

การพัฒนาแอปพลิเคชันระบุธนบัตรไทยด้วยเสียงสำหรับผู้พิการทางสายตา
ผ่านสมาร์ตโฟน

**Development Application for Identify Thai Banknote by Voice
for Blindness via Smartphone**

ณัฐวดี หงษ์บุญมี และ กาญจนา แสงตาล

Nattavadee Hongboonmee and Kanjana Sangtan

ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์

มหาวิทยาลัยนเรศวร

Department of Computer Science and Information Technology, Faculty of Science,

Naresuan University

Received: November 14, 2019; Revised: December 11, 2019; Accepted: December 17, 2019; Published: December 26, 2019

ABSTRACT – The aim of this research is to develop an application to help identify Thai banknotes by voice for the blindness on the Smartphone android operating system. This research applied deep learning technique for image recognition and text to speech for Smartphone pronunciation. The process began with a collection of 2,700 images of Thai banknotes by analyzing and creating an image classification model using a Convolution Neural Network, using MobileNet algorithm through Tensorflow library. Then, the model was developed to be a Smartphone application. The application is developed with Android studio, JAVA language and Text to Speech for pronunciation. The model performance test results obtained with an accuracy of 95.00%. The test of the application found that the banknote can be classified correctly 84.00%. The evaluation of user applications found that the average satisfaction was 4.33, the standard deviation 0.56, at a good level. It can conclude that this application makes it easier for the blindness to discover banknotes, because banknotes can be identified by voice and are convenient for the blindness to use anywhere, anytime. It also helps to promote a lifestyle for the blindness so that they can live their basic life more conveniently.

KEY WORDS -- Banknote Classification, Blindness, Convolution Neural Network, Deep Learning, Text to Speech

บทคัดย่อ -- งานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอการพัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยด้วยเสียงสำหรับผู้พิการทางสายตาผ่านสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ด้วยเทคโนโลยีการเรียนรู้เชิงลึกสำหรับการจำแนกภาพธนบัตรและเทคโนโลยี Text to Speech สำหรับการแปลงข้อความเป็นเสียงผ่านสมาร์ตโฟน การดำเนินงานเริ่มจากการรวบรวมกลุ่มตัวอย่างธนบัตรไทยจำนวน 2,700 ภาพ โดยการนำภาพมาวิเคราะห์และสร้างโมเดลจำแนกภาพด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันที่มีการใช้อัลกอริธึม MobileNet ผ่านไลบรารี Tensorflow จากนั้นนำโมเดลไปพัฒนาเป็นแอปพลิเคชันบน

สมาร์ตโฟน แอปพลิเคชันนี้พัฒนาด้วยโปรแกรม Android Studio ภาษา JAVA และไลบรารี Text to Speech ผลการวัดประสิทธิภาพโมเดลได้ค่าความถูกต้องเท่ากับ 95.00% ผลการทดสอบแอปพลิเคชันพบว่าสามารถจำแนกชนบัตร์ได้ถูกต้อง 84.00% ส่วนผลประเมินความพึงพอใจจากผู้ใช้งานพบว่ามีความพึงพอใจเฉลี่ยเท่ากับ 4.33 ส่วนเบี่ยงเบนมาตรฐาน 0.56 อยู่ในระดับดี สามารถสรุปได้ว่าแอปพลิเคชันนี้มีประสิทธิภาพสามารถช่วยให้ผู้พิการทางสายตาทราบถึงชนิดชนบัตร์ได้ง่ายขึ้น เนื่องจากสามารถระบุชนิดชนบัตร์ด้วยเสียงและอำนวยความสะดวกแก่ผู้พิการทางสายตาในการใช้งานได้ทุกที่ทุกเวลา ทั้งยังช่วยส่งเสริมการดำเนินชีวิตสำหรับผู้พิการทางสายตาให้สามารถดำรงชีวิตขั้นพื้นฐานได้สะดวกเพิ่มมากขึ้น

คำสำคัญ – การจำแนกชนบัตร์, ผู้พิการทางสายตา, โครงข่ายประสาทเทียมแบบคอนโวลูชัน, การเรียนรู้เชิงลึก, การแปลงข้อความเป็นเสียง

1. บทนำ

ข้อมูลของผู้พิการทางสายตาจากกรมส่งเสริมและพัฒนาคุณภาพชีวิตผู้พิการที่ได้รับการออกบัตรประจำตัวคนพิการ ในประเทศไทยปัจจุบัน พ.ศ.2562 มีจำนวน 186,081 คน [1] เป็นลำดับที่ 3 ของคนพิการทั้งหมด ซึ่งจากข้อมูลสถานการณ์ผู้พิการเห็นได้ชัดว่ามีผู้พิการเพิ่มขึ้นอย่างต่อเนื่องทุกปี ปัญหาที่ผู้พิการทางสายตาพบส่วนใหญ่เป็นปัญหาด้านการใช้ชีวิตประจำวัน โดยผู้พิการทางสายตาไม่สามารถใช้ชีวิตประจำวันได้เหมือนคนสายตาปกติจึงทำให้เกิดข้อจำกัดต่างๆ มากมาย เช่น ปัญหาเรื่องการเลือกซื้อสินค้าในชีวิตประจำวัน การเลือกซื้ออาหารเสื้อผ้าและผลิตภัณฑ์ต่างๆ ซึ่งจำเป็นที่จะต้องใช้จ่ายด้วยธนบัตร โดยที่ธนบัตรไทยในปัจจุบันยังไม่ออกแบบมาให้รองรับกับผู้พิการทางสายตา ทำให้ผู้พิการจำเป็นต้องอาศัยความคุ้นเคยหรือสอบถามจากผู้คนรอบข้างให้ช่วยบอกราคาธนบัตร ส่งผลให้ผู้พิการทางสายตาประสบปัญหาการถูกหลอกลวงจากมิฉาชีพได้ง่าย

