

# Enhancing Object Detection with Text-to-Speech

Machine Learning for Android (Summer 2023)



## Introduction

MediaPipe samples with a goal of showing the fundamental steps involved to create apps with our machine learning platform. These samples include apps like:

- Audio Classifiers
- Object detection
- Language detector

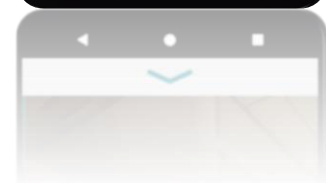
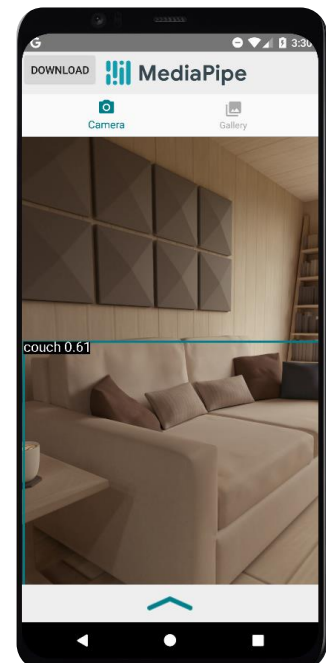
We have chosen to work with Object detection app and enhance it with a text to speech feature.

## App view

The application is fairly simple to use. The user will interact with only the main activity where the camera and the results are displayed.

You can either use the phones camera or use an image from gallery.

Also a download button for storing all objects detected as an audio file (.WAV)



## Contribution

1) The voice generated by AI was implemented using android's speech library, when an object is detected an handed to the results. The object is converted to a string and sent to the instance of Text to Speech which is initialized as the view is created. We make sure the TTS has finished the sentence before sending another word.

However the queue is *flushed* after each word and since the object detection process is faster than any speech speed some objects may not be spoken and therefore not saved in the audio file.

2) Saving audio of speech in .wav format. This was implemented by creating a new button for it. We use a strings hashmap for rendering the text to speech later on by using the **synthesizeToFile**. The audio is then stored (by default) in the devices downloads directory

```
Environment.getExternalStoragePublicDirectory(DIRECTORY_DOWNLOADS)
```

## Testing

For downloading:

**Android version below 5.0:** No permissions required works automatically

**Android version 5+:** Storage permission required

**Android version 10+:** May not work after permission due to new software restrictions

Text to speech: No issues

## References

<https://32net.id/bukaheula/share/iO3trUXFOxXudzkrUSXFeBKx1MeFT3dSyBdq9Xmx.pdf>

<https://developer.android.com/reference/android/speech/tts/TextToSpeech>

<https://stackoverflow.com/questions/26187851/android-how-can-i-save-longer-tts-wavefiles-from-synthesizeToFile>