ระบบการรู้จำเครื่องหมายจราจรของประเทศไทยโดยใช้โครงข่ายประสาทเทียม และซัพพอร์ตเวกเตอร์แมชชีน

ทศม วงศ์ช่วย[1] วิภาดา เวทย์ประสิทธิ์[2]

[1]หลักสูตรมหาบัณฑิตจัดการเทคโนโลยีสารสนเทศ คณะวิศวกรรมศาสตร์

[2]ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

มหาวิทยาลัยสงขลานครินทร์ อ.หาดใหญ่ จ.สงขลา 90112

E-mail:ThasamW@gsb.or.th, wwettayaprasit@yahoo.com

บทคัดย่อ

ปัจจุบันมีการพัฒนาระบบคอมพิวเตอร์มาใช้ร่วมกับรถยนต์ให้มีความสามารถมากขึ้นเช่นการนำคอมพิวเตอร์มาช่วยในการจำแนกเครื่องหมายจราจร บทความนี้ได้เสนอวิธีการจำแนกเครื่องหมายจราจรของประเทศไทยจำนวน 30 แบบ เครื่องหมายจราจรแต่ละแบบจะทำการสอนให้เครื่องคอมพิวเตอร์ได้เรียนรู้โดยใช้โครงข่ายประสาทเทียมเปรียบเทียบกับซัพพอร์ตเวกเตอร์แมชชีน การทดลองใช้ภาพถ่ายจากภาคใต้ของประเทศไทยจำนวน 960 ภาพ ผลการทดลองโครงข่ายประสาทเทียมมีค่าความถูกต้อง 99.84% ซัพพอร์ตเวกเตอร์แมชชีนมีค่าความถูกต้อง 99.94% ในขณะที่ซัพพอร์ตเวกเตอร์แมชชีนใช้เวลาเฉลี่ยในการทำนายน้อยกว่าโครงข่ายประสาทเทียม

คำหลัก — เครื่องหมายจราจร; การรู้จำ; โครงข่ายประสาทเทียม; ซัพพอร์ตเวกเตอร์แมชชีน

1. บทนำ

เครื่องหมายจราจรเป็นสิ่งสำคัญสำหรับการขับขี่ยานยนต์ในปัจจุบัน โดยที่เครื่องหมายจราจรกำหนดไว้เพื่อความปลอดภัยในการขับขี่ สาเหตุหนึ่งของอุบัติเหตุเนื่องจากผู้ขับขี่ไม่ปฏิบัติตามกฎจราจร และไม่เข้าใจความหมายของเครื่องหมายจราจร จากปัญหาดังกล่าวทำให้มีงานวิจัยเกี่ยวกับการการช่วยผู้ขับขี่ [1] [2] และการนำคอมพิวเตอร์เข้ามามีส่วนช่วยเตือนและแนะนำผู้ขับขี่ยานยนต์ [3] โดยนำคอมพิวเตอร์มาช่วยประกอบการตัดสินใจของผู้ขับขี่โดยอาศัยโครงข่ายประสาทเทียม [1] หรือมีการประยุกต์ใช้ซัพพอร์ตเวกเตอร์แมชชีน [2] [5] โดยมีจุดมุ่งหมายในการทำให้คอมพิวเตอร์สามารถจำแนกเครื่องหมายจราจรได้ [4] [5]

บทความนี้ประกอบด้วยส่วนต่างๆดังต่อไปนี้ส่วนที่ 2 ทฤษฎีและหลักการ ส่วนที่ 3 ระบบรู้จำเครื่องหมายจราจร ส่วนที่ 4 ผลการทดลอง และส่วนที่ 5 คือสรุป

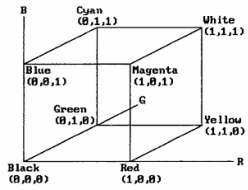
2. ทฤษฎีและหลักการ

2.1 รูปแบบของสี

มาตรฐานการนำภาพสีมาใช้ในระบบคอมพิวเตอร์สามารถใช้ 2 ลักษณะ คือ RGB หรือ HSV ดังนี้