จากที่กล่าวมาข้างต้นทำให้เห็นถึงความสำคัญของการใช้ธนบัตรในชีวิตประจำวันของผู้พิการทางสายตา ดังนั้นคณะผู้วิจัยจึงมีแนวคิดที่จะพัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยด้วยเสียงสำหรับผู้พิการทางสายตาดบนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ โดยมีวัตถุประสงค์เพื่อช่วยเหลือผู้พิการทางสายตาให้สามารถช่วยเหลือตนเองได้ โดยใช้หลักการประมวลผลภาพด้วยเทคโนโลยีการเรียนรู้เชิงลึกระบุค่าเงินในธนบัตร ผู้พิการทางสายตาสามารถใช้อุปกรณ์สมาร์ตโฟน

ถ่ายภาพธนบัตรเพื่อระบุราคาธนบัตรด้วยเสียงได้ ซึ่งจะช่วยให้ผู้พิการทางสายตาสามารถดำรงชีวิตขั้นพื้นฐานได้สะดวกมากขึ้น

2. ทฤษฎีที่เกี่ยวข้อง

2.1 ผู้พิการทางสายตา

การมองเห็นเป็นความสามารถของมนุษย์ที่จะสร้างภาพขึ้นในจิตใจเป็นกระบวนการชิมชาภาพที่มีอยู่ผ่านดวงตาไปสู่ระบบประสาททำให้เกิดความรู้สึกในการมองเห็น ซึ่งความสามารถในการมองเห็นสามารถแบ่งได้เป็น 2 ประเภท [2] ได้แก่ 1) การตาบอดเลือนกลาง (Low vision) หมายถึง การที่มีความบกพร่องทางการเห็นสามารถมองเห็นได้บ้างแต่ไม่เท่ากับคนปกติ ไม่สามารถเรียนรู้ได้ด้วยวิธีการเรียนการสอนที่ใช้กับคนปกติได้ ต้องใช้เครื่องมือหรืออุปกรณ์พิเศษบางอย่างที่จะช่วยให้สามารถใช้สายตาได้ดีขึ้น 2) การตาบอดสนิท (Blindness) หมายถึง การมองอะไรเห็นได้น้อยมากหรือไม่สามารถใช้สายตาให้เป็นประโยชน์ได้

ตารางที่ 1. แสดงระดับความพิการทางสายตาที่กำหนดโดยองค์การอนามัยโลก [2]

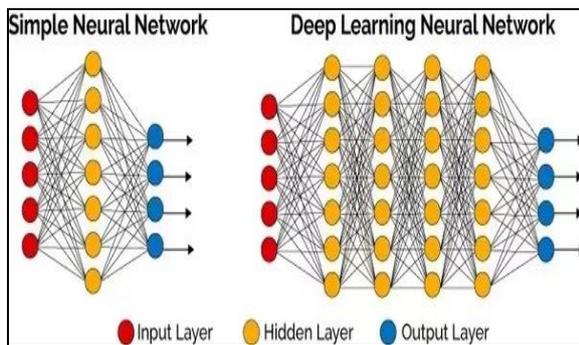
ระดับความพิการทางสายตา		ระดับความสมบูรณ์ของการมองเห็น
สายตาเลือนกลาง (Low Vision)	ระดับ 1	สามารถมองเห็นได้ 100.0%
		สามารถมองเห็นได้ 87.5%
		สามารถมองเห็นได้ 64.0%
		สามารถมองเห็นได้ 48.9%
	ระดับ 2	สามารถมองเห็นได้ 20.0%

สายตาบอด (Blindness)	ระดับ 3	สามารถมองเห็นได้ 10.0%
	ระดับ 4	สามารถมองเห็นได้ 2.0%
	ระดับ 5	มองไม่เห็นแม้แสงสว่าง

2.2 การเรียนรู้เชิงลึก

การเรียนรู้เชิงลึก (Deep Learning) [3] เป็นเทคนิคหนึ่งทางด้านการเรียนรู้ของเครื่อง (Machine Learning) ที่ถูกพัฒนาขึ้นมาอยู่บนพื้นฐานของโครงข่ายประสาทเทียมแบบเชิงลึก (Deep Learning Neural Network) (รูปที่ 1) โครงข่ายประสาทเทียมแบบเชิงลึกประกอบด้วยเส้นซึ่งจะเชื่อมต่อไปเป็นชั้นๆ แบ่งออกเป็น 3 Layers ได้แก่ 1) Input Layer ชั้นที่เป็นส่วนของการรับข้อมูล 2) Hidden Layer ชั้นที่เป็นเหมือนระบบเซลล์ประสาทในสมองและ 3) Output Layer ชั้นที่เป็นส่วนตอบสิ่งที่ประมวลผลออกมา

พื้นฐานของการเรียนรู้เชิงลึก คือ อัลกอริทึมที่พยายามจะสร้างแบบจำลองเพื่อแทนความหมายของข้อมูลในระดับสูง โดยการสร้างสถาปัตยกรรมข้อมูลที่ประกอบไปด้วยโครงสร้างย่อยๆ หลายส่วนและแต่ละส่วนนั้นได้มาจากการแปลงที่ไม่เป็นเชิงเส้น การเรียนรู้เชิงลึกถือว่าเป็นวิธีการที่มีศักยภาพสูงในการจัดการกับพีเจอร์สำหรับการเรียนรู้แบบไม่มีผู้สอนหรือการเรียนรู้แบบกึ่งมีผู้สอน



รูปที่ 1. แสดงโครงข่ายประสาทเทียมแบบเชิงลึก [4]

เทคนิคของการเรียนรู้เชิงลึกได้ถูกนำไปใช้แก้ไขปัญหาในหลายๆ ด้าน โดยเฉพาะอย่างยิ่งด้านที่เกี่ยวข้องกับปัญญาประดิษฐ์ (Artificial Intelligence) และการประมวลผลรูปภาพ (Image Processing) ซึ่งรวมทั้งการจำแนกวัตถุในรูปภาพด้วย

2.3 โครงข่ายประสาทเทียมแบบคอนโวลูชัน

โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network: CNN) [5] เป็นรูปแบบหนึ่งของการเรียนรู้ของเครื่อง (Machine Learning) ที่จำลองรูปแบบประเภทโครงข่ายประสาทเทียม (Artificial Neural Networks: ANN) ให้เป็นรูปแบบการเรียนรู้เชิงลึก (Deep Learning) ที่เป็นลักษณะโครงข่ายประสาทเทียมเชิงลึก (Deep Learning Neural Network) โดยมีวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำรูปแบบ (Pattern Recognition) และการสร้างความรู้ใหม่ (Knowledge Extraction) เช่นเดียวกับความสามารถที่มีในสมองมนุษย์ โครงข่ายประสาทเทียมเชิงลึกมีการเรียนรู้วิธีการแทนข้อมูลอย่างมีประสิทธิภาพ

