หมายเลขโครงงาน	

แบบฟอร์มขอบเขตโครงงาน Pre-Project

ชื่อภาษาไทย ระบบช่วยบรรยายภาพด้วยเสียงสำหรับผู้พิการทางสายตา

ชื่อภาษาอังกฤษ Audio Description System for the Visually Impaired

โดย

นางสาว ชนากานต์ พ้นภัย รหัสนักศึกษา 65010178

นางสาว ชนิดา สมแก้ว รหัสนักศึกษา 65010181

นางสาว ดีนาน จันทร์เชื้อ รหัสนักศึกษา 65010341

อาจารย์ที่ปรึกษาหลัก

(ผศ.ดร. สมปอง วิเศษพานิชกิจ) ลงนามวันที่ 27 / 11 / 2567

อาจารย์ที่ปรึกษาร่วม (ถ้ามี)

(ผศ.ดร. นภัทร สระเอี่ยม) ลงนามวันที่_____/_____/_____

วัตถุประสงค์โดยคร่าวของการนำเสนอโครงงาน Pre-project

โครงงานนี้มีเป้าหมายในการพัฒนาระบบที่ช่วยผู้พิการทางสายตาให้สามารถรับรู้สิ่งแวดล้อมรอบตัวได้อย่างถูกต้อง โดยใช้หลักการแปลงข้อมูลภาพเป็นเสียงบรรยายระบบจะประมวลผลข้อมูลภาพอย่างรวดเร็วและแม่นยำเพื่อสร้างคำอธิบายที่เข้าใจง่ายและเหมาะสมกับสถานการณ์จริงเป้าหมายคือการช่วยลดความกังวลในสถานการณ์ที่ไม่คุ้นเคย และส่งเสริมความปลอดภัย รวมถึงเพิ่มความสะดวกสบายในชีวิตประจำวันให้กับผู้พิการทางสายตาให้ผู้พิการทางสายตามีคุณภาพชีวิตที่ดีขึ้น และสามารถใช้ชีวิตได้อย่างอิสระมากขึ้นในสังคม

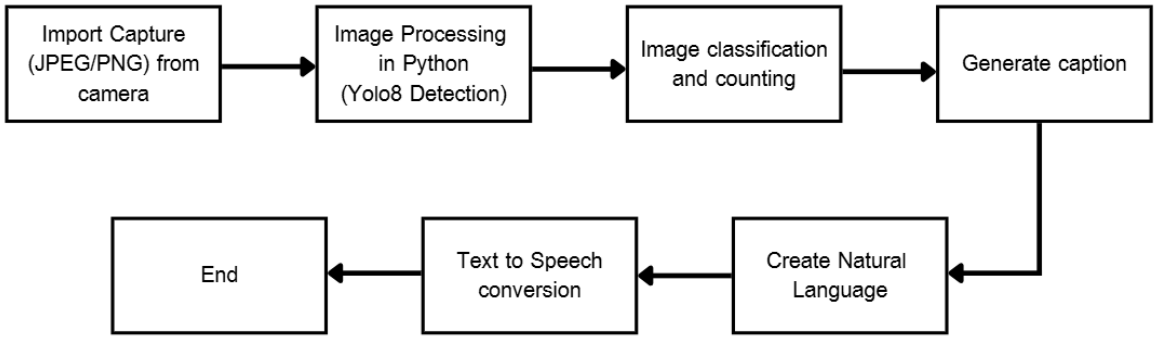
ขอบเขตของโครงงาน Pre-project

โครงงานนี้จะศึกษาเกี่ยวกับเทคโนโลยีที่ใช้ในการแปลงข้อมูลภาพเป็นเสียงบรรยาย โดยจะใช้เทคโนโลยี YOLO ร่วมกับ Open Image Dataset เพื่อเพิ่มความแม่นยำในการตรวจจับวัตถุ โดยการฝึกโมเดลบนชุดข้อมูลที่มีการจำแนกประเภทวัตถุที่หลากหลาย ในขั้นตอนของการออกแบบระบบ จะพัฒนาโปรแกรมที่รับภาพจากกล้องและแปลงข้อมูลภาพที่ได้เป็นเสียงบรรยายผ่านลำโพงขนาดเล็ก เพื่อให้ผู้ใช้งานสามารถได้ยินคำบรรยายจากระบบอย่างชัดเจน และเพื่อให้การทำงานของระบบสะดวกและประหยัดพลังงาน จะใช้ raspberry pi ซึ่งเป็นไมโครคอนโทรลเลอร์ที่เชื่อมต่อกับกล้องและลำโพงได้อย่างมีประสิทธิภาพ นอกจากนี้ ระบบจะถูกออกแบบให้

