**Capstone Project**

**Document Skeleton**

# Process overview

The following diagram shows the overall end-to-end process for defining, designing and delivering the Capstone project.



Note: The following are the candidate sections of the document. They are presented here for guidance. Questions in each section could be used as possible aspects to cover. Some questions may not be applied to each project. On the other hand, additional information may be needed.

# Problem statement

Individuals who have a disability currently face immense obstacles through their lives. They are more likely to experience more psychological distress, be unemployed and often face a higher cost of living. Although assistive technologies have been present throughout most of the 19th and 20th century, a majority are outdated and primitive at best.

However, technology has set the stage for a revolution in the field of assistive technology. Just as how the decrease in computing cost made personal computers ubiquitous today, cheap computing can allow for the development of advanced assistive devices to be made available to the greater public. There is now an opportunity for businesses to position themselves to develop these technologies and offer them to the public. This gives individuals with disability a hope at better standard of living, possibly even to a state where they can live life as if they did not have the disability.

Therefore the problem statement is:

**Can businesses create new forms of assistive technologies that can be mass produced for individuals with disability, with the goal of simulating a life without disability as closed as possible**

For the purposes of the capstone, the problem statement is further refined to understand if new forms of assistive technologies can assist those who have **visual disabilities**.

# Industry/ domain

The proposed developments in assistive technologies sit across the intersection of the medical and technology industry. Currently assistive technologies in the medical field focus on accessibility, often not employing new technology trends(e.g wheelchairs, crutches). In the technology industry, there is a spread of companies, ranging from niche players and research bodies at universities who focus on advanced prosthetics to start-ups who bring the element of accessibility to their technology.

The goal to success is to operate on scale. Assistive devices must be designed and produced in a way that the costs are kept reasonable and that there is a low learning barrier in order to use it effectively.

# Stakeholders

The main stakeholders considered are any individuals who experience a form of visual disability.

# Business question

**Can business provide assistive technologies that can reduce visual disability in a low cost way?**

To answer the above business question, use cases were assistive technologies were able to make the greatest impact was identified.

* Counting Money
  + Often, counting money is difficult for the visually impaired, and this often leaves them more likely to be short-changed
* Medicine labels
  + As labels are difficult to read, the risk of injury from ingesting a high dosage or taking the wrong medicine is quite high
* Identifying colour
  + The ability to identify colour is important to perform the basic chores. Therefore those who are visually impaired are unable to independently perform these themselves

To answer the business question, a proof-of-concept application will be built using computer vision that will contain the below functionality:

* Counting money through object detection
* Colour identification through clustering
* Reading medicine label through optical character recognition

The POC application will also come with text to speech capabilities, that will allow the outputs of the models to be verbally presented to the user.

# Data question

Data will be required to train a model to correctly detect the different notes and coins that are currently being circulated as a part of Australia’s currency.

# Data

To source the above data, images were downloaded from Google images. Image augmentation was then applied to increase the number of images. In total, 812 images were used.

In terms of the condition of the data, resolution and image size was an issue, as most pictures are of low quality which affected the results of the training

# Data science process

## Modelling

To fulfil the use cases, three models were created.

**Counting money – Object detection**

For object detection, a YOLOV4 model was used and trained as it provides the best in class performance. The POC utilises the darknet implementation of YOLOV4, which allows for greater learning and provides custom method that greatly increase accuracy with low/no impact in inference time.

To train the model, an image file and a corresponding text file containing the bounding boxes of the area of interest was generated. The method of generation was a labelling tool that allows users to mark the object of interest and automatically generating the bounding box coordinates

Model metrics are as below:

|  |  |  |
| --- | --- | --- |
| **Precision** | **Recall** | **F1-score** |
| 0.93 | 0.89 | 0.91 |

Overall the model has a mean average precision of 90.06%, with a prediction time of 3 seconds for 122 images.

**Colour detection – Clustering**

For colour detection, a clustering model was developed. The clustering model was designed to only find 2 clusters as the goal is to find the dominant colour in the picture. Based on results on the results of the clustering model, a sorting algorithm was created to translate the RGB values to basic colours.

**Reading medicine label – OCR**

For reading medicine labels, an OCR model was utilised. The OCR model utilised CRAFT based detection, where deep learning is used to detect characters and the area they belong to instead of using rigid bounding boxes. To recognise characters, a ResNet convolutional neural network was used, along with a connectionist temporal classification model to decode the results. A LSTM model was only used to maintain the sequence of characters and words.

All models were created offline and trained using a physical GPU.

## Outcomes

The POC application has successfully demonstrated that a low cost assistive technology can be developed by leveraging on the increasing advancements and decreasing cost of technology.

In regard to counting money, the model was able to successfully detect and count the total number of notes provided in the picture.