2.4 อัลกอริทึม MobileNet

MobileNet เป็นอัลกอริทึมภายใต้โครงข่ายประสาทเทียมแบบ CNN [6] ที่ออกแบบให้มีขนาดเล็กสำหรับอุปกรณ์พกพาเช่นสมาร์ทโฟนโดยบริษัทกูเกิล ซึ่งเป็นอัลกอริทึมที่มีความแม่นยำพอประมาณ แต่การใช้ทรัพยากรน้อยกว่าสถาปัตยกรรมแบบอื่นจึงสามารถนำไปใช้งานกับอุปกรณ์พกพาที่มีทรัพยากรจำกัดได้ดี

2.5 TensorFlow

TensorFlow เป็นไลบรารีสำหรับใช้พัฒนา Machine learning ของกูเกิล [6] สามารถนำมาประยุกต์ใช้งานได้อย่างยืดหยุ่นสามารถนำไปใช้งานได้แบบ Portable มีการใช้งานง่ายและเป็น Open Source เขียนด้วยภาษา Python ซึ่งเป็นโปรแกรมจำลองการทำงานของกระบวนการเรียนรู้ของเครื่อง TensorFlow และสามารถรองรับการทำงานบนทั้งบน Desktop และ Mobile

2.6 Text to Speech

Text to Speech ซอฟต์แวร์สังเคราะห์เสียง [7] เป็นเทคโนโลยีที่สามารถสร้างเสียงคำพูดได้ตามความต้องการ ปัจจุบันสามารถประยุกต์ใช้กับข้อความภาษาไทยได้ ความสามารถของ Text To Speech คือการส่งข้อความให้ระบบแปลงข้อความเป็นเสียงซึ่งถือว่าเป็นพื้นฐานบนอุปกรณ์พกพาแบบปฏิบัติการแอนดรอยด์

2.7 งานวิจัยที่เกี่ยวข้อง

จากการศึกษาพบว่ามิงงานวิจัยที่พัฒนาโดยใช้เทคโนโลยีการเรียนรู้เชิงลึกอยู่จำนวนหนึ่งเกี่ยวกับการศึกษาดังต่อไปนี้

ปองคุณและคณะ [8] ศึกษาเรื่องการใช้การเรียนรู้แบบเชิงลึกสำหรับการรู้จำตัวอักษรพิมพ์ไทย งานวิจัยนี้นำเสนอโครงข่ายประสาทเทียมแบบ CNN ซึ่งเป็นการเรียนรู้ของเครื่องประเภทหนึ่งสำหรับใช้ในการรู้จำตัวอักษรพิมพ์ไทยที่ประกอบด้วย 44 ตัวอักษร และรูปแบบอื่นๆ ที่ไม่ใช่ตัวอักษร 21 รูปแบบ TensorFlow ถูกนำมาประยุกต์ใช้สำหรับการเรียนรู้คุณลักษณะ ผลการทดลองแสดงให้เห็นว่าโครงข่ายประสาทเทียมแบบ CNN ประสบความสำเร็จที่สามารถจำแนกตัวอักษรพิมพ์ไทยได้ถูกต้อง 99.84% โดยข้อจำกัดคือสามารถจำแนกได้ถูกต้องกับขนาดภาพตัวอักษร 28x28 pixels

พระระดิษฐ์และคณะ [9] ได้ศึกษาวิจัยเรื่องการตรวจจับและจำแนกยานพาหนะด้วยอัลกอริทึมการเรียนรู้เชิงลึกสำหรับระบบวิดีโอรักษาความปลอดภัย งานวิจัยนี้ได้นำเทคนิค Inception และ MobileNet มาทำการทดลองเปรียบเทียบการจำแนกรถจักรยานยนต์ด้วยข้อมูลภาพและวิดีโอของยานพาหนะและใช้เทคนิคการประมวลผลรูปภาพ Single Shot Multibox Detector (SSD) ในการวิเคราะห์ขอบเขตของรูปภาพ ผลการวิจัยพบว่าเทคนิค Inception มีค่าความถูกต้องที่ใกล้เคียงกับเทคนิค MobileNet แต่จากการทดลองกับวิดีโอพบว่ายังไม่สามารถจำแนกภาพรถจักรยานยนต์ได้ทั้งหมด โดยเฉพาะที่เป็นภาพระยะไกล

อรรรถพลและธนภัทร [10] ได้ศึกษาเรื่องการทำนายพระเครื่อง โดยใช้เทคนิคการเรียนรู้เชิงลึก โดยใช้โครงข่ายประสาทเทียมแบบ CNN ประเภท Xception ในการสกัดคุณลักษณะเด่นของภาพและสร้างโมเดลที่มีการใช้จำนวนพารามิเตอร์เท่าที่จำเป็น ในขณะที่ยังมีประสิทธิภาพในการทำนายผลที่ดี จากการทดลองพบว่าโมเดลที่นำเสนอให้ค่าความแม่นยำเท่ากับ 96.90%

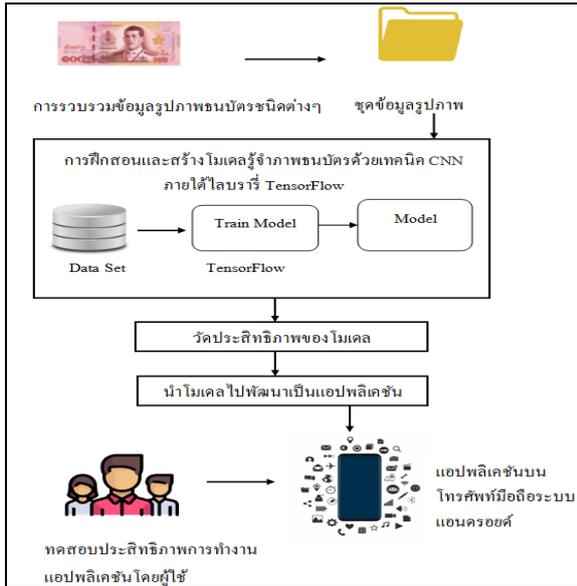
นพรัตน์และคณะ [11] นำเสนอการออกแบบและพัฒนาระบบแบ่งปันสูตรการทำอาหารและค้นหาสูตรการทำอาหารจากภาพวัตถุดิบด้วยเทคนิคการเรียนรู้เชิงลึกโดยใช้

