

บทที่ 4

ผลการดำเนินงาน

สำหรับผลการดำเนินงานการพัฒนากระบวนกรู้จำทำทางภาษามือไทยด้วยโครงข่ายประสาทเทียมแบบวนกลับ แบ่งออกได้เป็น 3 ส่วน ได้แก่

- 4.1 ผลการเตรียมข้อมูล
- 4.2 ผลการฝึกฝนโมเดล
- 4.3 ผลการวัดประสิทธิภาพโมเดล
- 4.4 ผลการเปรียบเทียบประสิทธิภาพโมเดล

4.1 ผลการเตรียมข้อมูล

4.1.1 การหาข้อมูลภาษามือไทย

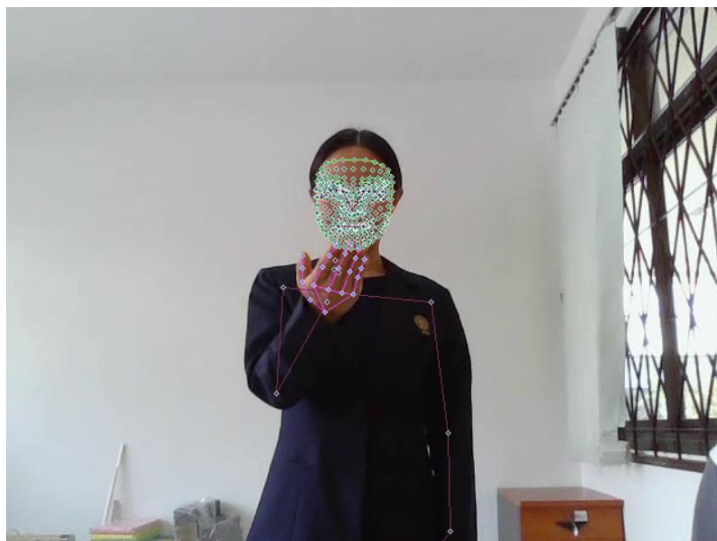
ในการรวบรวมข้อมูล สำหรับการสร้าง TSL10 (dataset ภาษามือไทย 10 ท่า) ผู้วิจัยต้องการวิดีโอท่าภาษามือที่ใช้ในชีวิตประจำวันของผู้พิการทางการได้ยินและการสื่อความหมาย เป็นจำนวน 10 คำ ผู้วิจัยได้มีการออกหนังสือขอความอนุเคราะห์จากศูนย์บริการสนับสนุนการนักศึกษาพิการระดับอุดมศึกษา (DSS) ประจำมหาวิทยาลัยราชภัฏสกลนครเพื่อเก็บข้อมูลสำหรับการเทรนโมเดลสำหรับการรู้จำภาษามือไทยจากทั้งผู้เชี่ยวชาญภาษามือและผู้พิการที่ใช้ภาษามือเป็นหลักในการสื่อสาร



ภาพที่ 4.1 ตัวอย่างภาษามือไทย ‘สวัสดี’ จากผู้เชี่ยวชาญภาษามือไทย

4.1.2 การสกัดลักษณะเด่นของข้อมูล

ภาษามือนั้นขึ้นอยู่กับการใช้มือและท่าทาง การนำวิดีโอที่เป็นภาษามือมาใช้ในการเทรนโมเดลนั้นจึงเป็นเรื่องยาก ผู้วิจัยจึงได้ใช้เครื่องมือ MediaPipe ที่เป็น Framework มาใช้ในการแก้ปัญหา ซึ่งวิธีการคือการใช้ MediaPipe ในการสร้าง Key Points ขึ้นตามจุดต่าง ๆ ของร่างกายเป็นค่า มิติ X, Y, Z ของหน้า, มือและท่าทางรูปภาพที่ 4.2



ภาพที่ 4.2 การใช้ MediaPipe ในการจำลอง Key Points

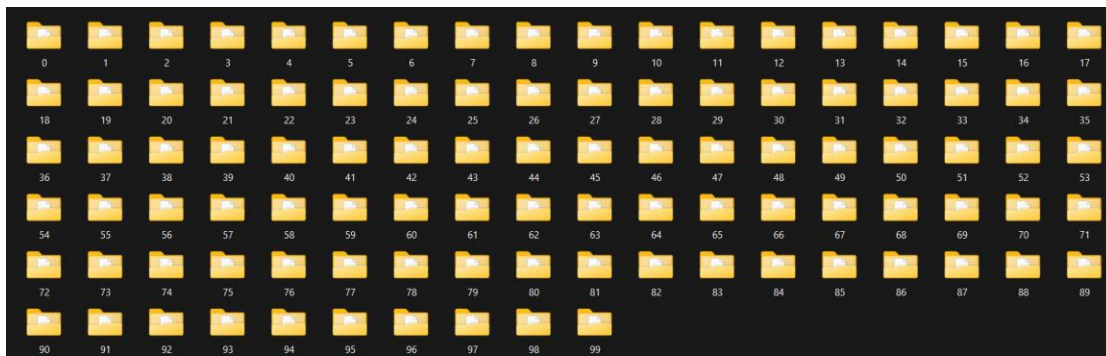
4.1.3 การเตรียมไฟล์สำหรับเทรนโมเดล

เมื่อสามารถสร้าง Key Points เสร็จขั้นตอนต่อไปคือการนำผลของค่า Key Points ของแต่ละจุดของร่างกายเขียนเป็น .npy ไฟล์ ซึ่งมีขั้นตอนดังนี้

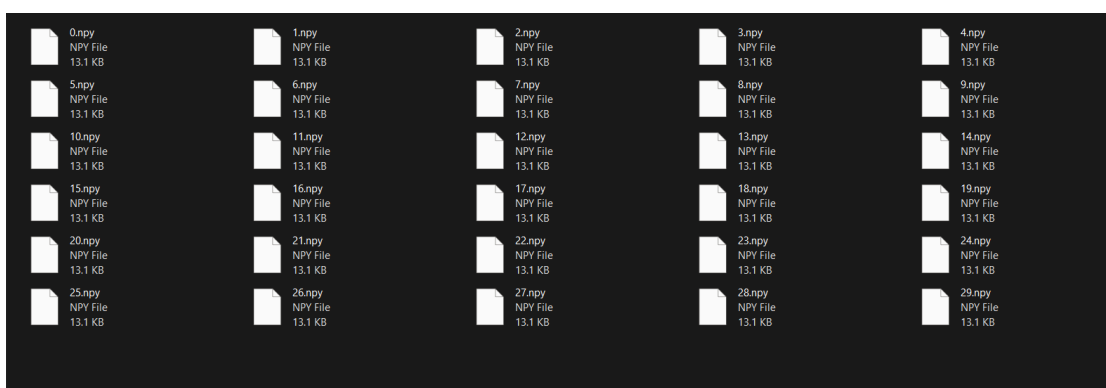
- 1) สร้างโฟลเดอร์สำหรับเก็บ Datasets
- 2) ในโฟลเดอร์ Datasets มี โฟลเดอร์ที่เป็นชื่อท่าภาษามือ ดังภาพที่ 4.3
- 3) ในโฟลเดอร์ที่เป็นชื่อท่าภาษามือจะมีโฟลเดอร์สำหรับเก็บวิดีโอท่าภาษามือ 85 วิดีโอ โดยแยกเป็น โฟลเดอร์ละ 1 วิดีโอ ดังภาพที่ 4.4
- 4) ในโฟลเดอร์เก็บวิดีโอท่าภาษามือจะมีไฟล์ .npy 30 ไฟล์ ซึ่ง 1 ไฟล์ จะเก็บค่าที่ได้จากการสกัด Key Points จาก Mediapipe X, Y, Z ใน 1 เฟรม ดังภาพที่ 4.5



