Hi! this is team **Innovative Incarnates** presenting **Virtual Eyes**

The socially innovative application meant to be used by visually impaired people, who may control the app via voice commands. The app basically takes an image of the view in front of the person and after running an Image processing Classifier (which we trained on an Nvidia GeForce GTX 1060 6GB GPU for almost 3 days) recites the various objects present in the scene ahead of him. Also, it tells the person about the distance between the camera and the object by creating a depth map, for which we use **3D Photograph Stereo Vision Camera Lens**, which allows a single camera to take photos of the same scene from 2 different views, displaced by a short fixed distance.

**Workflow:**The user asks the app to describe the scene in front of it, which then   
1) Recognizes the voice command   
2) Sends the image of the scene to server in the form of a string of Base 64 format which is then converted to ‘jpg’ format for further processing  
3) Image Processing model based upon YOLO (You Only Look Once) classifier for real time object detection which classifies different objects in the image by running a 30 layer Neural network with 24 convolutional layers and   
4) Sends the processed image (by converting it back into a string of Base 64 format ) to the mobile phone of the user.  
5) As the image reaches the cellphone (this takes almost 3 to 4 seconds), the user is informed about the scene using speech interface.

**USE CASE:**The user does not need to be dependent on someone else to locate the objects near him. He can simply give the command to the app to tell him about the scene before him and the app will do it for him. He can see the views around him by just talking to his device.