งานได้บนอุปกรณ์พกพา งานวิจัยนี้ประยุกต์ใช้การเรียนรู้เชิงลึกที่มีการใช้อัลกอริทึม YOLO ผ่านไลบรารี Darknet สำหรับการสร้างโมเดลในการรู้จำภาพ โมเดลได้ฝึกสอนให้สามารถรู้จำวัตถุดิบได้จำนวน 20 ประเภท โดยสอนด้วยภาพวัตถุดิบประเภทละ 100 ภาพ ผลการทดลองพบว่าโมเดลมีค่า Precision, Recall และ F1-score อยู่ที่ 0.96, 0.98 และ 0.97 ตามลำดับ

จากการศึกษางานวิจัยที่เกี่ยวข้องได้ข้อสรุปว่ามีงานวิจัยหลายชิ้นที่มีประสิทธิภาพดีในการจำแนกภาพถ่ายจากสภาพแวดล้อมที่คงที่ ซึ่งส่วนใหญ่ใช้เทคนิคการประมวลผลภาพหลายรูปแบบเพื่อสกัดคุณลักษณะต่างๆ (Features) และสร้างโมเดลจำแนกประเภทด้วยเทคนิคการเรียนรู้เชิงลึกภายใต้โครงข่ายประสาทเทียมแบบ CNN ซึ่งมีความสามารถในการประมวลผลภาพและรู้จำภาพที่มีผลลัพท์การทำนายผลที่แม่นยำและมีประสิทธิภาพสูง ดังนั้นงานวิจัยนี้จึงนำเทคนิคการเรียนรู้เชิงลึกโดยใช้โครงข่ายประสาทเทียมแบบ CNN ซึ่งสามารถเรียนรู้ชุดคุณลักษณะเฉพาะของรูปภาพโดยอัตโนมัติ ภายใต้อัลกอริทึม MobileNet มาประยุกต์ใช้สำหรับการจำแนกและรู้จำภาพธนบัตรแบบอัตโนมัติผ่านแอปพลิเคชันบนสมาร์ตโฟนสำหรับผู้พิการทางสายตาต่อไป

3. วิธีการดำเนินงาน

การพัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยด้วยเสียงสำหรับผู้พิการทางสายตาผ่านสมาร์ตโฟนสามารถสร้างกรอบแนวคิดในการวิจัยแสดงดังรูปที่ 2



รูปที่ 2. แสดงกรอบแนวคิดของงานวิจัย

จากรูปที่ 2 กรอบแนวคิดของงานวิจัยประกอบด้วย ขั้นตอนการดำเนินการศึกษาวิจัย 5 ขั้นตอน ได้แก่ 1) การรวบรวมและเตรียมข้อมูล 2) การสร้างโมเดลรู้จำภาพธนบัตร 3) การวัดประสิทธิภาพโมเดล 4) การพัฒนาแอปพลิเคชันและ 5) การทดสอบประสิทธิภาพแอปพลิเคชัน

3.1 การรวบรวมและเตรียมข้อมูล

การเก็บรวบรวมข้อมูลรูปภาพกลุ่มตัวอย่างได้มาจากการถ่ายภาพด้วยตนเองและรวบรวมจากเครือข่ายอินเทอร์เน็ต โดยทำการคัดเลือกธนบัตรที่นิยมใช้กันในปัจจุบันจำนวน 4 รูปแบบ ได้แก่ ธนบัตรแบบที่ 17 แบบที่ 16 แบบที่ 15 และแบบที่ระลึก แบ่งเป็นธนบัตรประเภท 20 บาท ด้านหน้าและด้านหลัง 50 บาท ด้านหน้าและด้านหลัง 100 บาท ด้านหน้าและด้านหลัง 500 บาท ด้านหน้าและด้านหลัง 1,000 บาท ด้านหน้าและด้านหลัง สามารถแบ่งรูปภาพได้เป็น 27 คลาส กลุ่มตัวอย่างคลาสละ 100 ภาพ รวมรูปภาพทั้งหมด 2,700 ภาพ การเก็บข้อมูลแต่ละคลาสแยกเป็น โฟลเดอร์ได้ทั้งหมด 27 โฟลเดอร์ ชุดข้อมูลภาพที่รวบรวมได้จะแบ่งเป็น 2 ส่วน ส่วนแรกคือข้อมูลฝึกสอน 80% ส่วนที่สอง 20% คือข้อมูลทดสอบ รายละเอียดดังรูปที่ 3-6 และตารางที่ 2



รูปที่ 3. แสดงภาพตัวอย่างธนบัตรแบบที่ 17



รูปที่ 4. แสดงภาพตัวอย่างธนบัตรแบบที่ 16



รูปที่ 5. แสดงภาพตัวอย่างธนบัตรแบบที่ 15





รูปที่ 6. แสดงภาพตัวอย่างธนบัตรที่ระลึก

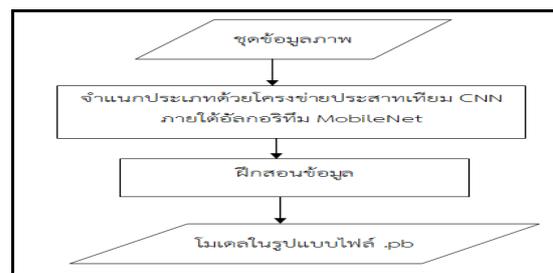
ธนบัตรแบบที่ระลึกด้านหน้าจะเหมือนกับธนบัตรแบบที่ 16 ด้านหน้า ดังนั้นจึงใช้เป็นคลาสเดียวกัน

