



รายงาน

เรื่อง ระบบตรวจจับอารมณ์บนใบหน้าเพื่อความปลอดภัยในการขับรถยนต์

Real-time Emotion Detector for driving

เสนอ

อาจารย์ชาคริต วัชรโรภาส

จัดทำโดย

นางสาวณัฐกฤตา	แก้วพัฒนสกุล	6410401043
นายธนดล	กฤตวีรนนท์	6410450974
นายกรณวิทย์	พัฒนโชควานิช	6410450079

รายงานนี้เป็นส่วนหนึ่งของรายวิชา 01418364 Practical Deep Learning

การเรียนรู้เชิงลึกในทางปฏิบัติ ภาคปลาย ปีการศึกษา 2566

คณะวิทยาศาสตร์ ภาควิชาวิทยาการคอมพิวเตอร์

มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตบางเขน

สารบัญ

คำนำ.....	3
บทที่ 1	4
บทนำ	4
ที่มาและความสำคัญ.....	4
ปัญหา	4
กลุ่มเป้าหมาย	4
ประโยชน์.....	4
บทที่ 2	5
การเก็บรวบรวมข้อมูล.....	5
ลักษณะของข้อมูลและแหล่งข้อมูล การได้มาของข้อมูล และการเก็บข้อมูลเองเพิ่มเติม.....	5
การแบ่งชุดข้อมูล	5
บทที่ 3	6
การเลือกประเภทและการพัฒนาโมเดล	6
บทที่ 4	8
การวิเคราะห์ผลลัพธ์	8
ข้อจำกัดและอุปสรรค.....	9
บรรณานุกรม	10

คำนำ

รายงานเรื่อง ระบบตรวจจับอารมณ์บนใบหน้าเพื่อความปลอดภัยในการขับรถยนต์ มีจุดมุ่งหมายในการใช้เทคโนโลยีและการวิเคราะห์ข้อมูลเพื่อป้องกันอุบัติเหตุในการเดินทางที่เกิดจากการปล่อยให้สภาวะทางอารมณ์ควบคุม โดยได้ศึกษาค้นคว้าจากงานวิจัยและ open source ในแหล่งข้อมูลออนไลน์ เพื่อเป็นหลักในการวิเคราะห์ข้อมูลประเภทอารมณ์ผ่านทางการเปลี่ยนแปลงใบหน้า

คณะผู้จัดทำรายงานขอขอบคุณอาจารย์ชาคริต วัชรโรภาสที่ให้คำปรึกษาและแนวคิดในการทำโครงงาน

หวังเป็นอย่างยิ่งว่ารายงานฉบับนี้จะเป็นประโยชน์แก่ผู้ขับขี่ที่ใช้รถยนต์ทุกประเภท และผู้ที่ต้องการศึกษาเกี่ยวกับระบบตรวจจับอารมณ์บนใบหน้าแบบเรียลไทม์

คณะผู้จัดทำ

บทที่ 1

บทนำ

ที่มาและความสำคัญ

อารมณ์เป็นภาวะความรู้สึกทางใจที่เปลี่ยนแปลงไปตามสิ่งกระตุ้นทั้งจากภายในและภายนอก สามารถแสดงออกได้ผ่านทั้งทางสีหน้า ท่าทาง ภาษากาย น้ำเสียง และคำพูด และมีทั้งประเภทที่สามารถควบคุมได้และไม่สามารถควบคุมได้

อารมณ์จึงเป็นหนึ่งในสาเหตุสำคัญของอุบัติเหตุจากการใช้รถยนต์ ที่เกิดจากการปล่อยให้อารมณ์ทางอารมณ์ควบคุมความคิด การกระทำ และการตัดสินใจ จนอาจนำไปสู่ความเสียหายทางชีวิตและทรัพย์สินได้

ปัญหา

อุบัติเหตุทางรถยนต์เป็นปัญหาที่เกิดขึ้นได้มากในประเทศไทยยุคปัจจุบัน โดยร้อยละ 60 เกิดขึ้นจากตัวผู้ขับขี่เอง โดยหนึ่งในสาเหตุสำคัญคือการปล่อยให้อารมณ์โกรธ หุนหัน โมโหร้าย ควบคุมสติและกระทำการโดยประมาท จนสร้างความเสียหายแก่ชีวิตและทรัพย์สิน รวมทั้งก่ออันตรายแก่ผู้ร่วมใช้ท้องถนนร่วมกัน

ผู้จัดทำจึงต้องการสร้างระบบที่สามารถตรวจจับอารมณ์ของผู้ขับขี่และทำการแจ้งเตือนเมื่ออารมณ์ของคนขับอยู่ในสภาวะเสี่ยงต่อการเกิดอุบัติเหตุ คือ อารมณ์โกรธ