ภาพที่ 4.3 โฟลเดอร์ชื่อท่าภาษามือ



ภาพที่ 4.4 โฟลเดอร์ 100 โฟลเดอร์สำหรับเก็บ .npy ไฟล์



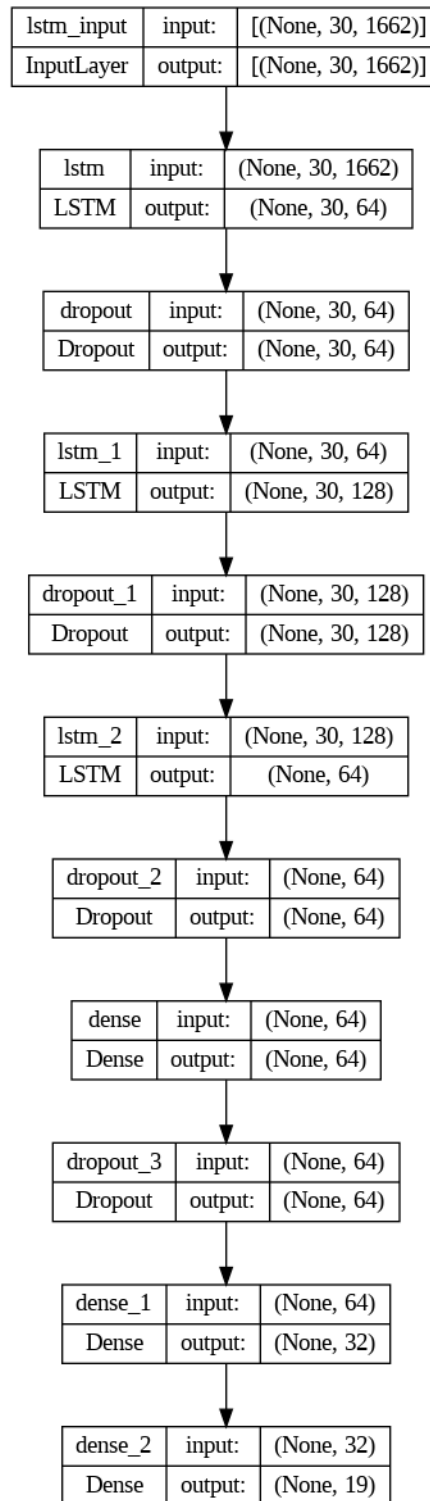
ภาพที่ 4.5 ไฟล์ .npy 30 ไฟล์ ใน 1 โฟลเดอร์วิดีโอ

```
[0.513763964176178, 0.2923258967631226, -0.7177266478538513, 0.9998338222503662, 0.5323629975318909, 0.23734638644321442, -0.6862671983149495, 0.9996179342269897, 0.5444337725539343, 0.23533668775558472, -0.68823530864715535, 0.9996473735493774, 0.556738078942078, 0.2371148109436835, -0.6864155723022461, 0.999553047688293, 0.4890727480779724, 0.24118662049770355, -0.6979305744171143, 0.9997079668452454, 0.47264736800792847, 0.24212203919887543, -0.6979529857635498, 0.9997654552228888, 0.4351204037666321, 0.2620096802711487, -0.4803236424922943, 0.9998923540115356, 0.5376855134963989, 0.34098589420318604, -0.6192570924758911, 0.9999128580093384, 0.48499661684036255, 0.3454931974411011, -0.630450725554199, 0.9999533891677856, 0.6755043268203735, 0.5413625836372375, -0.25669506192207336, 0.9995226860046387, 0.323110116797504395, 0.54801037940079, -0.39301383029060164, 0.9998277425765991, 0.6903421878814697, 0.4728837966918945, -0.16575846876011658, 0.6145241856575012, 0.280550818186569214, 0.8829171657562256, -0.3184884786609835, 0.7724840044975281, 0.6689577698707581, 1.1150187253952026, -0.37631329894065857, 0.06317772716283798, 0.254950881004335, 1.121402052227783, -0.5437986254692078, 0.2010273934106445, 0.6775882244110107, 1.202189326286316, -0.442077815326843, 0.06265565752983093, 0.23309989273548126, 1.1978400898220825, -0.6125809550285339, 0.16351358592510223, 0.6443684697151184, 1.1880269050598145, -0.44865655303001404, 0.0994851216673851, 0.25811303119659424, 1.1903650760650635, -0.6809645295143127, 0.2326664924621582, 0.63535761831319092, 1.1503536701202393, -0.4037608584295349, 0.10839639604091644, 0.27482539415359497, 1.1509642763137817, -0.5702068967819214, 0.2458571344614029, 0.5886954665184021, 1.100771966934204, 0.015387484803795815, 0.018652206380520484, 0.37841296195983807, 1.1060657501220703, -0.012911488302052021, 0.019229311496019363, 0.5669893622398376, 1.5478923320770264, 0.12271221721732797, 0.004864117596298456, 0.3690384328365326, 1.5520542860031128, 0.0941530391573906, 0.0018569874810054898, 0.5498352646827698, 1.9391964673995972, 0.6198827624320984, 0.00025886562070809305, 0.37422287464141846, 1.9356013536453247, 0.44789570569992065, 0.00007588174776174128, 0.5515167117118835, 2.0101349353790283, 0.6517901420593262, 0.0002654777199495584, 0.36904314160346985, 2.002102851867676, 0.47243815660476685, 0.00015105475904420018, 0.5283734798431396, 2.0731194019317627, 0.3326743543148041, 0.0002841882815118879, 0.4112731814384404, 2.0751466751098633, 0.0801382857950333, 0.000139289651973874, 0.5055079460144043, 0.324820302295685, -0.020756686106324196, 0.5043082237243652, 0.2893693447111037, -0.0376918762222287, 0.5046404004069085, 0.2908507453346224, 0.0196088829421997, 0.4069516604545746, 0.2551915943622589, -0.0236737838742723, 0.5039717555046082, 0.27899354696273804, -0.040126752108335495, 0.5035245418548584, 0.26505526900291443, -0.03733150288462639, 0.5024560689926147, 0.23065197467803955, -0.01832530837807655, 0.4502556324005127, 0.22837071120738983, 0.008382469415664673, 0.5018429160118103, 0.20549513399600983, -0.01401455039381981, 0.5015809535980225, 0.19139957427978516, -0.015624466352164745, 0.49968209862709045, 0.13673017919063568, -0.007687304634600878, 0.5056626796722412, 0.3276265263557434, -0.019792525097727776, 0.5057495832443237, 0.33172017335891724, -0.01718379184603691, 0.5058169960975647, 0.3354008197784424, -0.013908008113683456, 0.5061157631937371063, -0.01224687043574326, 0.5062152313232422, 0.3413931727409363, -0.01326871756464243, 0.50651937772315979, 0.34717822074800137, -0.014530476182699203, 0.506720244804491, 0.3528299927711487, -0.01316643113156271, 0.5068864822387695, 0.36117708683013916, -0.005900360178202391, 0.5044519901275635, 0.29446184635162354, -0.03405220806598663, 0.497595459227936, 0.2939923405647278, -0.024411851540207863, 0.42047008872032166, 0.1858855557441711, 0.04129388928413391, 0.4712504744529724, 0.23682130873203278, 0.00151264475993812, 0.46425706148147583, 0.2379876673221588, 0.0017867061542347074, 0.457370787858963, 0.23788440227508545, 0.00359665558974737, 0.4478225284286499, 0.23204606771469116, 0.00929336342960596, 0.4770136773586273, 0.23396410648007965, 0.002425867598503828, 0.4606583523750305, 0.21163778655910492, -0.001598997856490314, 0.4682215750217430, 0.21184209455413818, -0.0013061054050922394, 0.45353642508628045, 0.21344813704490662, 0.000375406472012401, 0.448919415473938, 0.21680809798526764, 0.0033341306261718273, 0.44226643441107605, 0.2391018122434616, 0.012676827609539032, 0.48109912872314453, 0.3766014575958252, 0.0014231280656531453, 0.4474794864654541, 0.2256166934967041, 0.010544038377782236, 0.4200647175312042, 0.2320365011692047, 0.044032201170921326, 0.43351414799690247, 0.23166775703430176, 0.019634438678622246, 0.4682251214981079, 0.2810328006744385, -0.005045172292739153, 0.4970826208591461, 0.32158970832824707, -0.0198683962225914, 0.4990517199039459, 0.3323628902453303,
```