ตารางที่ 2. แสดงชุดข้อมูลภาพที่ใช้ในการศึกษา

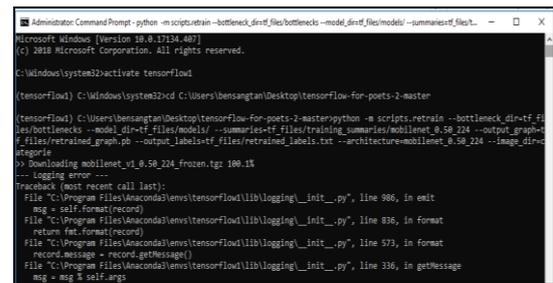
ลำดับ	แบบธนบัตร	ประเภท	ด้าน	จำนวน/ภาพ
1	แบบที่ 17	20	หน้า	100
2		หลัง	100	
3		50	หน้า	100
4		หลัง	100	
5		100	หน้า	100
6		หลัง	100	
7		500	หน้า	100
8		หลัง	100	
9		1,000	หน้า	100
10		หลัง	100	
11	แบบที่ 16	20	หน้า	100
12		หลัง	100	
13		50	หน้า	100
14		หลัง	100	
15		100	หน้า	100
16		หลัง	100	
17		500	หน้า	100
18		หลัง	100	
19		1,000	หน้า	100
20		หลัง	100	
21	แบบที่ 15	1,000	หน้า	100
22		หลัง	100	
23	แบบที่ระลึก	20	หลัง	100
24		50	หลัง	100
25		100	หลัง	100
26		500	หลัง	100
27		1,000	หลัง	100
รวม				2,700

3.2 การสร้างโมเดล

ขั้นตอนนี้เป็นกรนำข้อมูลที่จัดเตรียมไว้มาทำการประมวลผลเพื่อฝึกสอน (Train data) และสร้าง โมเดลเพื่อรู้จำและจำแนกภาพธนบัตรประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกเลือกใช้เทคนิคโครงข่ายประสาทเทียมแบบ CNN ภายใต้อัลกอริทึม MobileNet เครื่องมือสำหรับการดำเนินงาน ได้แก่ ไลบรารี TensorFlow ซึ่งเป็นไลบรารีสำหรับพัฒนาโมเดลและภาษา Python สำหรับฝึกสอนและสร้าง โมเดล ดำเนินการฝึกสอนจำนวน 500 รอบ ขั้นตอนการทำงานแสดงในรูปที่ 7 – 8



รูปที่ 7. แสดงขั้นตอนการฝึกสอนข้อมูลและสร้าง โมเดล



รูปที่ 8. แสดงตัวอย่างให้ผลการฝึกสอนข้อมูล

3.3 การวัดประสิทธิภาพโมเดล

การวัดประสิทธิภาพโมเดลจำแนกภาพธนบัตรโดยดำเนินการแบ่งข้อมูลออกเป็น 2 ส่วน ส่วนที่หนึ่ง 80% สำหรับการฝึกสอนและสร้าง โมเดล ส่วนที่สอง 20% สำหรับเป็นข้อมูลทดสอบ ซึ่งจะได้ข้อมูลจำนวน 540 ภาพ (คลาสละ 20 ภาพ) มาทำการทดสอบโมเดลและนำ โมเดลจำแนกภาพที่สร้างเสร็จเรียบร้อยแล้วทดสอบค่าความถูกต้องของการทำนาย การวัดประสิทธิภาพ โมเดลใช้ตัวชี้วัด ได้แก่ ค่าความถูกต้อง (Accuracy) ดังสมการที่ 1 [12]

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN}} \quad (1)$$

(TP+TN+FP+FN)

โดยที่

True Positive (TP) คือ คลาสบอกว่าจริงและทำนายว่าจริง

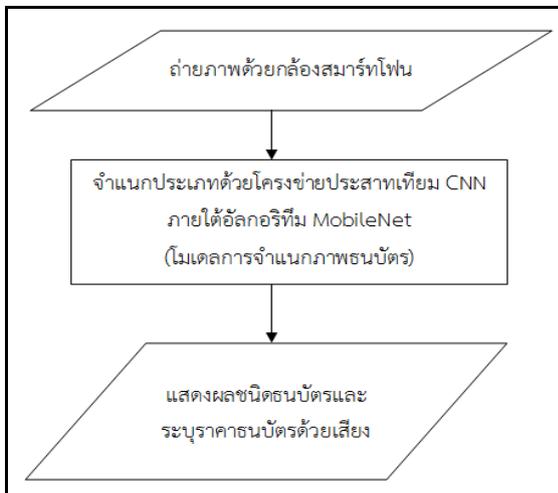
True Negative (TN) คือ คลาสบอกว่าไม่จริงและทำนายว่าไม่จริง

False Positive (FP) คือ คลาสบอกว่าจริงและทำนายว่าไม่จริง

False Negative (FN) คือ คลาสบอกว่าไม่จริงและทำนายว่าจริง

3.4 การพัฒนาแอปพลิเคชัน

การพัฒนาแอปพลิเคชันใช้โปรแกรม Android Studio และภาษา JAVA เป็นภาษาหลักในการพัฒนา ส่วนการระบุชนิดด้วยเสียงใช้เทคนิค Text to Speech ในการแปลงข้อความเป็นเสียง แสดงขั้นตอนการทำงานของแอปพลิเคชัน ดังรูปที่ 9



รูปที่ 9. แสดงขั้นตอนการทำงานของแอปพลิเคชัน

แอปพลิเคชันระบุชนิดธนบัตรไทยด้วยเสียงบนระบบปฏิบัติการแอนดรอยด์ ออกแบบขั้นตอนการทำงาน รูปที่ 9 ให้สามารถถ่ายภาพธนบัตรจากกล้องสมาร์ทโฟน จากนั้นทำการรู้จำและจำแนกภาพด้วยโมเดลที่ผ่านการฝึกสอนเรียบร้อยแล้วและทำการแสดงผลด้วยการระบุราคาธนบัตรด้วยเสียงของภาพธนบัตรที่ตรวจพบ ซึ่งมีการประมวลผลในแบบเรียลไทม์



