

## Allevi-Commerce

**Priyanshu Srivastava**

Scholar, Department of CSE, AKTU Associate,  
Uttar Pradesh, India  
priyanshu.abes@gmail.com

**Priyanka Kaushal**

Scholar, Department of CSE, AKTU Associate,  
Uttar Pradesh, India  
priyankakaushal8802@gmail.com

**Nitanshi Madan**

Scholar, Department of CSE, AKTU Associate,  
Uttar Pradesh, India  
nitanshi.madaan@gmail.com

**Rizwan Khan**

HOD, Department of CSE, AKTU Associate,  
Uttar Pradesh, India  
rizwan.khan@abesit.in

---

**Abstract-** *In this technical world, everyone is trying to alleviate their work by developing or utilizing the latest technologies. We made an effort to take advantages of object recognition technique in e-commerce. About 56.2% population of the world buys online in 2019. So by implementing visual search in e-commerce applications, we are trying to bring more comfort in online shopping. Google provided an API for analysing an image which is named as Cloud Vision API. So we are using Google's Cloud Vision API to detect the objects and its category from an image, in this way customers can purchase according to their choice just by capturing an image. The Amazon Web Services also provided an API for object detection named as amazon Rekognition. In this paper, we have demonstrated the implementation of visual search in online.*

**Keywords-** *Google Cloud Vision, E-Commerce, Object Detection, Visual Search.*

---

### I. INTRODUCTION

Electronic commerce is an online business which perform the activity of purchasing and selling of the commodities and services over an electronic network. There are many E-commerce applications like Amazon and Flipkart. Recently amazon added visual search feature in its application to bring comfort in purchasing, Visual Search is the next descent search and the research is going on this topic. The purpose of this study is to describe the implementation of Google Cloud Vision API and Amazon Recognition API in E-commerce application. Some objects can't be searched easily by using text search method or some objects can't be described in text like jewellery. So to make them search easily we have used visual search in E-commerce application. The user have to capture and upload the image of the item and then that image is processed by the visual search API which will collect the information like name, colour, shape, and another noticeable information from the image. After this user will get a list of similar objects from there he can click on the desired one and can purchase it from our application. We have used Google Cloud Vision API in our application.

### II. INTRODUCTION TO GOOGLE CLOUD VISION

Google Cloud Vision is a powerful API for extracting meaningful information from images like colour, shape, text etc. It rapidly categorises images (e.g. Cat, Bicycle, Tree), and it detects the every objects and faces within the image. It also reads the text available in image. You can easily find out the wide set of objects in your images from animals, electronics or vehicles to thousands of other items, usually found within pictures. Over time new concepts are introduced and improve accuracy. In the result it can give following information about an image with its confidence level:

1. Objects
2. Labels
3. Web
4. Text
5. Properties
6. Safe Search

#### 2.1 Objects

Objects are the materialistic things which can be seen or touched. This API detects all the available objects within an image for example person, man, woman, tree, camera etc.

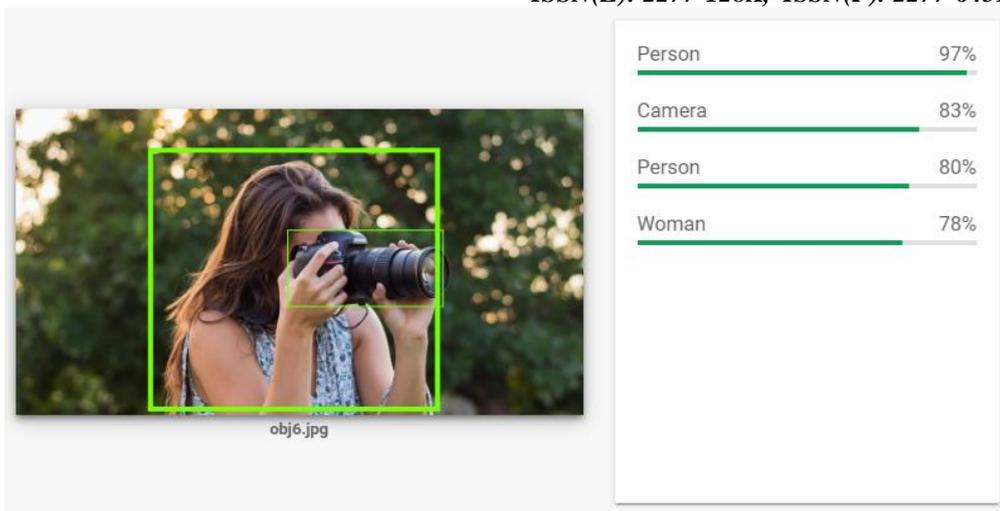


Figure 1. A photo of detected Objects by cloud vision API

### 2.2 Lables

A label gives the information about the objects. It can identify objects, places, activities, animal species, products and more. E.g. Photographer, Digital SLR, camera etc.