2.1.1 RGB

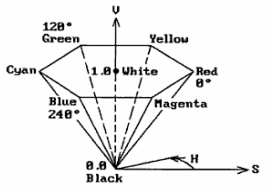
RGB (Red Green Blue) คือมาตรฐานสีโดยมีสีหลัก 3 สีคือ สีแดง สีเขียว สีน้ำเงิน ภาพสีที่ได้ที่ใช้หลักการผสมผสานของแม่สีทั้ง 3 สีนี้ [6] โดยสีที่ได้มีลักษณะใกล้เคียงกับสีธรรมชาติ และใช้แบบจำลองแสดงดังรูปที่ 1



รูปที่ 1 แบบจำลองสี RGB [6]

2.1.2 HSV

HSV (Hue-Saturation-Value) คือมาตรฐานของสี [1] ที่กำหนดความเข้มของสี โดยมีค่า H คือค่าสีของภาพ S คือ ค่าสเปกตรัมของสีโดยที่มีค่าตั้งแต่ 0 ถึง 255 โดยที่ทำการเปลี่ยนเป็นองศา ตั้งแต่ 0 ถึง 360 องศา โดยแต่ละองศามีค่าสีแตกต่างกัน และ V คือ ค่าความสว่างโดยมีค่าระหว่าง 0 ถึง 1 โดยใช้แบบจำลองแสดงดังรูปที่ 2



รูปที่ 2 แบบจำลอง HSV [1]

การแปลงมาตรฐาน RGB ไปเป็นมาตรฐาน HSV สามารถทำได้โดยใช้สมการ (1) ดังนี้

กำหนดให้ $m = \max(R,G,B)$ $n = \min(R,G,B)$ R คือค่าของสีแดง

G คือค่าสีเขียว และ B คือค่าของสีน้ำเงิน

$$H = 60° \times h \quad (1)$$

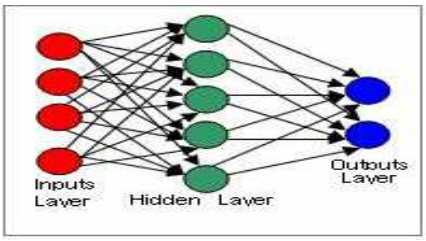
$$h = \begin{cases} (G-B)/(m-n) & if \ \ G = m \\ 2+(B-R)/(m-n) & if \ \ B = m \\ 4+(R-B)/(m-n) & if \ \ R = m \\ 0 & if \ \ m = 0 \end{cases} \quad (2)$$

$$S = \begin{cases} (m-n)/m & if \ \ m \neq 0 \\ 0 & if \ \ m = 0 \end{cases} \quad (3)$$

$$V = m \quad (4)$$

2.2 โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม หรือ เรียกว่า Neural Networks (NN) คือโมเดลทางคณิตศาสตร์ที่จำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ ซึ่งประกอบด้วย เซลประสาท หรือ “นิวรอน” รูปที่ 3 แสดงสถาปัตยกรรมของโครงข่ายประสาทเทียมที่ประกอบด้วยชั้นข้อมูลเข้า ชั้นซ่อนและชั้นข้อมูลออก แต่ละนิวรอนสามารถคำนวณได้จากสมการที่ (5) – (6)



รูปที่ 3 สถาปัตยกรรมโครงข่ายประสาทเทียม

$$y_j^r = f(v_j^r) \quad (5)$$

$$v_j^r = \sum_{i=1}^{n_{r-1}} w_{ij}^r y_i^{r-1} + b_j^r \quad (6)$$

โดยที่ w_{ij}^r คือ ค่าน้ำหนัก

y_i^r คือ ผลลัพธ์ที่ได้จากการคำนวณ

b_j^r คือ ค่าไบอัส

n คือ จำนวนโหนด

r คือ ชั้นที่1,2,3,….,L

2.3 ซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีนหรือเรียกว่า SVM เป็นการแบ่งข้อมูลสองกลุ่มออกจากกันโดยมีหลายระนาบ [8] กำหนดให้ $(y_1, x_1), ..., (y_i, x_i)$ เป็นตัวอย่างที่ใช้ในการสอนโดยที่ i มีค่าตั้งแต่ 1 ถึง n โดยที่ n คือจำนวนข้อมูลตัวอย่าง m คือจำนวนมิติข้อมูลเข้าจะได้ว่า $x \in R^m$ และ y คือ ผลลัพธ์มีค่า +1 หรือ -1 หรือ $y_i \in \{-1, 1\}$