รูปที่ 10. แสดงไฟล์โมเดลที่นำเข้าไปรรม Android studio

การนำโมเดลจำแนกภาพที่ผ่านการฝึกสอนเรียบร้อยแล้วแล้วเข้าไปรรม Android Studio เพื่อเรียกใช้ประมวลผลในแอปพลิเคชัน ดำเนินการดังตัวอย่าง รูปที่ 10 ซึ่งไฟล์ที่ได้จากการฝึกสอนจะมีจำนวนสองไฟล์ คือไฟล์ 'retrained_graph.pb' และไฟล์ 'retrained_labels.txt'

ส่วนการเรียกใช้งาน Text to Speech เพื่อแปลงข้อความเป็นเสียง ขั้นตอนนี้ดำเนินการเขียนโค้ดเรียกใช้งานการทำงานของเสียงเพื่อแปลงข้อความออกเป็นเสียง ตามจำนวนราคาธนบัตรไทย ดังตัวอย่างโค้ดในรูปที่ 11

```
private TextToSpeech tts;
tts=new TextToSpeech(this, this);

@Override
public synchronized void onDestroy() {
    super.onDestroy();
    tts.shutdown();
}

private void speak(CharSequence message) {
    if (Build.VERSION.SDK_INT >= Build.VERSION_CODES.LOLLIPOP) {
        tts.speak(message, TextToSpeech.QUEUE_FLUSH, params: null, utteranceid: "");
    } else {
        tts.speak(message.toString(), TextToSpeech.QUEUE_FLUSH, params: null);
    }
}

@Override
public void onInit(int status) {
```

รูปที่ 11. แสดงโค้ดเรียกใช้การแปลงข้อความเป็นเสียง

3.5 การทดสอบแอปพลิเคชัน

การทดสอบประสิทธิภาพของแอปพลิเคชันใช้วิธีการทดสอบแบบ Black Box Testing เป็นการทดสอบฟังก์ชันการทำงานต่างๆ ของแอปพลิเคชัน เพื่อทดสอบความสมบูรณ์ของแอปพลิเคชันโดยผู้ใช้งาน

4. ผลการศึกษา

ผลการทดสอบการจำแนกภาพธนบัตร โดยใช้เทคนิค CNN และผลการประเมินประสิทธิภาพแอปพลิเคชันของการจำแนกภาพธนบัตรด้วยการหาค่าความถูกต้อง แสดงรายละเอียดดังต่อไปนี้

4.1 ผลการทดสอบประสิทธิภาพโมเดล

ผลของการสร้างโมเดลและทดสอบประสิทธิภาพโมเดลจำแนกภาพธนบัตรด้วยเทคโนโลยีการเรียนรู้เชิงลึก ได้ผลค่าความถูกต้อง (Accuracy) แสดงรายละเอียดในตารางที่ 3

ตารางที่ 3. แสดงผลการทดสอบประสิทธิภาพโมเดล

ประเภท	แบบธนบัตร	ด้าน	Accuracy (%)
20	แบบที่ 17	หน้า	99.66%
		หลัง	99.72%
50		หน้า	88.87%
		หลัง	87.60%
100		หน้า	98.91%
		หลัง	97.56%
500		หน้า	99.65%
		หลัง	97.09%
1,000		หน้า	98.70%
		หลัง	96.11%
20	แบบที่ 16	หน้า	87.02%
		หลัง	78.77%
50		หน้า	99.16%
		หลัง	78.60%
100		หน้า	96.65%
		หลัง	97.56%
500		หน้า	99.99%
		หลัง	99.29%
1,000		หน้า	100.00%
		หลัง	89.70%
1,000	แบบที่ 15	หน้า	92.00%
		หลัง	93.10%
20	แบบที่ ละเอียด	หลัง	99.99%
50		หลัง	99.59%
100		หลัง	85.92%
500		หลัง	99.99%

1,000	หลัง	98.00%
ค่าเฉลี่ย		95.00%

จากตารางที่ 3 แสดงค่าความถูกต้องในการเปรียบเทียบการจำแนกภาพธนบัตรแต่ละชนิดจำนวน 27 คลาส พบว่าการจำแนกภาพธนบัตรชนิด 1,000 บาท ด้านหน้าแบบที่ 16 มีความถูกต้องสูงกว่าคลาสอื่น คือ 100.00% ส่วนการจำแนกภาพธนบัตรชนิด 50 บาท ด้านหลังแบบที่ 17 มีความถูกต้องต่ำกว่าคลาสอื่น คือ 78.60% เนื่องจากในชุดข้อมูลของคลาสธนบัตรชนิด 50 บาท ด้านหลังแบบที่ 17 ภาพธนบัตรมีความคล้ายคลึงใกล้เคียงกับธนบัตรชนิด 20 บาทและชนิด 100 บาทด้านหลัง (ดังรูปที่ 12) ซึ่งทำให้การทำนายผลมีค่าความถูกต้องต่ำกว่าคลาสอื่น



รูปที่ 12. แสดงภาพที่อยู่ในคลาสธนบัตร 20 50 บาท 100 บาท แบบที่ 17 ด้านหลัง

สามารถสรุปผลของประสิทธิภาพโมเดลได้ดังนี้ โมเดลซึ่งสร้างจากไลบรารี TensorFlow อัลกอริทึม MobileNet มีผลลัพธ์ค่าความถูกต้องเฉลี่ย 95.00% สามารถนำมาใช้พัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยสำหรับผู้พิการทางสายตาบนสมาร์ตโฟนได้

4.2 ผลการพัฒนาแอปพลิเคชัน

การพัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยสำหรับผู้พิการทางสายตาด้วยเสียงผ่านสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ เมื่อเปิดแอปพลิเคชันเข้ามาหน้าแรก ดังแสดงใน

รูปที่ 13 ผู้ใช้งานจะต้องใช้กล้องสมาร์ทโฟนถ่ายภาพธนบัตรแล้วแอปพลิเคชันจะเรียกใช้โมเดลการเรียนรู้เชิงลึกเพื่อจำแนกภาพธนบัตรและแสดงผลลัพธ์หน้าจอระบุชนิดราคาธนบัตรพร้อมเสียงบอกราคาสำหรับผู้พิการทางสายตากลับไปให้ผู้ใช้งานดังแสดงในรูปที่ 13-14 โดยแอปพลิเคชันจะแสดงว่าเป็นธนบัตรชนิด 20 บาท 50 บาท 100 บาท 500 บาท หรือ 1,000 บาท โดยไม่ระบุว่าเป็นธนบัตรรุ่นใด (แบบ 17 แบบ 16 แบบ 15 หรือแบบที่ระลึก) แอปพลิเคชันสามารถประมวลผลได้ทั้งแบบด้านหน้าและแบบกลับหัว เพื่ออำนวยความสะดวกให้ผู้พิการทางสายตาสามารถถ่ายภาพได้ทุกมุม