ภาพที่ 4.6 ไฟล์ .npy ที่เก็บค่า X, Y, Z ของ Key Points

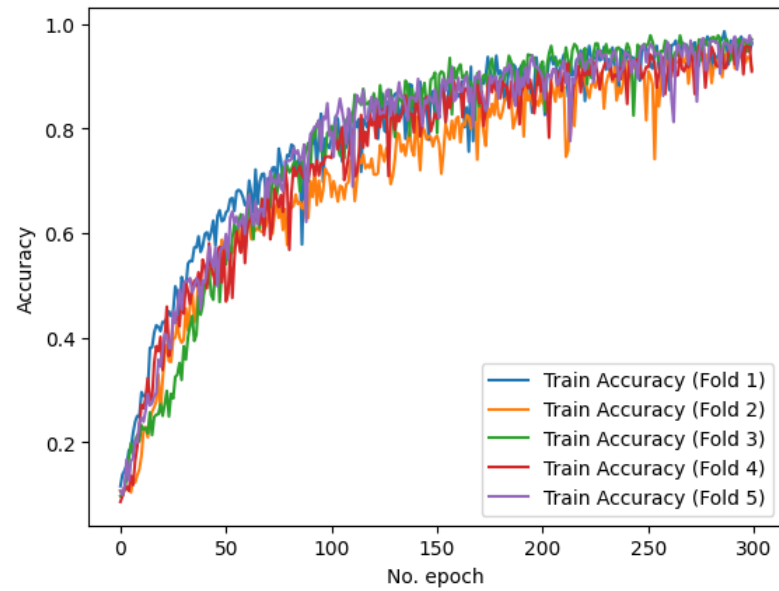
4.2 ผลการฝึกฝนโมเดล

4.2.1 โมเดล LSTM

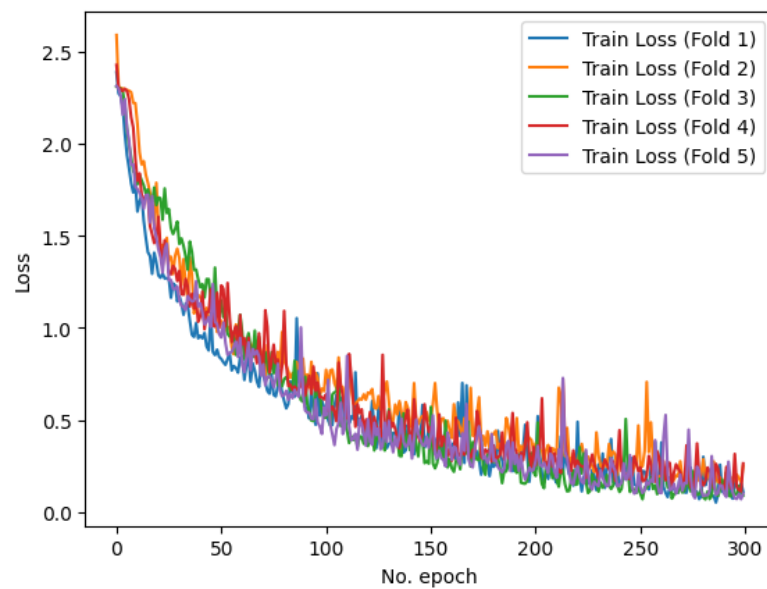


ภาพที่ 4.7 โครงสร้างโมเดล LSTM

ผลจากการพัฒนาระบบรู้จำท่าทางภาษาไทยด้วยโครงข่ายประสาทเทียมแบบวนกลับพบว่า การกำหนดค่าอัตราการเรียนรู้ Epochs เท่ากับ 300 และ Batch Size เท่ากับ 32 ของโมเดล LSTM ให้ค่า Accuracy และค่า Loss ดังภาพด้านล่าง