สำหรับปัญหาเชิงเส้นมิติข้อมูลสามารถแบ่งเป็น 2 กลุ่มโดยระนาบตัดสินใจดังสมการที่ (7) โดยที่ w คือ ค่าน้ำหนัก และ b คือค่าไบอัส

$$(w \cdot x) + b = 0 \quad (7)$$

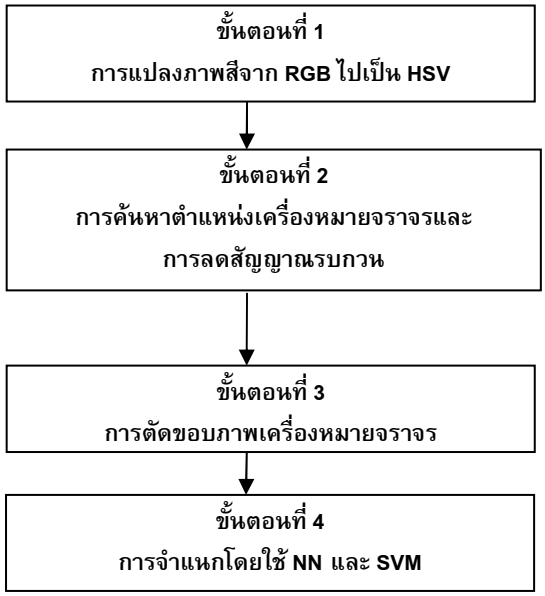
สำหรับการแบ่งกลุ่มให้เป็น 2 กลุ่มในระนาบการตัดสินใจสามารถคำนวณได้จากสมการที่ (8)-(9)

$$(w \cdot x) + b > 0 \ if \ \ y_i = +1 \quad (8)$$

$$(w \cdot x) + b < 0 \ if \ \ y_i = -1 \quad (9)$$

3. ระบบรู้จำเครื่องหมายจราจร

ระบบรู้จำเครื่องหมายจราจรมีการออกแบบเป็น 4 ส่วน ส่วนที่ 1 คือการแปลงภาพจาก RGB เป็น HSV ส่วนที่ 2 คือการค้นหาตำแหน่งเครื่องหมายจราจรและลดสัญญาณรบกวน ส่วนที่ 3 คือการตัดขอบภาพเครื่องหมายจราจร ส่วนที่ 4 คือการจำแนกโดยโครงขายประสาทเทียม และซัพพอร์ตเวกเตอร์แมชชีน ดังรูปที่ 4



รูปที่ 4 แสดงขั้นตอนการดำเนินการรู้จำเครื่องหมายจราจร

3.1 การแปลงภาพสีจาก RGB ไปเป็น HSV

ขั้นตอนนี้เป็นการแปลงภาพ ขนาด 2048 x 1536 พิกเซล ที่มีมาตรฐาน RGB ไปเป็นมาตรฐาน HSV เพื่อหาตำแหน่งของสีเครื่องหมายจราจร รูปที่ 5 แสดงตัวอย่างการแปลงภาพ RGB ไปเป็น HSV ของเครื่องหมายจราจรโค้งขวา