Figure 2. A photo of extracted Labels by cloud vision API.

### 2.3 Web

Web detection detects the web reference of an image.

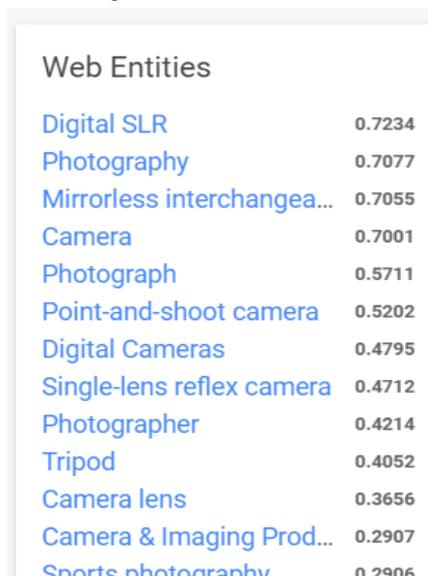
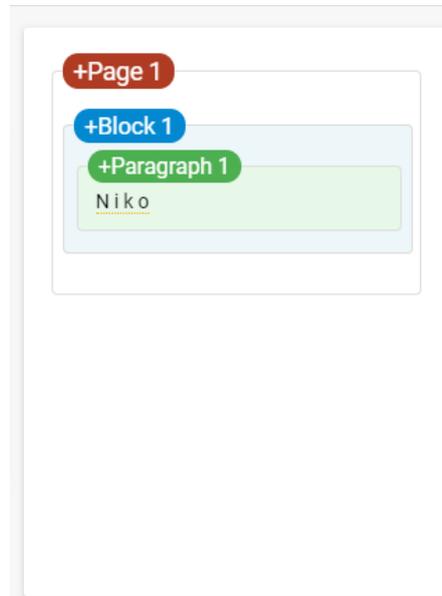


Figure 3. This picture is showing the Web Entities extracted by cloud vision API.

## 2.4 Text

The Google Cloud Vision API can detect Text within an image, this process is known as Optical Character Recognition. There are two annotation features that support optical character recognition:

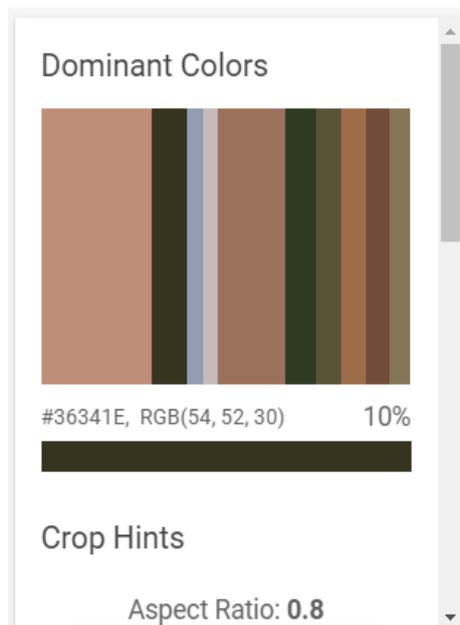
- **TEXT\_DETECTION** detects and extracts text from any image. For example, a photograph might contain a brand name or street name. The JSON includes the entire extracted string, as well as individual words, and their bounding boxes.
- **DOCUMENT\_TEXT\_DETECTION** also extracts the text from an image, but the acknowledgment is optimized for intensive text and documents. JSON includes information about pages, blocks, paragraphs, words, and braves.



**Figure 4.** This photo is showing the extracted text by Cloud Vision API.

## 2.5 Properties

It detects the general attributes of the image like dominant color, crop hints.