รูปที่ 13. ตัวอย่างหน้าจอแสดงผลแอปพลิเคชัน



รูปที่ 14. แอปพลิเคชันสามารถถ่ายภาพธนบัตรแบบกลับหัวได้

ผลการพัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยด้วยเสียง สำหรับผู้พิการทางสายตาผ่านสมาร์ทโฟน ระบบปฏิบัติการแอนดรอยด์ด้วยเทคนิคการเรียนรู้เชิงลึก จากผลการพัฒนาพบว่าแอปพลิเคชันสามารถระบุราคาของธนบัตรแต่ละภาพได้ทั้งแบบหน้าตรงและแบบกลับหัว พร้อมทั้งมีเสียง

บอกราคาธนบัตรเพื่ออำนวยความสะดวกให้ผู้พิการทางสายตาได้

ผลการทดสอบประสิทธิภาพการจำแนกภาพธนบัตร โดยการทดสอบถ่ายภาพธนบัตรผ่านกล้องสมาร์ทโฟนจำนวน ชนิดธนบัตรละ 10 ครั้ง คละแบบธนบัตร พบว่าระบบที่พัฒนาขึ้นสามารถใช้ในการระบุธนบัตรได้ โดยให้ค่าความถูกต้องเฉลี่ย 84.00% รายละเอียดดังตารางที่ 4 ธนบัตรที่แอปพลิเคชันสามารถทำนายได้ถูกต้องมากที่สุดได้แก่ ธนบัตรชนิด 1,000 บาท ส่วนธนบัตรที่ทำนายได้ถูกต้องน้อยที่สุด ได้แก่ ธนบัตรชนิด 50 บาท

ตารางที่ 4. แสดงผลการทดสอบประสิทธิภาพแอปพลิเคชัน

ธนบัตร	จำนวนความถูกต้อง	จำนวนความผิดพลาด	Accuracy (%)
20	8	2	80.00%
50	7	3	70.00%
100	8	2	80.00%
500	9	1	90.00%
1,000	10	-	100.00%
ค่าเฉลี่ย			84.00%

ข้อจำกัดที่พบจากการทดสอบการใช้งานจริงแอปพลิเคชัน คือ 1) การประมวลผลภาพควรถ่ายภาพธนบัตรในที่ที่มีแสงสว่างเพียงพอ 2) ระยะห่างระหว่างกล้องสมาร์ทโฟนกับธนบัตรไม่ควรห่างกันมากกว่า 8 นิ้ว ถ้ามักกว่าระยะที่กำหนดจะมีผลต่อความแม่นยำในการทำนายได้

4.3 ผลการทดสอบแอปพลิเคชัน

ผลการทดสอบแอปพลิเคชันใช้วิธีการทดสอบแบบแบ็กบ็อกซ์ (Black Box Testing) เพื่อเป็นตัวบ่งชี้คุณภาพของระบบที่พัฒนาขึ้น การทดสอบระบบประกอบด้วย การประเมินประสิทธิภาพของระบบ การประเมินความพึงพอใจของผู้ใช้ที่มีผลต่อระบบโดยผู้ใช้งานทั่วไปจำนวน 30 คน โดยการหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน โดยกำหนดเกณฑ์ของค่าเฉลี่ยการประเมิน แสดงในตารางที่ 5

ตารางที่ 5. แสดงเกณฑ์การให้คะแนน

ระดับเกณฑ์	ระดับการประเมิน
4.51 – 5.00	ดีมาก
3.51 – 4.50	ดี
2.51 – 3.50	พอใช้
1.51 – 2.50	น้อย
1.00 – 1.50	น้อยที่สุด

ตารางที่ 6. แสดงผลการประเมินแอปพลิเคชันจากผู้ใช้งาน

รายการประเมิน	Mean	S.D.	ระดับ
1. ด้านความสามารถระบบตามหน้าที่การทำงาน	4.27	0.57	ดี
2. ด้านความสามารถระบบตามความต้องการของผู้ใช้	4.37	0.71	ดี
3. ด้านความสะดวกและความง่ายต่อการใช้งาน	4.33	0.60	ดี
4. ด้านความสามารถในการทำงานของระบบ	4.13	0.43	ดี
5. ด้านประโยชน์ของระบบ	4.57	0.50	ดีมาก
สรุป	4.33	0.56	ดี

ผลการประเมินความพึงพอใจของผู้ใช้ที่มีต่อการใช้งานแอปพลิเคชัน โดยผู้ใช้งานทั่วไปจำนวน 30 คน พบว่าได้ค่าเฉลี่ยเท่ากับ 4.33 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.56 ซึ่งผลที่ได้อยู่ในระดับดี แสดงว่าแอปพลิเคชันนี้มีประสิทธิภาพสามารถอำนวยความสะดวกให้ผู้พิการทางสายตาให้สามารถดำรงชีวิตประจำวันด้วยตนเองได้สะดวกเพิ่มมากขึ้น

5. บทสรุปและอภิปรายผล

การศึกษาวิจัยนี้เป็นการวิจัยเชิงประยุกต์มีวัตถุประสงค์เพื่อพัฒนาแอปพลิเคชันที่ช่วยผู้พิการทางสายตาให้สามารถทราบราคาธนบัตรแต่ละชนิดด้วยเสียงผ่านการถ่ายภาพจากกล้องสมาร์ทโฟน โดยประยุกต์ใช้เทคโนโลยีการเรียนรู้เชิงลึก

การดำเนินงานเริ่มจากการรวบรวมข้อมูลกลุ่มตัวอย่างภาพธนบัตร โดยแบ่งคลาสที่ต้องการจำแนกออกเป็น 27 คลาส คลาสละ 100 ภาพ รวมจำนวนภาพกลุ่มตัวอย่างทั้งหมด 2,700