3.2 การค้นหาตำแหน่งเครื่องหมายจราจรและการลดสัญญาณรบกวน

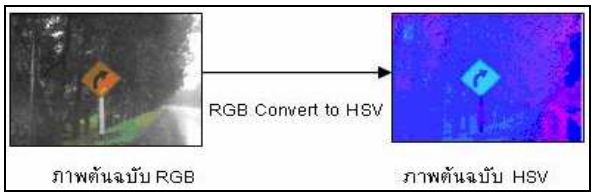
นำภาพที่อยู่ในรูปแบบ HSV และนำภาพที่ได้มาดำเนินการ 2 ขั้นตอนย่อย คือ A) คือการหาสีป้ายโดยกรองเลือกเอาเฉพาะสีเหลืองที่เป็นตำแหน่งป้ายที่ต้องการ และ B) คือการลดสัญญาณรบกวนภาพที่มีอยู่ในภาพออกโดยลดจำนวนจุดสีที่อยู่นอกป้ายออกจากกรอบของเครื่องหมายจราจร แสดงดังรูปที่ 6

3.3 การตัดขอบภาพเครื่องหมายจราจร

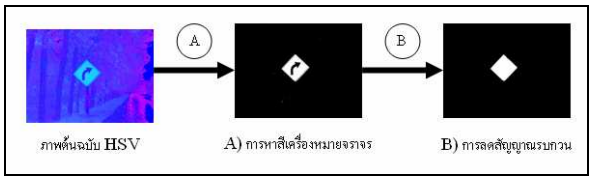
ขั้นตอนนี้เป็นการตัดขอบภาพเครื่องหมายจราจรโดยทำการย่อขนาดเป็น 50x50 พิกเซล โดยทำการแทนค่าพื้นหลังของสี่เหลี่ยมที่อยู่นอกตำแหน่งของเครื่องหมายจราจรด้วยสีขาว เพื่อพร้อมในการทำเป็นข้อมูลเพื่อใช้สอนและทดสอบ กับ NN และ SVM ต่อไป แสดงได้ดังรูปที่ 7

3.4 การจำแนกโดยใช้ NN และ SVM

เมื่อได้ภาพมาก็ดำเนินการจำแนกโดยทำการเปรียบเทียบ 2 วิธี คือ NN และ SVM



รูปที่ 5 การแปลงภาพจาก RGB ไปยัง HSV



รูปที่ 6 ค้นหาเครื่องหมายจราจรและลดสัญญาณรบกวน

รูปที่ 7 ภาพที่ตัดขอบภาพโดยย่อขนาด 50x50 พิกเซล

3.4.1 โครงข่ายประสาทเทียม (NN)

จากภาพขนาด 50x50 พิกเซล เมื่อพิจารณาเฉพาะตำแหน่งของเครื่องหมายจราจรที่เป็นรูปสี่เหลี่ยมข้าวหลามตัด(สีขาวขอบนอกไม่ต้องนำมาพิจารณา)จะสามารถลดขนาดการนำเข้าข้อมูลเพียง 25x25 พิกเซล ทำการส่งเข้าโครงข่ายประสาทเทียมโดยขนาดของโหนดข้อมูลเข้าคือ 625 โหนด (25x25) ขนาดของโหนดซ่อนคือ1,300 โหนด และขนาดโหนดผลลัพธ์คือ 1 โหนด สถาปัตยกรรมของโครงข่ายประสาทเทียมที่ได้ (โหนดข้อมูลนำเข้า:โหนดซ่อน:โหนดผลลัพธ์) คือ (625:1,300:1) โดยใช้ฟังก์ชั่นการเรียนรู้โครงข่ายเป็นแบบ Gradient Descent with Momentum Method (traingdx) ดังรูปที่ 8

3.4.2 ซัพพอร์ตเวกเตอร์แมชชีน (SVM)

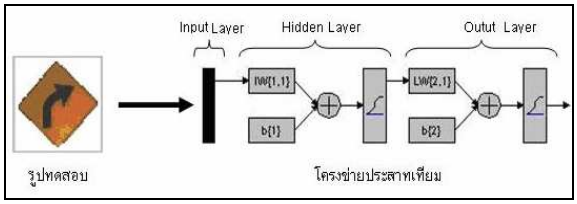
ขนาดของโหนดข้อมูลเข้าคือ 625 โหนด และโหนดผลลัพธ์คือ 1 โหนด และกำหนดเคอร์เนลเป็นแบบ Linear