**Figure 5.** The general attributes detected by cloud vision API.

## 2.6 Safe Search

Explicit content such as adult content or violent content is detected in safe search detection. This feature uses five categories ("adult", "spoof", "medical", "violence", and "racy") and returns the possibility that everyone is present in a given picture.

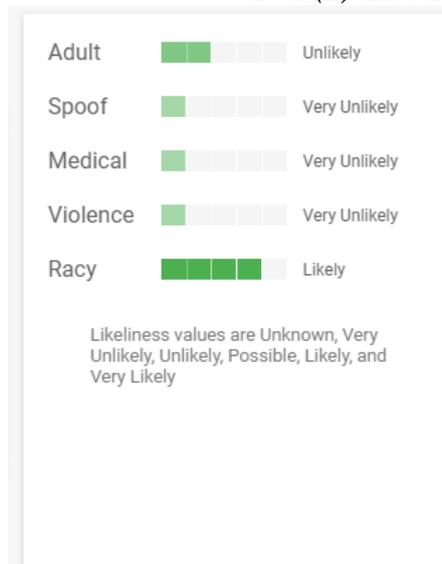


Figure 6. This image is showing the result of safe search detection of cloud vision API.

The Cloud Vision API gives the response in JSON format, the following is the example of JSON report of the request and response of the information detected by this API:

**REQUEST**

```
{
  "requests": [
    {
      "image": {
        "content": "/9j/7QBEUGhvdG9...image contents...eYxxxzj/Coa6Bax//Z"
      },
      "features": [
        {
          "type": "SAFE_SEARCH_DETECTION ",
          "maxResults": 50
        }
      ]
    }
  ]
}
```

**RESPONSE**

```
"safeSearchAnnotation": {
  "adult": "UNLIKELY",
  "medical": "VERY_UNLIKELY",
  "racy": "LIKELY",
  "spooof": "VERY_UNLIKELY",
  "violence": "VERY_UNLIKELY"
},
"textAnnotations": [
  {
    "boundingPoly": {
```

```
"vertices": [  
  {  
    "x": 1144,  
    "y": 453  
  },  
  {  
    "x": 1175,  
    "y": 453  
  },  
  {  
    "x": 1175,  
    "y": 469  
  },  
  {  
    "x": 1144,  
    "y": 469  
  }  
]  
,  
"description": "Niko\n",  
"locale": "sr"  
},
```

### **III. INTRODUCTION TO AMAZON REKOGNITION**

Just like Google Cloud Vision API, Amazon Rekognition API also extracts the information from the images. The image or video is provided to the Rekognition API, and then it identifies objects, people, texts, scenes and inappropriate content. Amazon Rekognition provides highly accurate face analysis and facial recognition on the images and videos you provide. You can find, analyse, and compare different types of user verification, people counting and public safety access issues. It is simple and easy to use API which can rapidly any image or video stored in Amazon S3 bucket. There are two main features of Amazon Rekognition API:

#### **3.1 Amazon Rekognition Image**

Rekognition image is powered by deep learning that extracts or detects objects, faces, texts and scenes and it also recognizes celebrities and inappropriate content. Following are the key features of Rekognition image:

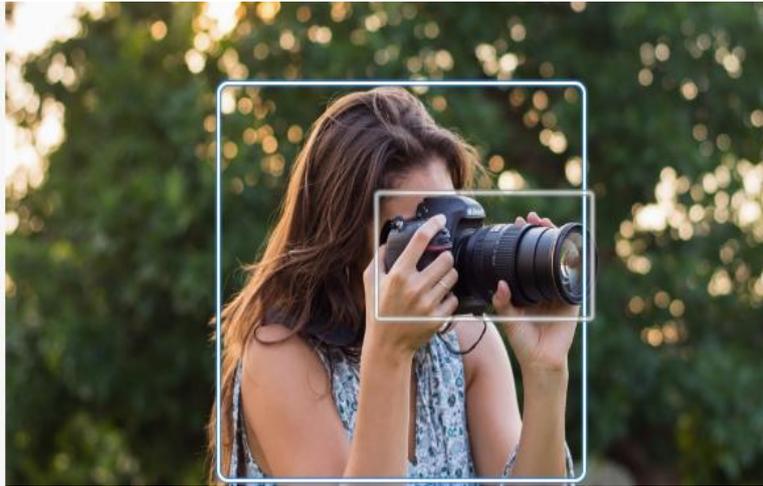
- Object and scene detection
- Facial recognition
- Facial analysis
- Face comparison
- Unsafe image detection
- Celebrity Recognition
- Text in Image