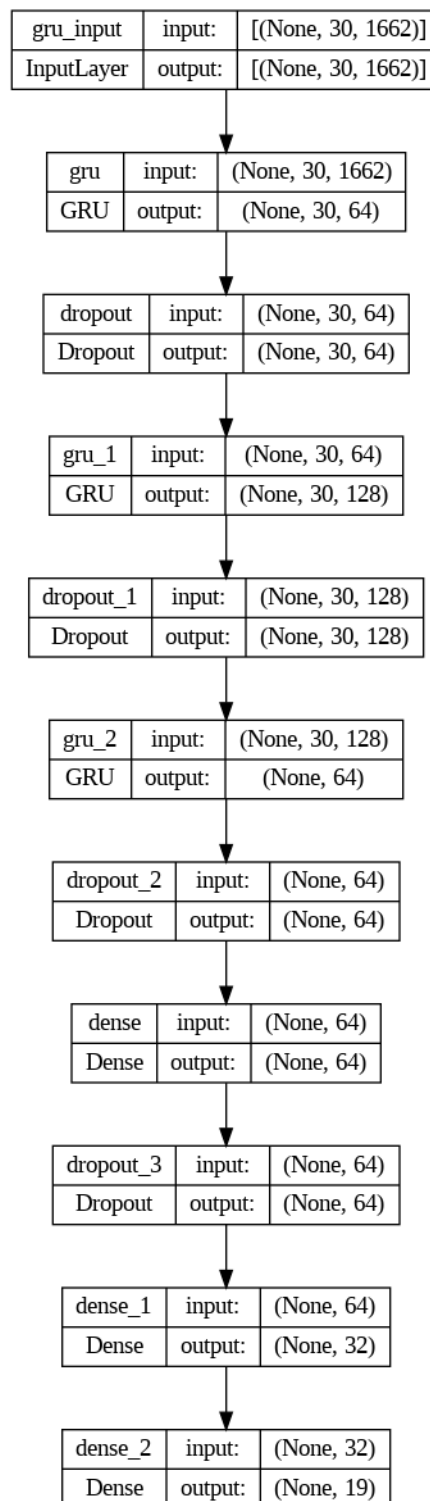


ภาพที่ 4.8 กราฟแสดงจำนวนรอบการเทรนและค่า Accuracy ของโมเดล LSTM



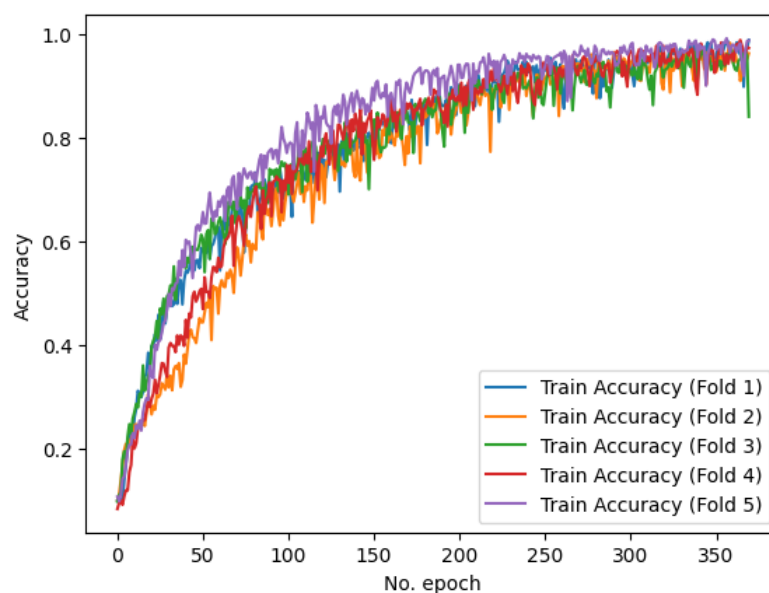
ภาพที่ 4.9 กราฟแสดงรอบการเทรนและค่า Loss ของโมเดล LSTM

4.2.2 โมเดล GRU

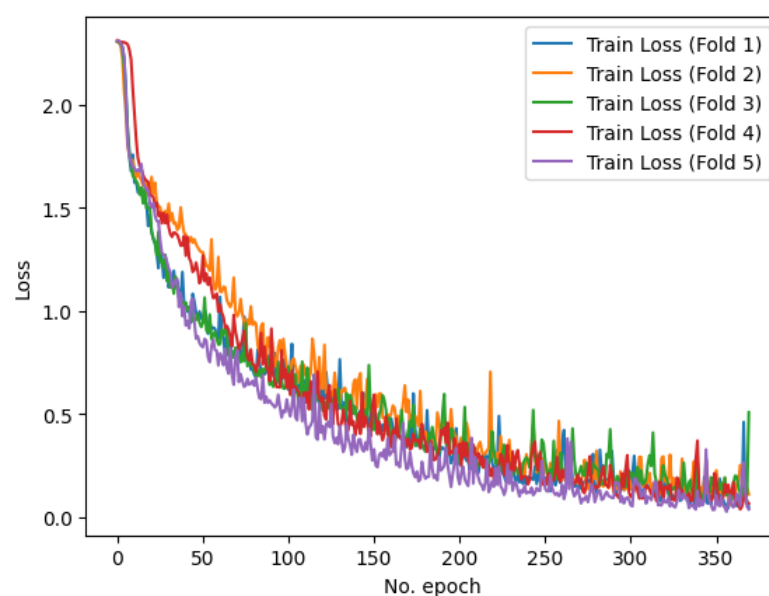


ภาพที่ 4.10 โครงสร้างโมเดล GRU

ผลจากการพัฒนาระบบรู้จำท่าทางภาษามือไทยด้วย RNN พบว่า การกำหนดค่าอัตราการเรียนรู้ Epochs เท่ากับ 370 และ Batch Size เท่ากับ 32 ของโมเดล GRU ให้ค่า Accuracy และค่า Loss ดังภาพด้านล่าง