4. ผลการทดลอง

การทดลองมีการเก็บข้อมูลจากภาพถ่ายจากเครื่องหมายจราจรจริงจากถนนในภาคใต้ของประเทศไทยในจังหวัด สงขลา พัทลุง นครศรีธรรมราช ปัตตานี ยะลา และนราธิวาส จำนวน 960 ภาพ โดยนำภาพเครื่องหมายจราจรที่พบบ่อย 30 แบบดังตารางที่ 1 แบบละ 32 ภาพ รวมต้นแบบจากกรมการขนส่งทางบกอีกแบบละ 1 ภาพ โดยที่เครื่องหมายจราจรแต่ละแบบมีการข้อมูลสำหรับสอนจำนวน 46 ภาพและทดสอบจำนวน 300 ภาพ ดังแสดงตารางที่ 2 ซึ่งเป็นตัวอย่างของเครื่องหมายจราจรเลี้ยวขวา

จากการทดลองเปรียบเทียบความถูกต้องโดยทำการทดลองแบบโครงข่ายประสาทเทียม 1 โครงข่ายต่อเครื่องหมายจราจร 1 แบบ ดังนั้นจะมีโครงข่ายประสาทเทียมจำนวน 30 โครงข่ายในทำนองเดียวกันกับซัพพอร์ตเวกเตอร์แมชชีน ผลการทดลองแสดงดังรูปที่ 9 กลุ่ม a b และ d นั้นมีความถูกต้องเท่ากันทั้ง NN และ SVM คือ 100 % กลุ่ม c นั้นมีความถูกต้องเท่ากันทั้ง NN และ SVM คือ 99.93% แต่ในกลุ่ม e และ f พบว่า SVM มีความถูกต้องมากกว่า NN จากค่าเฉลี่ยพบว่า NN มีค่าเฉลี่ยความถูกต้อง 99.84% และ SVM มีค่าเฉลี่ยความถูกต้องซึ่งสูงกว่าเล็กน้อยคือ 99.94% สำหรับรายละเอียดของการทดลองของ NN และ SVM หาค่า True Positive (TP) False Positive (FP) False Negative (FN) และ True Negative (TN) แสดงดังรูปที่ 10(a) และ (b) ตามลำดับ

การทดลองเวลาของการทำนายเปรียบเทียบระหว่าง NN กับ SVM แสดงดังรูปที่ 11 ผลการทดลองของ SVM ใช้เวลาน้อยกว่า NN โดยที่ SVM ใช้เวลาเฉลี่ยที่ 0.007 วินาที แต่ NN ใช้เวลาเฉลี่ยที่ 0.356 วินาที



