



أكاديمية سدايا
SDAIA Academy

SMART CITY

PROJECTS PROPOSAL - T5 BOOTCAMP



Table of Contents

Smart City Projects Proposal	2
Speech recognition Project	3
Emergency Sirens Detection Project	4
Waste Classification Project.....	5
Face Mask Detection Project.....	6
Arabic Sign Language Translation Project.....	7
Data Description.....	7
Speech Emotion Recognition Project	8
Data Description.....	8
Road Damage Detection Project.....	9
Tools.....	9
Conclusion.....	9

Smart City Projects Proposal

Under Vision 2030, Saudi Arabia aims to become the place where the best of data and AI is made reality. Under the strategy, Saudi Arabia will implement a multi-phase, multi-faceted plan that includes skills, policy and regulation, investment, research and innovation, and ecosystem development. This will enable the country to become a global hub for data and AI, and to rank among the top data-driven economies in the world. Moreover, Artificial Intelligence and machine learning will be the most important products and means available to deal with the growth and changes of future cities, which will often be smart cities. Artificial intelligence is one of the most important engines of the Fourth Industrial Revolution, and it will not only change the business world but also transform cities and the ways in which we operate and live radically. However, In order to contribute on 2030 vision, we have created seven subprojects that accomplish one main idea. The main goal is to show how we can take advantage of deep learning in a smart city. Let us remember that the primary goal of artificial intelligence in general is to improve and develop ways of life for humans, as well as for smart cities by rapidly obtaining and analyzing large amounts of data very quickly to make more informed decisions aimed at improving the ways of life of the inhabitants of these cities in various fields. The following are explanations of the projects along with the responsible teams.

Speech recognition Project

By: Ahmed Almuaybid, Abdulrahman Alrifae, Naif Alzahrani

Speech recognition, also known as automatic speech recognition (ASR) or speech-to-text conversion, is the way that we can process human speech and turn it into a written format. While it's common to confuse it with voice recognition, speech recognition focuses on translating speech from verbal to textual, while voice recognition only seeks to determine the voice of the individual user.

This project will focus on building a speech recognition system that understands simple spoken commands.

Data Description

The Speech Commands Datasets are provided by TensorFlow which includes 65,000 of 30 short words, by thousands of different people.

Emergency Sirens Detection Project

By: Mai Aljuaid, Basma Alduaiji

Due to heavy traffic, emergency vehicles are delayed in responding to an incident. This costs people their lives and therefore, this project will assist in solving this problem by detecting emergency sirens and clearing the road by changing the traffic signals according to the direction of where the emergency vehicle is headed before reaching the signal.

Data Description

The dataset that will be used in this project includes two wav files. One represents emergency sounds (Siren) and the other one represents non-emergency sounds (normal noise). The emergency wav file is 23 minutes long and the non-emergency file is 27 minutes long. As these are audio files, it will be read in python as arrays and will be split into chunks. Feature extraction will be implemented during the project to use in modeling.

Waste Classification Project

By: *Rahaf Alqahtani*

Waste is an important global issue. Increasing amounts of waste are being generated as the world's population and living standards rise. Therefore, recycling is the solution so that the recyclability of a material depends on its ability to restore the properties it had in its original state for further use. Which helps reduce: energy use, air pollution (from incineration), and water pollution (from landfilling). So I will classify waste as Organic or Recyclable using Convolutional Neural Network (CNN).

Data Description

The data that will be used in this project has been taken from kaggle website (<https://www.kaggle.com>). Contain more than 25K images that segregated into two classes (Organic and recyclable).

Face Mask Detection Project

By: *Muneera Alshunaifi, Tarfah Alabbar*

Under the light of the corona pandemic and the health restrictions, wearing a face mask is the best thing to avoid the exposure to the infection so most countries decided the rule of wearing masks in public places and punishes the people who don't wear it for their own safety. In this project we will create a deep learning model that detects the people who are wearing masks and who are not. The developed model will be CNN, in which we find it the most suitable for our problem and for detecting images.