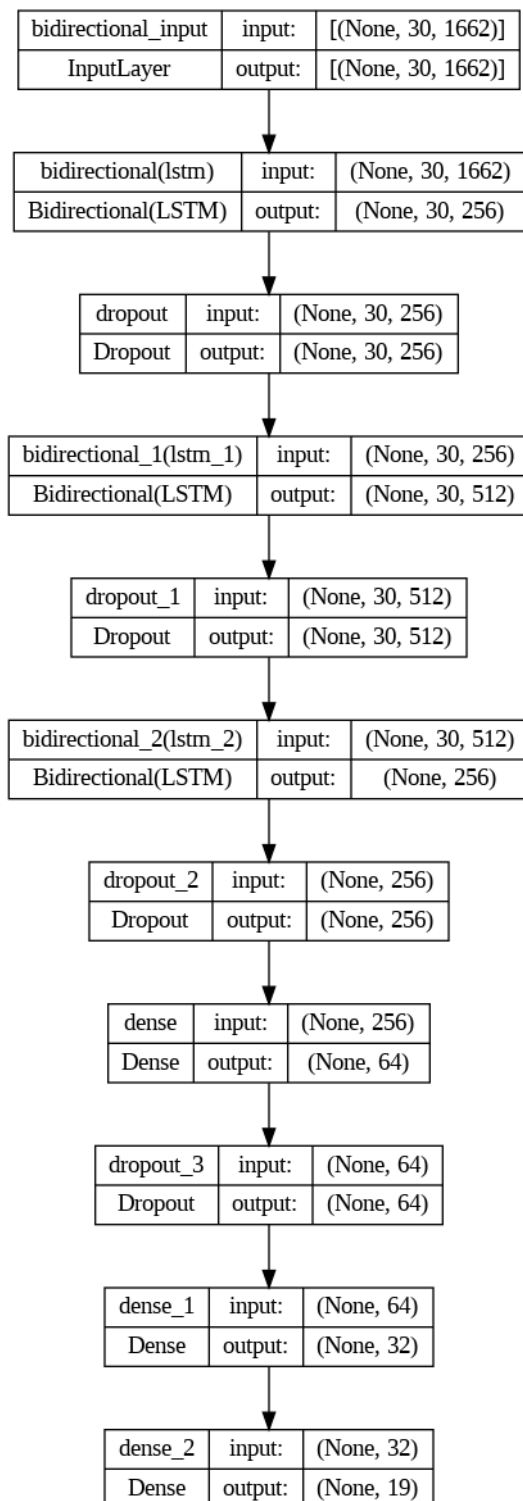


ภาพที่ 4.11 กราฟแสดงจำนวนรอบการเทรนและค่า Accuracy ของโมเดล GRU



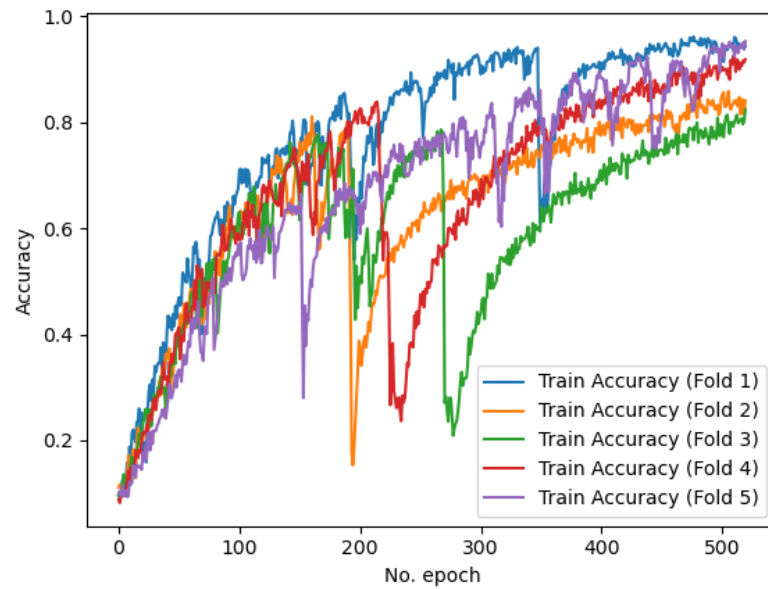
ภาพที่ 4.12 กราฟแสดงรอบการเทรนและค่า Loss ของโมเดล GRU

4.2.3 โมเดล BiLSTM

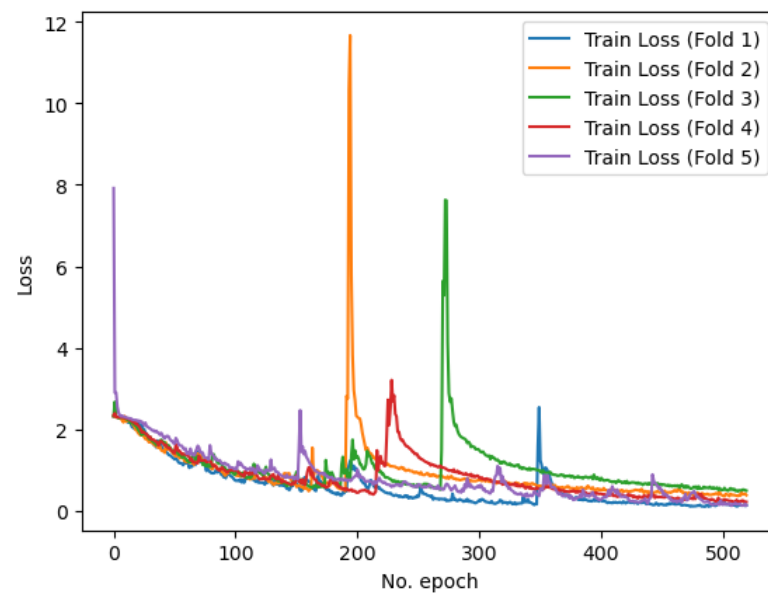


ภาพที่ 4.13 โครงสร้างโมเดล BiLSTM

ผลจากการพัฒนาระบบรู้จำท่าทางภาษามือไทยด้วย RNN พบว่า การกำหนดค่าอัตราการเรียนรู้ Epochs เท่ากับ 520 และ Batch Size เท่ากับ 512 ของโมเดล BiLSTM ให้ค่า Accuracy และค่า Loss ดังภาพด้านล่าง



ภาพที่ 4.14 กราฟแสดงจำนวนรอบการเทรนและค่า Accuracy ของโมเดล BiLSTM



ภาพที่ 4.15 กราฟแสดงรอบการเทรนและค่า Loss ของโมเดล BiLSTM

จากภาพของโครงสร้างโมเดลทั้ง 3 โมเดล ผู้วิจัยได้มีการใช้ Dense และ Dropout ซึ่ง Dense layer เป็นชั้นของโครงข่ายประสาทเทียมที่มีการเชื่อมต่อ (connection) ระหว่างโนด (node) ในชั้นก่อนหน้ากับชั้นต่อไป โดยทุกโนดในชั้นจะมีการเชื่อมต่อกับโนดทุกตัวในชั้นต่อไป ชั้น Dense จะมีการปรับค่าน้ำหนัก (weight) และค่าไบแอส (bias) ของโนดในแต่ละชั้น ซึ่งช่วยให้โมเดลสามารถเรียนรู้และสร้างรูปแบบ (pattern) ที่ซับซ้อนได้มากขึ้น

Dropout layer เป็นชั้นที่มีไว้เพื่อลดการเกิด overfitting ในโมเดล โดย Dropout layer จะสุ่มตัดการเชื่อมต่อระหว่างโนดในชั้นก่อนหน้ากับชั้นต่อไป โดยตัดการเชื่อมต่อเหล่านี้ด้วยการกำหนดค่าเป็นศูนย์ (zero) โดยสุ่มตัดบางโนดออกจากการคำนวณในแต่ละรอบการฝึกฝน การทำ Dropout จะช่วยให้โมเดลสามารถเรียนรู้และสร้างรูปแบบที่เหมาะสมกับข้อมูลได้ดีขึ้นโดยไม่เกิดการเรียนรู้ที่ผิดพลาดจาก overfitting

การใช้ Dense layer และ Dropout layer ในโมเดลประสาทเทียมช่วยให้โมเดลมีความแม่นยำในการทำนายข้อมูลมากขึ้น โดย Dense layer เป็นชั้นหลักในการสร้างโมเดลและ Dropout layer เป็นชั้นช่วยลดการเกิด overfitting ในโมเดลให้มีประสิทธิภาพสูงขึ้น การใช้ Dense layer และ Dropout layer นั้นเป็นเพิ่มความแม่นยำมากขึ้นนั่นเอง

4.3 ผลการวัดประสิทธิภาพโมเดล

4.3.1 ผลการวัดประสิทธิภาพโมเดล LSTM

จากการเทรนโมเดลด้วยชุดข้อมูลสำหรับเทรนได้ค่า Accuracy และค่า Loss ดังนี้

ตารางที่ 4.1 ผลการวัดประสิทธิภาพการเทรนโมเดล LSTM

อัตราส่วน	Mean Accuracy	Mean Loss
60:40	0.81	0.22
70:30	0.88	0.19
80:20	0.79	0.25
90:10	0.97	0.16

จากตารางพบว่าในอัตราส่วน 60:40 มีค่า Accuracy เฉลี่ยเท่ากับ 0.81 และค่า Loss เฉลี่ยเท่ากับ 0.22 อัตราส่วน 70:30 มีค่า Accuracy เฉลี่ยเท่ากับ 0.88 และค่า Loss เฉลี่ยเท่ากับ 0.19 อัตราส่วน 80:20 มีค่า Accuracy เฉลี่ยเท่ากับ 0.79 และค่า Loss เฉลี่ยเท่ากับ 0.25 และอัตราส่วน 90:10 มีค่า Accuracy เฉลี่ยเท่ากับ 0.97 และค่า Loss เฉลี่ยเท่ากับ 0.16

4.3.2 ผลการวัดประสิทธิภาพโมเดล GRU

จากการเทรนโมเดลด้วยชุดข้อมูลสำหรับเทรนได้ค่า Accuracy และค่า Loss ดังนี้

ตารางที่ 4.2 ผลการวัดประสิทธิภาพการเทรนโมเดล GRU

อัตราส่วน	Mean Accuracy	Mean Loss
60:40	0.88	0.24
70:30	0.90	0.18
80:20	0.82	0.25
90:10	0.97	0.15