รูปที่ 8 สถาปัตยกรรมโครงข่ายประสาทเทียม

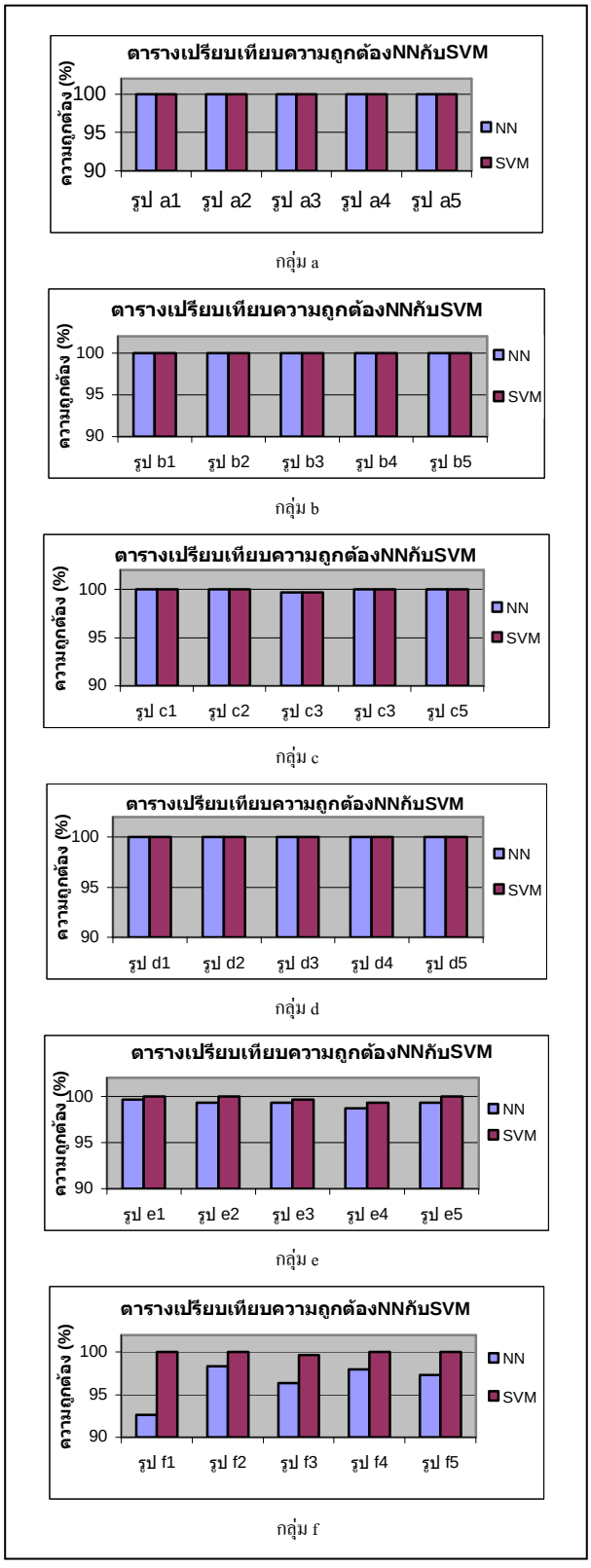
ตารางที่ 1 แสดงเครื่องหมายจราจรจำนวน 30 แบบ

ภาพ / กลุ่ม	เครื่องหมายจราจร				
	1	2	3	4	5
a					
b					
c					
d					
e					
f					

ตารางที่ 2 ตัวอย่างข้อมูลสำหรับการสอนและการทดสอบของเครื่องหมายจราจรเลี้ยวขวา

จำนวนภาพ	ภาพใช่	ภาพไม่ใช่ (สุ่มมา)
ใช้สอน(Train Set) 46 ภาพ	23 ภาพ	23 ภาพ
ใช้ทดสอบ(Test Set)300 ภาพ	10 ภาพ	290 ภาพ
ภาพที่เก็บในแต่ละแบบ	**33**	**957**