กลุ่มเป้าหมาย

กลุ่มเป้าหมายของโครงการนี้ คือ ผู้ขับขี่ที่ใช้รถยนต์

ประโยชน์

1. ระบบที่สร้างสามารถตรวจสอบอารมณ์ของผู้ขับขี่รถยนต์และสามารถแจ้งเตือนได้หากเข้าสู่สภาวะที่อาจทำให้เกิดความเสี่ยงในการขับรถ
2. ผู้ใช้งานสามารถนำไปใช้งานได้จริง เพิ่มความปลอดภัยในการใช้รถใช้ถนนได้

บทที่ 2

การเก็บรวบรวมข้อมูล

ลักษณะของข้อมูลและแหล่งข้อมูล การได้มาของข้อมูล และการเก็บข้อมูลเองเพิ่มเติม

ลักษณะของโครงการเป็นปัญหาแบบ Classification ซึ่งรับข้อมูลเข้าเป็นรูปภาพใบหน้าที่แสดงอารมณ์ต่าง ๆ ของมนุษย์ โดยมีชุดข้อมูลรูปภาพที่ใช้จาก Kaggle.com จำนวนประมาณ 9,000 รูป และข้อมูลรูปภาพที่เก็บด้วยตนเองจำนวน 30 รูป คิดเป็น 0.3% ของข้อมูลเดิม ในการนำข้อมูลที่ได้ด้วยตนเองมาใช้ทางผู้จัดทำได้ทำการขออนุญาตจากบุคคลใกล้ชิดและใช้รูปจากบุคคลสาธารณะมาใช้ในการเทรนโมเดล

ข้อมูลภาพที่ใช้เทรนเป็นภาพมนุษย์เฉพาะส่วนที่เป็นใบหน้า ขนาด 48 x 48 pixels

การแบ่งชุดข้อมูล

ชุดข้อมูลเราได้แบ่งเป็นทั้งหมด 6 ประเภท โดยแบ่งเป็นดังต่อไปนี้

1. Angry	เป็นอารมณ์โกรธ	จำนวนทั้งหมด 1,505 รูป
2. Fear	เป็นอารมณ์ความหวาดกลัว	จำนวนทั้งหมด 1,505 รูป
3. Happy	เป็นอารมณ์ความสุข	จำนวนทั้งหมด 1,505 รูป
4. Neutral	เป็นสีหน้าตามธรรมชาติ	จำนวนทั้งหมด 1,505 รูป
5. Sad	เป็นอารมณ์ความเศร้า	จำนวนทั้งหมด 1,505 รูป
6. Surprise	เป็นความประหลาดใจหรือตกใจ	จำนวนทั้งหมด 1,505 รูป

บทที่ 3

การเลือกประเภทและการพัฒนาโมเดล

โครงงานเรื่องระบบตรวจจับอารมณ์บนใบหน้าเพื่อความปลอดภัยในการขับรถครั้งนี้ ใช้เทคนิค CNN (Convolutional Neural Network) ในการสร้างโมเดล ซึ่งเป็นโมเดลที่มีข้อมูลเข้าเป็นรูปภาพ เนื่องจากการตรวจจับใบหน้าของผู้ขับขี่รถยนต์จะถูกถ่ายจากกล้องที่ตั้งอยู่บริเวณหน้ารถ ข้อมูลเข้าที่ใช้จึงเป็นไฟล์วิดีโอซึ่งเป็น sequence ของรูปภาพ

ขั้นตอนในการสร้างโมเดล

ข้อมูลเข้า : dataset ที่บรรจุรูปใบหน้าของมนุษย์แต่ละอารมณ์ ขนาด 48 x 48 pixel

โครงสร้างโมเดล : CNN โดยใช้ activation function เป็น ReLU มีโครงสร้าง layer โมเดลดังนี้

```
model.add(Conv2D(32, kernel_size=(3, 3), activation='relu', input_shape=(48,48,3)))
model.add(Conv2D(64, kernel_size=(3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))

model.add(Conv2D(128, kernel_size=(3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Conv2D(128, kernel_size=(3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))

model.add(Flatten())
model.add(Dense(1024, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(7, activation='softmax'))
```

ในส่วนสุดท้ายของโมเดลจะได้ค่าข้อมูลออกเป็นค่าความน่าจะเป็นของอารมณ์ที่โมเดลคาดการณ์จากฟังก์ชัน Softmax