จากตารางพบว่าในอัตราส่วน 60:40 มีค่า Accuracy เฉลี่ยเท่ากับ 0.88 และค่า Loss เฉลี่ยเท่ากับ 0.24 อัตราส่วน 70:30 มีค่า Accuracy เฉลี่ยเท่ากับ 0.90 และค่า Loss เฉลี่ยเท่ากับ 0.18 อัตราส่วน 80:20 มีค่า Accuracy เฉลี่ยเท่ากับ 0.82 และค่า Loss เฉลี่ยเท่ากับ 0.25 และอัตราส่วน 90:10 มีค่า Accuracy เฉลี่ยเท่ากับ 0.97 และค่า Loss เฉลี่ยเท่ากับ 0.16

4.3.3 ผลการวัดประสิทธิภาพโมเดล BiLSTM

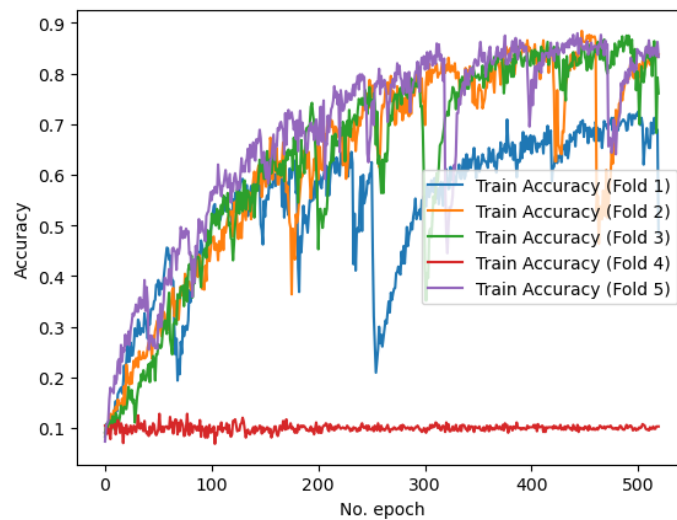
จากการเทรนโมเดลด้วยชุดข้อมูลสำหรับเทรนได้ค่า Accuracy และค่า Loss ดังนี้

ตารางที่ 4.3 ผลการวัดประสิทธิภาพการเทรนโมเดล BiLSTM

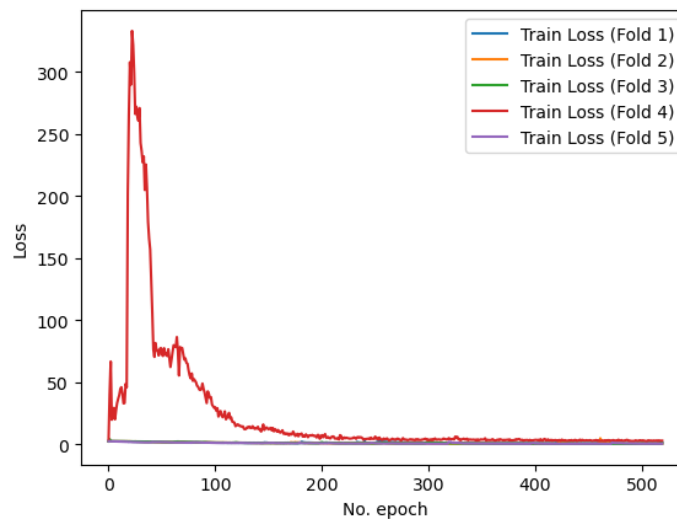
อัตราส่วน	Mean Accuracy	Mean Loss
60:40	0.50	3.12
70:30	0.36	2.24
80:20	0.51	1.12
90:10	0.91	0.28

จากตารางพบว่าในอัตราส่วน 60:40 มีค่า Accuracy เฉลี่ยเท่ากับ 0.50 และค่า Loss เฉลี่ยเท่ากับ 3.12 อัตราส่วน 70:30 มีค่า Accuracy เฉลี่ยเท่ากับ 0.36 และค่า Loss เฉลี่ยเท่ากับ 2.24 อัตราส่วน 80:20 มีค่า Accuracy เฉลี่ยเท่ากับ 0.51 และค่า Loss เฉลี่ยเท่ากับ 1.12 และอัตราส่วน 90:10 มีค่า Accuracy เฉลี่ยเท่ากับ 0.97 และค่า Loss เฉลี่ยเท่ากับ 0.28

จากผลการเทรนของโมเดล BiLSTM จะเห็นได้ว่าในอัตราส่วน 60:40, 70:30, 80:20 ค่า Loss เฉลี่ยสูงมากเมื่อเทียบโมเดล GRU และ LSTM เนื่องจากโมเดล BiLSTM นั้นมีความต้องการจำนวนข้อมูลสำหรับเทรนที่มากกว่า 2 โมเดลนั้นและเมื่อมีข้อมูลสำหรับเทรนน้อยแล้วยังนำข้อมูลมาแบ่ง 5-Fold Cross Validation อีกยิ่งทำให้ข้อมูลสำหรับเทรนน้อยลงไปอีก จึงง่ายต่อการเกิดค่า Loss ที่สูงของโมเดล BiLSTM ในอัตราส่วนนี้ดังภาพที่ 4.16

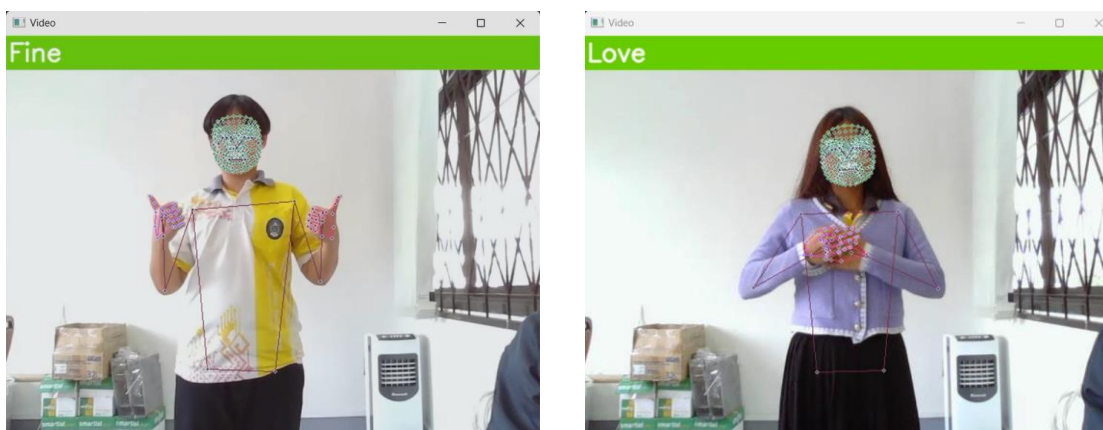


ภาพที่ 4.16 กราฟแสดงจำนวนรอบการเทรนและค่า Accuracy ของโมเดล BiSTM อัตราส่วน 80:20



ภาพที่ 4.17 กราฟแสดงจำนวนรอบการเทรนและค่า Loss ของโมเดล BiSTM อัตราส่วน 80:20

จากผลการวัดประสิทธิภาพของโมเดลทั้ง 3 โมเดล ผู้วิจัยจึงตัดสินใจนำโมเดลทั้ง 3 ที่มีการเทรนในอัตราส่วน 90:10 เนื่องจากมีค่า Accuracy ของโมเดลทั้ง 3 มีมากกว่า 90% และค่า Loss อยู่ในระดับที่สามารถยอมรับได้