หมายเหตุ: ภาพที่นำมาทดสอบเป็นภาพที่ไม่ซ้ำกับภาพที่ใช้สอน



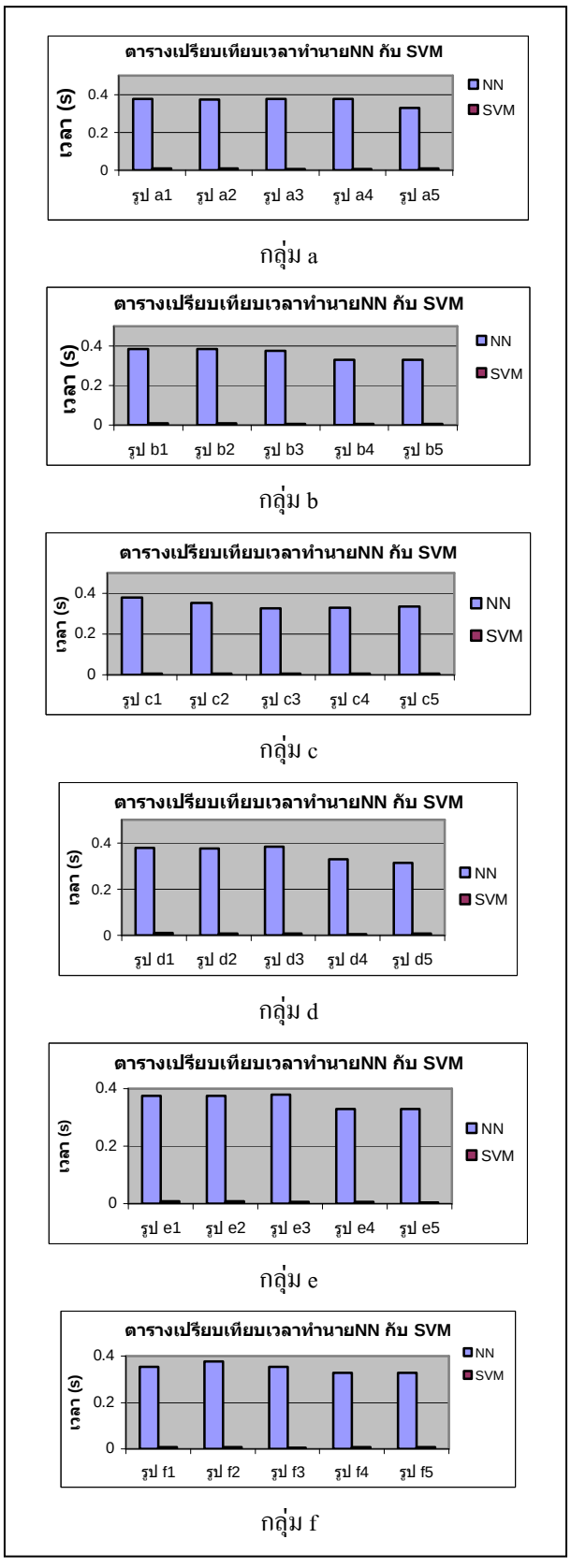
รูปที่ 9 ค่าความถูกต้องของ NN และ SVM

ชนิดข้อมูล / ภาพ	ข้อมูลจริง (Actual)		ข้อมูลทำนาย (Prediction)		ความถูกต้อง (%)
	True Positive	False Positive	False Negative	True Negative	
ภาพ a1	10	0	0	290	100
ภาพ a2	10	0	0	290	100
ภาพ a3	10	0	0	290	100
ภาพ a4	10	0	0	290	100
ภาพ a5	10	0	0	290	100
ภาพ b1	10	0	0	290	100
ภาพ b2	10	0	0	290	100
ภาพ b3	10	0	0	290	100
ภาพ b4	10	0	0	290	100
ภาพ b5	10	0	0	290	100
ภาพ c1	10	0	0	290	100
ภาพ c2	10	0	0	290	100
ภาพ c3	10	0	1	289	99.67
ภาพ c4	10	0	0	290	100
ภาพ c5	10	0	0	290	100
ภาพ d1	10	0	0	290	100
ภาพ d2	10	0	0	290	100
ภาพ d3	10	0	0	290	100
ภาพ d4	10	0	0	290	100
ภาพ d5	10	0	0	290	100
ภาพ e1	10	0	1	289	99.67
ภาพ e2	8	2	0	290	99.33
ภาพ e3	10	0	2	288	99.33
ภาพ e4	6	4	0	290	98.67
ภาพ e5	8	2	0	290	99.33
ภาพ f1	10	0	22	268	92.67
ภาพ f2	7	3	2	288	98.33
ภาพ f3	10	0	11	279	96.33
ภาพ f4	10	0	6	284	98
ภาพ f5	9	1	8	282	97