#### **3.2 Amazon Rekognition Video**

Rekognition video is also powered by deep learning service for analysing videos which detects activities, perceives the movement of people in frame, and recognizes people, objects, celebrities and inappropriate content. Following are the key features of Rekognition video:

- Real-Time analysis of streaming video
- Person identification and pathing
- Face recognition
- Facial analysis
- Objects, scenes and activities detection
- Inappropriate video detection
- Celebrity recognition

### 3.3 Demonstration of Amazon Rekognition



Results	
Person	99.9 %
Human	99.9 %
Electronics	97.2 %
Camera	97.2 %
Photography	96.6 %
Photo	96.6 %
Photographer	67.2 %
Portrait	63.3 %
Face	63.3 %

#### REQUEST

```
{  
  "Image": {  
    "Bytes": "..."  
  }  
}
```

#### RESPONSE

```
{  
  "LabelModelVersion": "2.0",  
  "Labels": [  
    {  
      "Confidence": 99.90416717529297,  
      "Instances": [  
        {  
          "BoundingBox": {  
            "Height": 0.8157051801681519,  
            "Left": 0.2739938795566559,  
            "Top": 0.1624937206506729,  
            "Width": 0.474600225687027  
          },  
          "Confidence": 99.90416717529297  
        }  
      ],  
      "Name": "Person",  
      "Parents": []  
    },  
    {  
      "Confidence": 99.90416717529297,  
      "Instances": [],  
      "Name": "Human",  
      "Parents": []  
    },  
    {  
      "Confidence": 97.28659057617188,  
      "Instances": [],  
      "Name": "Electronics",  
      "Parents": []  
    }  
  ]  
}
```

```
{
  "Confidence": 97.28659057617188,
  "Instances": [
    {
      "BoundingBox": {
        "Height": 0.26328498125076294,
        "Left": 0.47704359889030457,
        "Top": 0.38717103004455566,
        "Width": 0.28267747163772583
      },
      "Confidence": 97.28659057617188
    }
  ],
  "Name": "Camera",
  "Parents": [
    {
      "Name": "Electronics"
    }
  ]
},
{
  "Confidence": 96.6955795288086,
  "Instances": [],
  "Name": "Photography",
  "Parents": [
    {
      "Name": "Person"
    }
  ]
},
{
  "Confidence": 96.6955795288086,
  "Instances": [],
  "Name": "Photo",
  "Parents": [
    {
      "Name": "Person"
    }
  ]
},
{
  "Confidence": 67.2134780883789,
  "Instances": [],
  "Name": "Photographer",
  "Parents": [
    {
      "Name": "Person"
    }
  ]
},
{
  "Confidence": 63.34048843383789,
  "Instances": [],
```

```

    "Name": "Portrait",
    "Parents": [
      {
        "Name": "Photography"
      },
      {
        "Name": "Person"
      },
      {
        "Name": "Face"
      }
    ]
  },
  {
    "Confidence": 63.34048843383789,
    "Instances": [],
    "Name": "Face",
    "Parents": [
      {
        "Name": "Person"
      }
    ]
  }
]
}
}
}
    
```

#### IV. INTRODUCTION TO XAMARIN

Xamarin is an open source cross platform application development platform, integrated with Microsoft Visual Studio. With the help of Xamarin Framework, we can build cross platform applications which works on mobile platforms like iOS, Android and Windows Phone. It allows programmers to write their code using C# Language, which is then compiled for every platforms. It supports both patterns, MVC(Model-View Controller) and MVVM (Model-View ViewModel)