For both colour detection and the reading of labels, those functions were also successfully in fulfilling their use cases. (for sample test cases, please see the jupyter notebooks that are provided in the same google drive)

## Implementation

Key notes for implementation are:

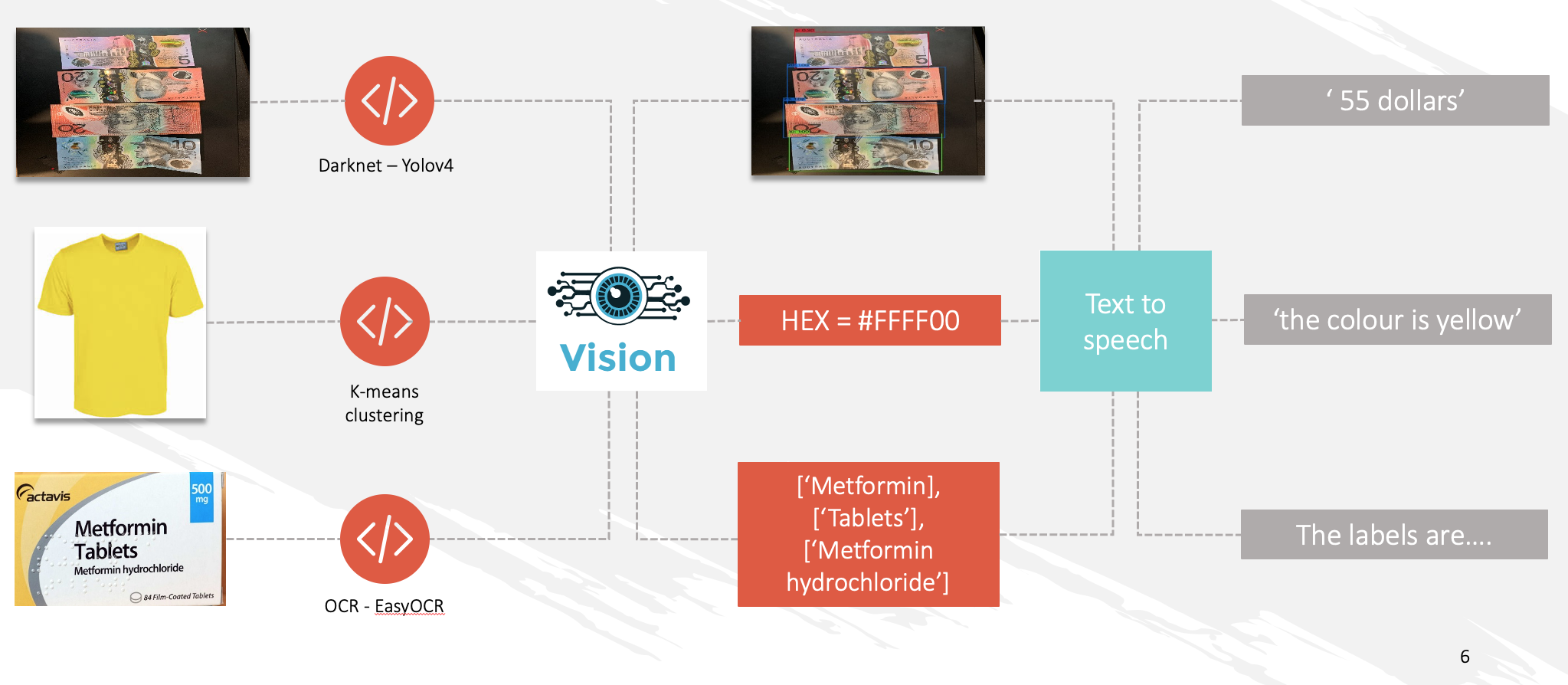
* Voice activated commands
  + Currently buttons are still used to prompt the application so functions can be launched. However an improvement can be to integrate mobile voice assistants that are found of devices to launch the application
* Mobile Deep Learning
  + Although the application is currently developed on a PC platform, the true advantage would be the ability to provide these functions on a mobile platform as mobile devices are ubiquitous in the world today. Therefore, a future implementation would consider mobile deployment, leveraging frameworks such as TensorflowLite and utilising edge/cloud computing to bring processing power to the average user.
* Power efficiency
  + As this application may be used over a long period of time, power efficiency is critical in ensuring that the user will be able to continuously access the capabilities of the application

# Business answer

Through the success in meeting all the use cases, the application has demonstrated that:

**Yes, businesses are able to revolutionize the field of assistive technologies and create new ways on how humans can overcome disabilities**

# End-to-end solution



# References

* Alexey AB – Darknet YOLOv4
  + <https://github.com/AlexeyAB/darknet>
* JaidedAI – EasyOCR
  + https://github.com/JaidedAI/EasyOCR