ภาพที่ 4.18 ตัวอย่างผลการทดสอบโมเดล GRU

ตารางที่ 4.5 ผลการทดสอบโมเดล LSTM

คำศัพท์	Confusion Matrix				Accuracy
	TP	FP	TN	FN	
สบายดี	90	0	10	0	1.0
สวัสดี	90	0	10	0	1.0
หิว	90	0	5	5	0.95
ฉันท	87	3	10	0	0.97
ชอบ	89	1	10	0	0.99
รัก	90	0	10	0	1.0
ป่วย	90	0	10	0	1.0
ขอโทษ	90	0	10	0	1.0
ขอบคุณ	90	0	10	0	1.0
คุณ	89	1	10	0	0.99
รวม	895	5	95	5	0.99

จากตารางที่ 4.5 พบว่า คำศัพท์ "สบายดี", "สวัสดิ์", "รัก", "ป่วย", "ขอโทษ" และ "ขอบคุณ" มีความแม่นยำสูงสุดที่ 100%

คำศัพท์ "ชอบ" และ "คุณ" มีความแม่นยำที่ 99%

คำศัพท์ "ฉัน" มีความแม่นยำที่ 97%

คำศัพท์ "หิว", "คุณ" มีความแม่นยำที่ 95%

รวมแล้วค่า Accuracy รวมทั้งหมดเท่ากับ 99%

ตารางที่ 4.6 ผลการทดสอบโมเดล GRU

คำศัพท์	Confusion Matrix				Accuracy
	TP	FP	TN	FN	
สบายดี	90	0	10	0	1.0
สวัสดิ์	90	0	10	0	1.0
หิว	89	1	9	1	0.98
ฉัน	89	1	9	1	0.98
ชอบ	90	0	10	0	1.0
รัก	90	0	10	0	1.0
ป่วย	90	0	10	0	1.0
ขอโทษ	90	0	10	0	1.0
ขอบคุณ	90	0	10	0	1.0
คุณ	90	0	10	0	1.0
รวม	898	2	98	2	0.99

จากตารางที่ 4.6 พบว่าคำศัพท์ "สบายดี", "สวัสดิ์", "ชอบ", "รัก", "ป่วย", "ขอโทษ", "ขอบคุณ" และ "คุณ" มีความแม่นยำสูงสุดที่ 100%

คำศัพท์ "หิว" และ "ฉัน" มีความแม่นยำอยู่ที่ 98%

รวมแล้วค่า Accuracy รวมทั้งหมดเท่ากับ 99%

ตารางที่ 4.7 ผลการทดสอบโมเดล BiLSTM

คำศัพท์	Confusion Matrix				Accuracy
	TP	FP	TN	FN	
สบายดี	90	0	10	0	1.0
สวัสดิ์	90	0	9	1	0.99
หิว	85	5	9	1	0.94
ฉัน	90	0	7	3	0.97
ชอบ	90	0	10	0	1.0
รัก	88	2	10	0	0.98
ป่วย	89	1	10	0	0.99
ขอโทษ	90	0	8	2	0.98
ขอบคุณ	90	0	10	0	1.0
คุณ	89	1	8	2	0.97
รวม	891	9	91	9	0.98

จากตารางที่ 4.7 พบว่าคำศัพท์ "สบายดี", "ชอบ" และ "ขอบคุณ" มีความแม่นยำอยู่ที่ 100%
 คำศัพท์ "ป่วย" และ "สวัสดิ์" มีความแม่นยำอยู่ที่ 99%
 คำศัพท์ "รัก" และ "ขอโทษ" มีความแม่นยำอยู่ที่ 98%
 คำศัพท์ "ฉัน" และ "คุณ" มีความแม่นยำอยู่ที่ 97%
 คำศัพท์ "หิว" มีความแม่นยำที่ 94%
 รวมแล้วค่า Accuracy รวมทั้งหมดเท่ากับ 98%

4.4 ผลการเปรียบเทียบประสิทธิภาพโมเดล

ผลจากการพัฒนาและวัดประสิทธิภาพของโมเดลทั้ง 3 แบบได้แก่ LSTM, GRU และ BiLSTM ที่ใช้ในการพัฒนาระบบรู้จำท่าทางภาษามือไทยด้วย RNN ซึ่งแสดงได้ดังต่อไปนี้

ตารางที่ 4.4 เปรียบประสิทธิภาพของโมเดลทั้ง 3 โมเดล

Model	Epoch	Train Accuracy	Train Loss	Test Accuracy
LSTM	300	0.97	0.16	0.99
GRU	370	0.97	0.15	0.99
BiLSTM	520	0.91	0.28	0.98

จากตารางเปรียบเทียบประสิทธิภาพโมเดลทั้ง 3 โมเดล GRU และ LSTM มีค่า Accuracy สูงอยู่ที่ 0.99 ส่วนโมเดล BiLSTM มีค่า Accuracy อยู่ที่ 0.91 โมเดล GRU มีจำนวนรอบการในการเทรนอยู่ที่ 370 รอบและโมเดล LSTM ที่มีจำนวนการเทรน 300 รอบ แต่ให้ค่า Accuracy เท่ากับโมเดล GRU นอกจากความแตกต่างของจำนวนในการเทรนของโมเดลแล้วยังมีเรื่องของเวลาที่ใช้ในการเทรนที่แตกต่างระหว่างโมเดล LSTM และโมเดล GRU แม้ว่าจำนวนในการเทรนของโมเดล LSTM จะน้อยกว่าโมเดล GRU แต่ใช้เวลาในการเทรนมากกว่าเนื่องจากจำนวนพารามิเตอร์ของโมเดล GRU นั้นน้อยกว่าจึงทำให้ใช้เวลาในการเทรนน้อยกว่าแม้จำนวนรอบที่ใช้ในการเทรนจะมากกว่าส่วนโมเดล BiLSTM มีจำนวนรอบในการเทรนมากที่สุดและใช้เวลาในหนึ่งรอบการเทรนมากที่สุดเนื่องจากตัวโมเดลมีการใช้จำนวนโหนดมากกว่า 2 โมเดล จึงทำให้ใช้เวลานานในการเทรน

ตารางที่ 4.8 การเปรียบเทียบความเหมือนความแตกต่างในงานวิจัย

งานวิจัย	Language	Class	Key Points	Dataset (videos)	Video size	Train Splitting	Model	Train Accuracy	Train Loss	Test Accuracy	Test Loss
งานวิจัยของ A. Chaikaew	Thai	5	42	500	-	Train: 60 Validate: 30 Test:10	LSTM, GRU, BiLSTM	0.94 0.93 0.97	0.16 0.20 0.39	0.97 0.94 0.94	0.06 0.14 0.23
งานวิจัยของ Gerges H. Samaan	English	10	258, 1662	750	640 x 480	Train: 60 Test: 40	LSTM, GRU, BiLSTM	0.99 1.0 0.99	-	0.99 1.0 0.99	-
งานวิจัยของ ผู้วิจัย	Thai	10	1662	1000	640 X 480	60:40 70:30 80:20 90:10	LSTM, GRU, BiLSTM	0.97 0.97 0.91	0.16 0.15 0.28	0.99 0.99 0.98	-