#### V. PROPOSED METHOD

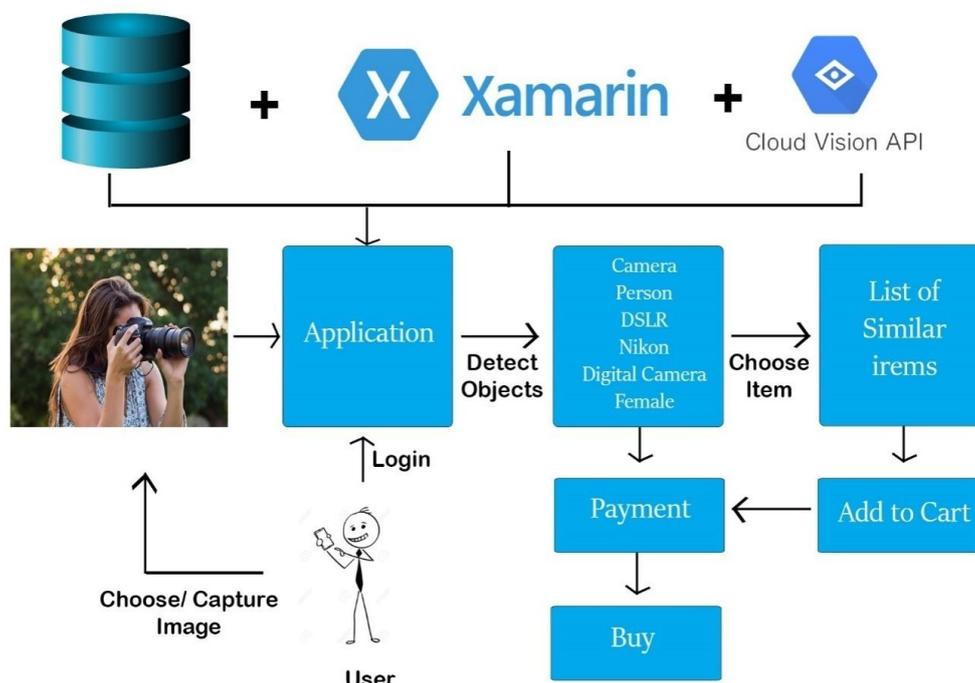


Figure 7. Proposed Method of Visual Search in E-commerce Application

Following are the steps to achieve the desired result

- Developed a cross-platform application using Xamarin which includes basic feature of an e-commerce application with basic features like search, cart, buy etc.
- Now developed a utility for user to capture or upload the image, and then the selected image is sent to Google Cloud Vision API.
- Cloud Vision API analyses the image and detects the objects within that image.
- The list of detected objects is available to the user, here user can select his desired object.
- The name of the selected object is taken from the label and the products with similar name is visible to user, now user can add his desired product in cart.
- At last user can complete his payment and then buy that object.

## VI. CONCLUSION

In this paper, we showed that by using Google Cloud Vision API we can enhance the online shopping experience of the users. It allows users to check anything on the go by just uploading the image. This API detects the objects in the uploaded image and shows the relevant results to the user. We can integrate Google Cloud Vision API very easily in any application. The use of this API is also very simple. The Google Cloud Vision API works well for both i.e. for users as well as for developers.

## ACKNOWLEDGMENT

We are thankful to Head of Department Prof.(Dr.) Rizwan Khan and all the faculty member of the Computer Science and engineering department of ABES Institute of Technology for motivation and encouragement which helped us to complete this review paper.

## REFERENCES

- [1] GAFFAR KHAN, Abdul. Electronic Commerce: A Study on Benefits and Challenges in an Emerging Economy. Global Journal of Management And Business Research, [S.l.], feb. 2016. ISSN 2249-4588.
- [2] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
- [3] C.-Y. Lee, S. Xie, P. W. Gallagher, Z. Zhang, and Z. Tu, "Deeply supervised nets.," in AISTATS, vol. 2, p. 5, 2015.
- [4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pp. 248– 255, IEEE, 2009.
- [5] Chen, Shih-Hsin & Chen, Yi-Hui. (2017). A Content-Based Image Retrieval Method Based on the Google Cloud Vision API and WordNet. 651-662. 10.1007/978-3-319-54472-4\_61.
- [6] Vision AI. <https://cloud.google.com/vision/>
- [7] Amazon Rekognition <https://aws.amazon.com/rekognition/>
- [8] Mukesh Prajapati, Dhananjay Phadake & Archit Poddar, Study on xamarin cross-platform framework, International Journal of Technical Research and Applications, Volume 4, Issue 4 (July-Aug, 2016), PP. 13-18.