รูปที่ 10 (a) ค่าConfusion matrix ของ NN

ชนิดข้อมูล / ภาพ	ข้อมูลจริง (Actual)		ข้อมูลทำนาย (Prediction)		ความถูกต้อง (%)
	True Positive	False Positive	False Negative	True Negative	
ภาพ a1	10	0	0	290	100
ภาพ a2	10	0	0	290	100
ภาพ a3	10	0	0	290	100
ภาพ a4	10	0	0	290	100
ภาพ a5	10	0	0	290	100
ภาพ b1	10	0	0	290	100
ภาพ b2	10	0	0	290	100
ภาพ b3	10	0	0	290	100
ภาพ b4	10	0	0	290	100
ภาพ b5	10	0	0	290	100
ภาพ c1	10	0	0	290	100
ภาพ c2	10	0	0	290	100
ภาพ c3	10	0	0	290	100
ภาพ c4	10	0	0	290	100
ภาพ c5	10	0	0	290	100
ภาพ d1	10	0	0	290	100
ภาพ d2	10	0	0	290	100
ภาพ d3	10	0	0	290	100
ภาพ d4	10	0	0	290	100
ภาพ d5	10	0	0	290	100
ภาพ e1	10	0	0	290	100
ภาพ e2	8	2	0	290	99.33
ภาพ e3	10	0	1	289	99.67
ภาพ e4	10	0	2	288	99.33
ภาพ e5	8	2	0	290	99.33
ภาพ f1	10	0	22	268	92.67
ภาพ f2	7	3	2	288	98.33
ภาพ f3	10	0	1	289	99.67
ภาพ f4	10	0	6	284	98
ภาพ f5	9	1	8	282	97

รูปที่ 10 (b) ค่าConfusion matrix ของ SVM



รูปที่ 11 เวลาของ NN และ SVM

5. สรุป

บทความนี้เป็นการทดลองรู้จำเครื่องหมายจราจรโดยใช้โครงข่ายประสาทเทียมกับซัพพอร์ตเวกเตอร์แมชชีน จากการทดลองรู้จำเครื่องหมายจราจรทั้ง 30 แบบของประเทศไทยที่พบบ่อย พบว่าค่าความถูกต้องของซัพพอร์ตเวกเตอร์แมชชีนมีความแม่นยำ 99.94% สูงกว่าโครงข่ายประสาทเทียมเล็กน้อย 99.84% นอกจากนี้ซัพพอร์ตเวกเตอร์แมชชีนยังใช้เวลาในการทดสอบเฉลี่ยที่ 0.007 วินาทีน้อยกว่าโครงข่ายประสาทเทียมที่ใช้เวลาเฉลี่ยที่ 0.356 วินาที เนื่องจากโครงข่ายประสาทเทียมมีชั้นซ่อนภายในที่มีขนาดใหญ่จึงทำให้ใช้เวลาในการเรียนรู้มากกว่าซัพพอร์ตเวกเตอร์แมชชีนนั้นเอง

เอกสารอ้างอิง

[1] S. Zhu, Y. Zhany and X. Lu, “Detection for Triangle Traffic Sign Based on Neural Network,” Advances in Neural Networks Vol. 3973, pp. 40-45 , 2006.

[2] C. G. Kiran, V. Lekhesh, V. Abdu Rahiman, and K. Rajeev, “Traffic Sign Detection and Pattern Recognition Using Support Vector Machine,” in the 7th International Conference on Advances Pattern Recognition, pp. 87-90, 2009.

[3] W. Woong-Jae, L. Minho and S. Joon-Woo, “Implementation of Road Traffic Signs Detection Based on Saliency Map Model,” IEEE Intelligent Vehicles Symposium, pp. 542-547, 2008.

[4] L. Lanlan and Z. Shuangdong, “Research of Intelligence Classifier for Traffic Sign Recognition,” The 6th International Conference on ITS Telecommunications Proceedings, pp.78-81, 2006.

[5] P. Gil-Jimenez, H. Gomez-Moreno, P. Siegmann, S. Lafuente-Arroyo and S. Maldonado-Bascon, “Traffic Sign Shape Classification Based on Support Vector Machines and the FFT of the Signature of Blob,” in Proceeding of the 2007 IEEE Intelligent Vehicles Symposium, pp. 584-589, 2007.

[6] K. Max, *Computer Graphics and Geometric Modeling : Implementation and Algorithms*, London: Springer, pp. 300-302, 2005.

[7] บุญเสริม กิจศิริกุล,. **ปัญญาประดิษฐ์.** กรุงเทพฯ:ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย,หน้า 169-181, 2546.

[8] C. Cortes and V. Vapnik, “Support Vector Networks,” Machine Learning,Vol. 30, No. 3, pp. 273-297, 1995.