จากในตารางที่ 4.8 ที่เป็นการเปรียบเทียบความเหมือนความแตกต่างในงานวิจัยสามารถอธิบายได้ดังต่อไปนี้

ในส่วนที่ 1 คืองานวิจัยที่นำมาเปรียบเทียบกับงานวิจัยนี้ได้แก่ งานวิจัยของ A. Chaikaew, K. Somkuan and T. Yuyen (2564) และงานวิจัยของ Gerges H. Samaan, Abanoub R. Widie, Abanoub K. Attia, Abanoub M. Asaad, Andrew E. Kamel, Salwa O. Slim, Mohamed S. Abdallah and Young-Im Cho (2022)

ในส่วนที่ 2 คือ ภาษาที่ใช้ในการทำระบบรู้จำภาษามือ งานวิจัยที่ 1 ใช้ภาษาไทย งานวิจัยที่ 2 ใช้ภาษาอังกฤษ และงานวิจัยของผู้วิจัยนั้นใช้ไทยเหมือนกับงานวิจัยที่ 1

ในส่วนที่ 3 คือ จำนวน class หรือจำนวนท่าภาษามือที่ใช้ในการวิจัย งานวิจัยที่ 1 ใช้ท่าทั้งหมด 5 ท่าภาษามือ งานวิจัยที่ 2 ใช้ท่าทั้งหมด 10 ท่าภาษามือ และงานวิจัยนี้ใช้ท่าทั้งหมด 10 ท่าภาษามือ

ในส่วนที่ 4 คือ จำนวน Key Points ที่จะสกัดนำมาให้โครงข่ายประสาทเทียมรู้จำ งานวิจัยที่ 1 ใช้ Key Points 42 Key Points ซึ่งได้มาจากการนำ Key Points ของมือซึ่งมีทั้งหมด 21 Key Points ของมือแต่ละข้างมารวมกันจึงได้ 42 Key Points ส่วนในงานวิจัยที่ 2 ใช้ แบ่งการทดลองออกเป็น 2 แบบได้แก่ การทดลองที่ 1 คือ การใช้ Key Points ในส่วนของมือและตัวซึ่งจะได้ Key Points ทั้ง 258 Key Points ส่วนในการทดลองที่ 2 จะใช้ 1662 ซึ่งเป็น Key Points ที่มีการรวม Key Points ของมือ ตัวและหน้าเข้าไปด้วย และในงานวิจัยนี้ได้ใช้ 1662 เหมือนกับการทดลองที่ 2 ของงานวิจัยที่ 2

ในส่วนที่ 5 คือ dataset หรือวิดีโอที่จะนำมาเทรนโมเดลงานวิจัยที่ 1 ใช้วิดีโอทั้งหมด 500 วิดีโอหรือ 100 วิดีโอต่อ 1 ท่าภาษามือ งานวิจัยที่ 2 ใช้วิดีโอทั้งหมด 750 วิดีโอทั้งหมด 750 วิดีโอหรือ 75 วิดีโอต่อ 1 ท่าภาษามือและในงานวิจัยนี้ใช้วิดีโอทั้งหมด 850 วิดีโอหรือ 85 วิดีโอต่อ 1 ท่าภาษามือ

ในส่วนที่ 6 คือ ขนาดของวิดีโอที่ถ่ายงานวิจัยที่ 1 นั้นไม่ได้ระบุไว้ ส่วนงานวิจัยที่ 2 และงานวิจัยนี้ใช้ขนาดวิดีโอเท่ากับ 640×480

ในส่วนที่ 7 คือการแบ่งข้อมูลสำหรับเทรนโมเดลและทดสอบโมเดลในงานวิจัยที่ 1 นั้นมีการแบ่งข้อมูลออกเป็น 3 ส่วนได้แก่ ส่วนที่เทรน 60% ส่วนในการทำ validation 30% และในการทดสอบ 10% ในงานวิจัยที่ 2 นั้นแบ่งข้อมูลออกเป็น 2 ส่วนได้แก่ 60% สำหรับเทรนโมเดลและ 40% สำหรับทดสอบ ส่วนในงานวิจัยครั้งนี้ได้แบ่งข้อมูลออกเป็น 60:40, 70:30, 80:20, 90:10 แล้วนำข้อมูลสำหรับการเทรนมาทำ 5-Fold Cross Validation และอีก 10% สำหรับทดสอบและหา accuracy ด้วยการทำ Confusion Matrix

ในส่วนที่ 8 คือโมเดลที่ทั้ง 3 งานวิจัยใช้ในการเทรนโมเดลและทดสอบโมเดล ได้แก่ Long Short-Term Memory : LSTM, Gate Recurrent Unit : GRU, Bidirectional Long Short-Term Memory : BiLSTM

ในส่วนที่ 9 คือ ผลค่า Accuracy ของการเทรนโมเดลในแต่ละโมเดล งานวิจัยที่ 1 มีค่า Train Accuracy ได้แก่ LSTM, GRU, BiLSTM เป็น 0.94, 0.93, 0.97 ตามลำดับ งานวิจัยที่ 2 มีค่า Train Accuracy ได้แก่ LSTM, GRU, BiLSTM เป็น 0.99, 1.0, 0.99 ตามลำดับและงานวิจัยนี้ มีค่า Train Accuracy ได้แก่ LSTM, GRU, BiLSTM เป็น 0.97, 0.97, 0.91 ตามลำดับ

ในส่วนที่ 10 คือ ผลค่า Loss ของการเทรนโมเดลในแต่ละโมเดล งานวิจัยที่ 1 มีค่า Train Loss ได้แก่ LSTM, GRU, BiLSTM เป็น 0.16, 0.20, 0.39 ตามลำดับ งานวิจัยที่ 2 ไม่ได้ระบุค่า Loss ไว้ และงานวิจัยนี้ มีค่า Train Loss ได้แก่ LSTM, GRU, BiLSTM เป็น 0.16, 0.15, 0.28 ตามลำดับ

ในส่วนที่ 11 คือ ผลค่า Accuracy ของการทดสอบโมเดลในแต่ละโมเดล งานวิจัยที่ 1 มีค่า Train Accuracy ได้แก่ LSTM, GRU, BiLSTM เป็น 0.97, 0.94, 0.94 ตามลำดับ งานวิจัยที่ 2 มีค่า Train Accuracy ได้แก่ LSTM, GRU, BiLSTM เป็น 0.99, 1.0, 0.99 ตามลำดับและงานวิจัยนี้ มีค่า Train Accuracy ได้แก่ LSTM, GRU, BiLSTM เป็น 0.97, 0.97, 0.91 ตามลำดับ

ในส่วนที่ 12 คือ ผลค่า Loss ของการทดสอบโมเดลในแต่ละโมเดล งานวิจัยที่ 1 มีค่า Train Loss ได้แก่ LSTM, GRU, BiLSTM เป็น 0.06, 0.14, 0.23 ตามลำดับ งานวิจัยที่ 2 ไม่ได้ระบุค่า Loss ไว้ และงานวิจัยนี้เนื่องจากการหา Accuracy ด้วยวิธีการทำ Confusion Matrix จึงทำให้ไม่มีค่า Loss