ภาพ นำมาสร้างโมเดลจำแนกภาพธนบัตรด้วยไลบรารี Tensorflow และภาษา Python ทำการแบ่งชุดข้อมูลสำหรับฝึกสอน (Train Data Set) 80% และชุดข้อมูลทดสอบ (Test Data Set) 20% รอบการเทรนจำนวน 500 รอบ ผลการสร้างโมเดลจำแนกภาพ พบว่าโมเดลมีค่าความถูกต้องเฉลี่ยเท่ากับ 95.00% ซึ่งมีประสิทธิภาพอยู่ในระดับดี

จากนั้นนำโมเดลไปพัฒนาเป็นแอปพลิเคชันบนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ โดยใช้โปรแกรม Android Studio ภาษา JAVA และไลบรารี Text to Speech สำหรับการแปลงข้อความเป็นเสียง ผลการทดสอบการใช้งานจริงของแอปพลิเคชัน พบว่าแอปพลิเคชันสามารถจำแนกและระบุชนิดธนบัตรได้ถูกต้องเฉลี่ย 84.00% ธนบัตรที่สามารถจำแนกได้ถูกต้องมากที่สุด ได้แก่ ธนบัตรชนิด 1,000 บาท ส่วนธนบัตรที่จำแนกได้ถูกต้องน้อยที่สุด ได้แก่ ธนบัตรชนิด 50 บาท เนื่องจากในชุดข้อมูลของคลาสธนบัตรชนิด 50 บาท ภาพธนบัตรมีความคล้ายคลึงใกล้เคียงกับธนบัตรชนิด 20 บาทและชนิด 100 บาทด้านหลัง ซึ่งอาจทำให้การทำงานมีค่าความถูกต้องน้อยกว่าคลาสนั้น

ส่วนผลการทดสอบความพึงพอใจแอปพลิเคชันจากผู้ใช้งานจำนวน 30 คน พบว่าได้ค่าเฉลี่ยรวมเท่ากับ 4.33 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.56 สามารถสรุปได้ว่าแอปพลิเคชันที่พัฒนาขึ้นนี้มีประสิทธิภาพอยู่ในระดับดี สามารถจำแนกและระบุธนบัตรด้วยเสียงได้ ช่วยอำนวยความสะดวกแก่ผู้พิการทางสายตาในการใช้งานได้ทุกที่ทุกเวลา ทั้งยังช่วยส่งเสริมการดำเนินชีวิตสำหรับผู้พิการทางสายตาได้อีกทางหนึ่ง

แนวทางการพัฒนาต่อไปในอนาคต ได้แก่ การปรับปรุงกระบวนการเตรียมข้อมูลเพิ่มการทำ Validation set ของข้อมูล เพื่อให้การฝึกสอนมีประสิทธิภาพเพิ่มมากขึ้นและศึกษาอัลกอริทึมอื่นๆ ของเทคนิคการเรียนรู้เชิงลึกเพิ่มเติมเพื่อปรับปรุงประสิทธิภาพการจำแนกให้มีประสิทธิภาพดีขึ้น

เอกสารอ้างอิง

- [1] กรมส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการ, “สถานการณ์คนพิการ มีนาคม 2562,” [Online]. Available: <http://dep.go.th/Content/View/4232/1> [Accessed: April 10, 2019].

- [2] สุภา คงแสงไชย, กลยุทธ์การกระตุ้นพัฒนาการทางสายตาสำหรับเด็กพิการทางสายตาและเด็กด้อยโอกาสอายุ 0-3 ปี, กรุงเทพฯ: สำนักพิมพ์สุภา, 2014.
- [3] L. Deng and D. Yu, "Deep Learning Methods and Application," *Foundations and Trends in Signal Processing*, vol. 17, No.3-4, pp. 197–387, June. 2014.
- [4] "The Connect Between Deep Learning and AI". [Online]. Available: <https://opensourceforu.com/2018/01/connect-deep-learning-ai/>. [Accessed: April 14, 2019].
- [5] Y. Gao and H. J. Lee, "Vehicle Make Recognition Based on Convolutional Neural Network," 2015 2nd International Conference on Information Science and Security (ICISS), Seoul, 2015, pp. 1–4.
- [6] S. Rabano, M. Cabatuan, E. Sybingco, E. Dadios and E. Calilung, "Common Garbage Classification Using MobileNet," 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Baguio City, Philippines, 2018, pp. 1–4.
- [7] S. Lukose and S. Upadhya, "Text to Speech Synthesizer-Formant Synthesis," in Proceeding of IEEE International Conference on Nascent Technologies in the Engineering Field, Navi Mumbai, India, 2017, pp. 121–124.
- [8] P. Pankerd, T. Tirabutr, T. Tanduang and S. Pongpinigpinyo, "Using deep learning for printed thai charaters recognition," in Proceeding of 13th National Conference on Computing and Information Technology (NCCIT), King Mongkut's University of Technology North Bangkok, Bangkok, 2017, pp. 181–186.
- [9] W. Wairotchanaphuttha, M. Uimin, N. Boonsirisumpun and W. Puarungroj, "Detection and classification of vehicles using deep learning algorithm for video surveillance systems," in Proceeding of 14th National Conference on Computing and Information Technology (NCCIT), King Mongkut's University of Technology North Bangkok, Bangkok, 2018, pp. 402–407.
- [10] A. Reangsuk and T. Kungkajit, "Classification of amulets using deep learning techniques," in Proceeding of 10th National Conference on Information Technology (NCIT), Mahasarakham University, Mahasarakham, 2018, pp. 190–194.
- [11] N. Manoi, A. Bunjanda and C. Rattanapoka, "A System for cooking recipe sharing and cooking recipe finding by an image of ingredients using deep learning technique," *The Journal of Industrial Technology*, vol. 15, No.2, pp. 97–111, May-August. 2019.
- [12] S. Gupta, W. Zhang, and F. Wang, "Model Accuracy and Runtime Tradeoff in Distributed Deep Learning: A Systematic Study," in Proceeding IEEE 16th International Conference on Data Mining (ICDM), Barcelona, Spain, 2016, pp. 171–180.