobileNets. 2020 International Conference on Inventive Computation Technologies (ICICT). doi:10.1109/icict48043.2020.9112490

[2] Ling Shao, Fan Zhu, & Xuelong Li. (2015). Transfer Learning for Visual Categorization: A Survey. IEEE Transactions on Neural Networks and Learning Systems, 26(5), 1019–1034. doi:10.1109/tnnls.2014.2330900

[3] Kozek, M. (2020). Transfer Learning algorithm in image analysis with Augmented Reality headset for Industry 4.0 technology. 2020 International Conference Mechatronic Systems and Materials (MSM). doi:10.1109/msm49833.2020.9201739

[4] Scheibmeir, J., & Malaiya, Y. K. (2019). Quality Model for Testing Augmented Reality Applications. 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). doi:10.1109/uemcon47517.2019.8992974

[5] Wikipedia: Transfer Learning: https://en.wikipedia.org/wiki/Transfer_learning

[6] Mukhi, Rumi, and Maulik Lakhani. 2016. "An overview of next-generation technologies and its applications." International Conference on Automatic Control and Dynamic OptimizationTechniques(ICACDOT) https://ieeexplore.ieee.org/document/7877614

[7] Youtube: Transfer Learning | Deep Learning Tutorial 27 (Tensorflow, Keras & Python) https://www.youtube.com/watch?v=LsdxvjLWkIY&ab_ch an nel=codebasics

[8] Nazir, Souha, Ammar Assoum, Bachar E. Hassan, and Fadi Dornaika. 2016. "Augmented Reality Application: A New Implementation Chain." IEEE International Multidisciplinary ConferenceonEngineeringTechnology(IMCET). https://ieeexplore.ieee.org/document/7777429

[9] R. Agarwal and Y. Pradeep, "Bridging the gap between traditional and online shopping methods for indian customers through digital interactive experience," in Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on, pp. 1184–1189, IEEE, 2013.

© 2021, www.IJARIIT.com All Rights Reserved