หมายเลขโครงงาน	

สามารถใช้งานได้เฉพาะในพื้นที่ที่เหมาะสม เช่น บริเวณทางเดินเท้าริมถนน เพื่อเพิ่มความปลอดภัยและความเหมาะสมสำหรับการใช้งานของผู้พิการทางสายตา

บล็อกไดอะแกรมของโครงงานที่นำเสนอ



แผนการปฏิบัติงานตลอดภาคการศึกษา

ช่วงการดำเนินงาน	**แผนงานที่จะดำเนินการ**	
เดือนที่ 1 (ธ.ค. 2567)	สัปดาห์ที่ 1	หาข้อมูลทำความเข้าใจเกี่ยวกับการใช้ YOLO สำหรับตรวจจับวัตถุ วางแผน และสรุปการทำงานโดยคร่าวของโปรเจคที่ทำ
	สัปดาห์ที่ 2	ติดตั้ง YOLO และเพิ่มการทำงานของ Open Image Dataset ในโปรแกรม VS Code และเขียนโปรแกรมการทำงาน จากนั้นทดลองการทำงานของโปรแกรม
	สัปดาห์ที่ 3	ฝึกฝนโมเดลให้ตรวจจับได้แม่นยำขึ้นในหลายๆรูปแบบ
	สัปดาห์ที่ 4	ระบบจะทำการจำแนกประเภท (class) ของวัตถุที่ตรวจจับได้ เช่น คน สัตว์ ยานพาหนะ และสิ่งกีดขวาง พร้อมทั้งกำหนดเงื่อนไขเบื้องต้นโดยจำกัดการตรวจจับเฉพาะวัตถุที่มีความเกี่ยวข้องหรือพบได้ในพื้นที่ทางเดินเท้า เพื่อให้ระบบแสดงข้อมูลที่มีความเหมาะสมและสอดคล้องกับบริบทการใช้งานของผู้พิการทางสายตาในพื้นที่ดังกล่าว

หมายเลขโครงงาน	

เดือนที่ 2 (ม.ค. 2568)	สัปดาห์ที่ 1	เพิ่มเงื่อนไขให้ครอบคลุมกับการตรวจจับในรูปแบบต่างๆ ให้มีความเหมาะสมและแม่นยำมากขึ้น
	สัปดาห์ที่ 2	
	สัปดาห์ที่ 3	ทำการรวมข้อมูลที่ได้จากการสร้างเงื่อนไขของวัตถุที่ตรวจจับ จากนั้นในอัลกอริทึมสร้าง Natural Language หรือ Scripts เพื่อเชื่อมประโยคให้เหมาะสม เข้าใจง่าย
	สัปดาห์ที่ 4	
เดือนที่ 3 (ก.พ. 2568)	สัปดาห์ที่ 1	แปลงข้อความที่ผ่านการเรียบเรียงแล้วมาเป็นเสียงบรรยาย (Text to speech Conversion) รวมถึงติดตั้ง Raspberry pi เพื่อเชื่อมต่อกับ VS Code
	สัปดาห์ที่ 2	ติดตั้งกล้องและลำโพง และทำการเชื่อมต่อกับ Raspberry pi เพื่อให้สามารถรับข้อมูลภาพและส่งเสียงบรรยายได้ และประกอบอุปกรณ์ให้สมบูรณ์
	สัปดาห์ที่ 3	
	สัปดาห์ที่ 4	ทดสอบการทำงานของระบบ โดยตรวจสอบการรับข้อมูลจากกล้องและการแปลงข้อมูลภาพเป็นข้อความ พร้อมทดสอบการพูดบรรยายของระบบผ่านลำโพง
เดือนที่ 4 (มี.ค. 2568)	สัปดาห์ที่ 1	ปรับปรุงระบบให้เหมาะสมกับสถานการณ์จริง และทดสอบความแม่นยำในการตรวจจับวัตถุ
	สัปดาห์ที่ 2	ตรวจสอบและปรับปรุงระบบการบรรยายเสียงในสภาพแวดล้อมต่างๆ เพื่อเพิ่มความแม่นยำในการทำงาน(พร้อมทำงาน)
	สัปดาห์ที่ 3	เขียนและจัดทำรายงานโครงงาน รวมถึงสรุปผลการทดสอบและการใช้งานจริง พร้อมนำเสนอผลลัพธ์ที่ได้จากการพัฒนา
	สัปดาห์ที่ 4	