Data Description

To perform this project, a large dataset of masked and unmasked faces are necessary for training deep learning model. The data set obtained from Github, it contains 67,049 images with Masked Face and 66,734 images with Incorrectly Masked Face at 1024×1024 pixels.

Arabic Sign Language Translation Project

By: *Hayat Aldhahri, Juri Alsayigh*

Sign language recognition is an important pillar for the development of societies. There are multiple solutions and approaches to improve and develop sign language recognitions systems. One approach is through machine learning. Developing Neurol Network Model a robust visual recognition algorithm for ArSL (Arabic Sign Language) will aid improving the communities with hearing disabilities living in a smart city.

Data Description

The dataset contains 54,049 images of Arabic Sign Language alphabets. The images were performed of 32 Arabic signs and alphabets (class).

- ArASL_Database_54k_Final.zip: contains all the images for the 32 signs and alphabet.
- ArSL_Data_Labels.csv: Label of each ArASL image.
- Signs_32_New.png: sample of each of the 32 ArSL

Speech Emotion Recognition Project

By: *Rawabi Alkhalfaf, Dimah Albunayyih*

Speech Emotion Recognition (SER) is the act of recognizing human emotion from speech. This is the advantage of the fact that voice often reflects underlying emotion through tone and pitch. SER can be beneficial for call centers that are giving health care support for elderly people and emergency call centers. The proposed project is based on speech emotion recognition, giving greater priority to calls with emotions such as fear, anger, and sadness, and less priority to calls with neutral speech and happiness.

Data Description

There are four different datasets that will be used in this project.

- Crowd-sourced Emotional Multimodal Actors Dataset (CREMA-D)
It is a data set of 7,442 original clips from 91 actors, and the sentences were presented using one of six different emotions (Anger, Disgust, Fear, Happy, Neutral, and Sad)
- Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS)
It is a data set of 1440 audio files from 24 actors, 60 trials per actor. Speech emotions include (calm, happy, sad, angry, fearful, surprise, and disgust expressions).
- Surrey Audio-Visual Expressed Emotion (SAVEE)
It is a data set of 480 files recorded from four native English male speakers. Emotion has been described in seven categories, which are (anger, disgust, fear, happiness, sadness, surprise, and neutral).
- Toronto emotional speech set (TESS)
It is a data set of 2800 audio files that contains a set of 200 target words spoken by two actresses (aged 26 and 64 years). Emotion has been described in seven categories, which are (anger, disgust, fear, happiness, pleasant surprise, sadness, and neutral).

Road Damage Detection Project

By: Waleed Aldalham, Abdulaziz Wali

In a real-world scenario, when the road managers from a governing body need to repair such damage, they need to clearly understand the type of damage in order to take effective action. So, using deep neural networks with images captured, road damage will be detected and fixed by the governing body or whom are responsible.

Data Description

This dataset is composed of 9,053 road damage images captured with a smartphone installed on a car, with 15,435 instances of road surface damage included in these road images. Tool will be used to perform our task is Python and its helpful libraries.

Tools

The tools used to accomplish the seven project are provided as modules in Python such as Pandas, Matplotlib, Seaborn, NumPy, Librosa, IPython, Keras, Scipy, and sklearn, Opencv-python (Video Capturing), LabelImg library (Training Images Labeling tool), Tensorflow (Tensorflow Object Detection), sounddevice.

The python libraries support multiple data analysis and cleaning methods to ensure the data is clean and ready to be visualized. In order to make the model more user-friendly, Flask or Android Studio might be deployed as tools to develop a web or mobile application.

Conclusion

To conclude, Smart cities can be new cities that have been designed and created in a smart way from the beginning, or a traditional city that has gradually been transformed into a fully smart city. From the previous proposed projects, we hope that they will contribute to our understanding of deep learning and be a core to future work in the scope of Saudi Arabia to help achieve the objectives of AI under 2030 vision. Therefore, we would like to get your approval on these bundles of projects in order to carry on with the next step.