กระบวนการพัฒนาและเทรนโมเดล

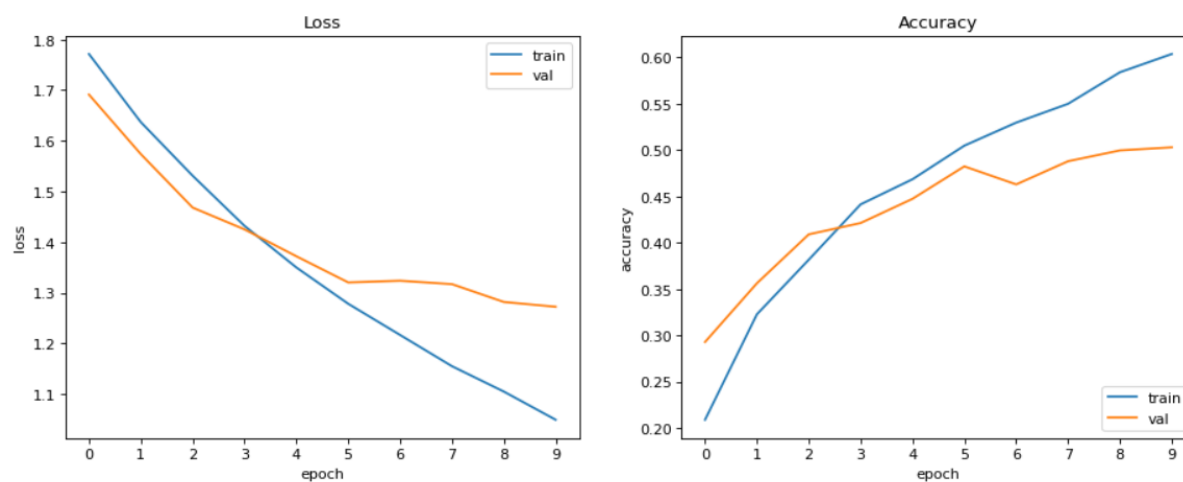
ในช่วงของการพัฒนาช่วงแรกนั้น จะเริ่มจากการที่ให้ model ทำนายประเภทอาชญากรรมใน dataset ว่ามีอาชญากรรมแบบใด และพัฒนา accuracy ซึ่งช่วงนี้ผู้จัดทำได้ปรับแต่งพารามิเตอร์ส่วนของ model เพิ่มลด filter หรือ Dense เข้าไป โดยจะประเมินได้จาก graph และค่า accuracy ขั้นตอนนี้ถูกทำซ้ำเพื่อให้ได้ค่า accuracy มากกว่าหรือเท่ากับ 70% จากการทำนายซึ่งเป็นเป้าหมายของคณะผู้จัดทำ

หลังจากที่ได้ model ที่ต้องการก็จะมาทำ detection อ่านใบหน้าคนจากกล้องถ่ายคลิปวิดีโอเรียลไทม์ โดยการ crop ส่วนที่เป็นใบหน้าจากวิดีโอและนำภาพที่ crop ได้เอาไว้ให้ model ทำนายว่าเป็นอาชญากรรมอะไรจากใน dataset แล้วแสดงผลออกมาข้างบนหัวของคนที่ถูก detect อยู่ตอนนั้น

บทที่ 4

การวิเคราะห์ผลลัพธ์

จากการ train และ test โมเดลจากชุดข้อมูลและโครงสร้างโมเดล ได้กราฟและค่า Loss, Accuracy ดังนี้



Train Loss: 0.865263, Accuracy: 0.699028
Test Loss: 1.272459, Accuracy: 0.502778

จะเห็นได้ว่ากราฟในช่วงแรกของผล Test นั้นมีค่าที่ค่อนข้างใกล้เคียงกับผล Train แต่หลังจากนั้นทั้งค่า Loss และ Accuracy ต่างออกห่างจากผล Train ทำให้เกิดภาวะ Overfitting

ข้อจำกัดและอุปสรรค

1. จำนวน dataset ที่ใช้ในการเทรนมีจำนวนมาก ต้องใช้เวลานาน จึงมีการปรับลดจำนวนข้อมูลลง แต่ก็ทำให้ได้ค่า accuracy ที่น้อยลงตามไปด้วย
2. ข้อมูลรูปภาพที่นำมาใช้ในการเทรนบางครั้งอาจมีปัญหาเนื่องจากความเพี้ยนบนใบหน้าหรือจากลักษณะทางกรรมพันธุ์ที่ทำให้การทำนายผลผิดพลาด เช่น ลักษณะตาชั้นเดียว
3. โปรแกรมอาจมีความคาดเคลื่อนในการ detect เนื่องจากใส่แว่นตาหรือทำทรงผมที่ปิดบังใบหน้า
4. ในช่วงแรกของการปรับปรุงโมเดล โปรแกรมสามารถตรวจจับใบหน้าและทำนายผลได้แค่ 1-2 อารมณ์ เนื่องจากขนาดของรูปภาพผิดไปจากโมเดลที่เทรนมา แต่ทางผู้จัดทำได้ทำการแก้ไขโค้ดในส่วนนี้แล้ว
5. ผู้จัดทำยังไม่สามารถปรับปรุงโมเดลให้ได้ค่า accuracy เท่าที่ต้องการได้
6. โปรแกรมยังไม่สามารถส่งเสียงแจ้งเตือนได้จริง

บรรณานุกรม

- Manas Sambare. (2020). *FER-2013*. Retrieved 7 February 2024, <https://www.kaggle.com/datasets/msambare/fer2013?resource=download>
- Atul Balaji. (2020). *Emotion-detection*. Retrieved 11 February 2024, <https://github.com/atulapra/Emotion-detection/tree/master>
- Travis Tang. (2019). *How Are You Feeling?—Answered with A Real-Time Emotion Detector*. Retrieved 11 February 2024, <https://medium.com/analytics-vidhya/building-a-real-time-emotion-detector-towards-machine-with-e-q-c20b17f89220>