หมายเหตุ

รายงานความก้าวหน้าที่จะมีกำหนดส่งของทุกๆเดือน ตามประกาศของภาควิชาฯ โดยในรายงานจะต้องแสดงหลักฐานผลการดำเนินงานสอดคล้องตามแผนการปฏิบัติงานที่ได้แสดงไว้

หมายเลขโครงงาน	

บทคัดย่อโครงงาน Pre-Project

ชื่อภาษาไทย ระบบช่วยบรรยายภาพด้วยเสียงสำหรับผู้พิการทางสายตา

ชื่อภาษาอังกฤษ Audio Description System for the Visually Impaired

บทคัดย่อ

โครงงานนี้มุ่งเน้นการพัฒนาระบบช่วยเหลือผู้พิการทางสายตาในการเดินทางบริเวณทางเดินเท้าริมถนน เพื่อเพิ่มความปลอดภัยและลดความเสี่ยงในการเผชิญอันตราย ระบบจะทำหน้าที่เปลี่ยนข้อมูลภาพให้เป็นเสียงบรรยายที่ชัดเจนและเข้าใจง่าย ช่วยส่งเสริมความสามารถในการรับรู้สิ่งแวดล้อมรอบตัวสำหรับผู้พิการทางสายตา ระบบนี้ได้รับการออกแบบโดยใช้เทคโนโลยี YOLO ซึ่งมีความสามารถในการตรวจจับและจำแนกวัตถุจากภาพถ่ายได้อย่างรวดเร็วและแม่นยำ ร่วมกับชุดข้อมูล Open Image Dataset ที่หลากหลายเพื่อการฝึกฝนโมเดลให้รองรับสถานการณ์จริงได้อย่างครอบคลุม การทำงานของระบบจะเริ่มจากการรับภาพจากกล้อง ซึ่งข้อมูลภาพดังกล่าวจะถูกประมวลผลและแปลงเป็นคำบรรยายผ่านเทคนิค Natural Language Processing (NLP) เพื่อให้เสียงบรรยายมีความเป็นธรรมชาติและเหมาะสมกับเหตุการณ์ จากนั้นคำบรรยายจะถูกเปลี่ยนเป็นเสียงผ่านกระบวนการ Text-to-Speech (TTS) และส่งผ่านลำโพงขนาดเล็กสำหรับการรับฟัง ระบบทั้งหมดถูกสร้างขึ้นบนแพลตฟอร์ม Raspberry pi ซึ่งเป็นไมโครคอนโทรลเลอร์ที่เชื่อมต่อกับกล้องและลำโพง การพัฒนาโปรแกรมควบคุมระบบดำเนินการผ่าน Visual Studio Code เพื่อความสะดวกและประสิทธิภาพในการพัฒนา เป้าหมายหลักของโครงงานนี้คือการช่วยเพิ่มความปลอดภัยให้กับผู้พิการทางสายตาขณะเดินทางบริเวณทางเดินเท้าริมถนน โดยลดความกังวลในการเผชิญสถานการณ์ที่ไม่คุ้นเคย พร้อมส่งเสริมให้ผู้พิการทางสายตามีคุณภาพชีวิตที่ดีขึ้นและสามารถดำเนินชีวิตในสังคมได้อย่างมั่นใจและปลอดภัยมากยิ่งขึ้น

หมายเลขโครงงาน	

Abstract

This project focuses on developing a system to assist visually impaired individuals in navigating sidewalks along roads safely, reducing risks and enhancing their awareness of surrounding environments. The system transforms visual data into clear and comprehensible audio descriptions, enabling visually impaired users to better perceive their surroundings. The system is designed using YOLO technology, which provides fast and accurate object detection and classification from images. It is integrated with the Open Image Dataset to train the model for comprehensive real-world scenarios. The process begins by capturing images through a camera, with the data being processed and converted into descriptive text using Natural Language Processing (NLP) techniques. These descriptions are then converted into speech through Text-to-Speech (TTS) technology and delivered via a small speaker for user accessibility. The entire system is implemented on a Raspberry pi platform, a microcontroller that connects efficiently with the camera and speaker. The program controlling the system is developed using Visual Studio Code (VS Code) for ease of coding and deployment. The primary goal of this project is to enhance the safety of visually impaired individuals when navigating sidewalks along roads. It aims to reduce concerns related to unfamiliar environments, promote independence, and improve the overall quality of life. This project showcases the potential of technology to create equality and improve living standards for everyone in society.