# บทที่ 2

# เอกสารและงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยเรื่อง การรู้จำใบหน้ามนุษย์โดยใช้วิธีวิเคราะห์องค์ประกอบหลักร่วมกับวิธีการวิเคราะห์เชิงภูมิศาสตร์ของใบหน้าและโครงข่ายประสาทเทียมนี้ ได้นำเสนอวิธีการประยุกต์ใช้ทฤษฎีการวิเคราะห์องค์ประกอบหลัก เพื่อลดผลกระทบของสัญญาณรบกวนและลดมิติข้อมูลภาพบนใบหน้า มาทำงานร่วมกับวิธีการวิเคราะห์เชิงภูมิศาสตร์ของใบหน้าและโครงข่ายประสาทเทียม มีเอกสารที่เกี่ยวข้องดังนี้

1. การประมวลผลภาพเชิงตัวเลข (Digital Image Processing)
2. การรู้จำรูปแบบ (Pattern Recognition)
3. ทฤษฎีการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis : PCA)
4. โครงข่ายประสาทเทียม (Artificial Neural Network)
5. วิธีการวิเคราะห์เชิงภูมิศาสตร์ (Geometric Feature-based Analysis)
6. งานวิจัยที่เกี่ยวข้อง

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

## การประมวลผลภาพเชิงตัวเลข (Digital Image Processing)

การประมวลผลภาพเชิงตัวเลข (Digital Image Processing) หมายถึงการนำภาพที่พบทั่วไปมาประมวลผลด้วยเครื่องคอมพิวเตอร์ โดยภาพที่นำมาประมวลผลด้วยเครื่องคอมพิวเตอร์นี้ จะถูกแทนที่ด้วยตัวเลขให้อยู่ในรูปของเมตริกซ์ แต่ภาพที่ได้โดยส่วนมากแล้วจะเป็นภาพที่ได้จากตัวรับสัญญาณ ซึ่งอยู่ในรูปของฟังก์ชัน $f(x,y)$ ที่ต่อเนื่องในระนาบสองมิติ (คือแกน $X$ และแกน $Y$) ดังนั้นข้อมูลภาพที่จะนำไปประมวลผลเชิงตัวเลขนั้นจะต้องทำการแปลงหรือดิจิไตซ์ (Digitization) ฟังก์ชั่นต่อเนื่อง $f(x,y)$ นั้น ให้เป็นฟังก์ชันไม่ต่อเนื่อง $g(x,y)$ เสียก่อน โดยการควอนไทต์ (Quantized) หรือแบ่ง $f(x,y)$ ออกเป็นช่วงๆ ซึ่งผลลัพธ์ที่ได้ก็คือ ค่าความเข้มของจุดภาพที่มีอยู่ $L$ ระดับ รวมทั้งค่าพิกัด $(x,y)$ โดยค่าเหล่านี้จะเป็นจำนวนเต็มที่ไม่เป็นลบ ซึ่งแต่ละจุดภาพนั้นเรียกว่า พิกเซล (Pixel) (Gonzalez และคณะ, 2002)

โดยทั่วไปแล้วภาพหนึ่งๆ ประกอบด้วยพิกเซลหรือจุดภาพมากมาย โดยทั่วไปจุดภาพแต่ละจุดจะเป็นสมาชิกของเมตริกซ์ที่มีขนาด $M$ แถวและ $N$ คอลัมภ์ $(M \times N)$ เพราะฉะนั้น $x, y$ จะมีค่าอยู่ในช่วง $(1 \le y \le M, 1 \le x \le N)$ โดยทั่วไปจุดภาพแต่ละจุดจะเป็นสมาชิกของเมตริกซ์ที่มีขนาด $M$ แถวและ $N$ คอลัมภ์ $(M \times N)$ เพราะฉะนั้น $x, y$ จะมีค่าอยู่ในช่วง $(1 \le y \le M, 1 \le x \le N)$ ซึ่งภาพแต่ละภาพที่สร้างขึ้นจะมีความหนาแน่นของพิกเซลเหล่านี้แตกต่างกันออกไป ความหนาแน่นบ่งบอกถึงความละเอียดของภาพ (Resolution Image) โดยมีหน่วยเป็น PPI (Pixel Per Inch) คือจำนวนจุดต่อนิ้ว ซึ่งภาพโดยทั่วไปที่มีความละเอียดสูงหรือคุณภาพดี ควรจะมีความละเอียด 300 x 300 ppi ขึ้นไป ค่า ppi ยิ่งสูงขึ้น ภาพนั้นจะมีความละเอียดและคมชัดมากขึ้นในแต่ละจุดหรือพิกเซล สำหรับภาพแต่ละภาพนอกจากจะมีความละเอียดหรือความหนาแน่นของจุดภาพแล้วยังมีช่วงระดับความเข้มของจุดภาพ $L$ ซึ่งจะบ่งบอกถึงระดับความละเอียดของภาพเชิงตัวเลข และจะแสดงคุณสมบัติทางสีให้แก่ภาพด้วย โดยแต่จุดจะเป็นตัวสร้างสีประกอบกันเป็นหนึ่งภาพ โดยทั่วไปแล้ว $L$ จะมีค่าความเข้มตั้งแต่ 2 ระดับขึ้นไป แต่ที่ใช้กันมากจะใช้กันที่ค่าระดับความเข้มเท่ากับ 256 ระดับ ซึ่งจะทำให้ค่าของจุดภาพอยู่ในช่วง [0-255] โดยใช้เนื้อที่การเก็บข้อมูลภาพ 1 ไบต์ (byte) หรือ 8 บิท (bit) สำหรับข้อมูล 1 จุดภาพ ($2^8 = 256$) ในกรณีที่ต้องการภาพที่มีความละเอียดของระดับความเข้มสูงๆ ค่า $L$ อาจจะต้องการจำนวนบิทสำหรับการเก็บข้อมูลมากกว่า 8 บิท คืออาจเป็น 16 หรือ 24 บิท โดยค่าระดับความเข้มของจุดภาพจะเท่ากับ $2^{16}$ และ $2^{24}$ ตามลำดับ

โดยทั่วไปจุดภาพแต่ละจุดจะเป็นสมาชิกของเมตริกซ์ที่มีขนาด $M$ แถวและ $N$ คอลัมภ์ $(M \times N)$ เพราะฉะนั้น $x, y$ จะมีค่าอยู่ในช่วง $(1 \le y \le M, 1 \le x \le N)$ และจำนวนช่วงระดับความเข้มของจุดภาพ $L$ จะบ่งบอกถึงระดับความละเอียดของภาพเชิงตัวเลข โดยทั่วไปแล้ว $L$ จะมีค่าความเข้มตั้งแต่ 2 ระดับขึ้นไป แต่ที่ใช้กันมากจะใช้กันที่ค่าระดับความเข้มเท่ากับ 256 ระดับ ซึ่งจะทำให้ค่าของจุดภาพอยู่ในช่วง [0-255] โดยใช้เนื้อที่การเก็บข้อมูลภาพ 1 ไบต์ (byte) หรือ 8 บิท (bit) สำหรับข้อมูล 1 จุดภาพ ($2^8 = 256$) ในกรณีที่ต้องการภาพที่มีความละเอียดของระดับความเข้มสูงๆ ค่า $L$ อาจจะต้องการจำนวนบิทสำหรับการเก็บข้อมูลมากกว่า 8 บิท คืออาจเป็น 16 หรือ 24 บิท โดยค่าระดับความเข้มของจุดภาพจะเท่ากับ $2^{16}$ และ $2^{24}$ ตามลำดับ

โดยทั่วไปวิธีการประมวลผลภาพเชิงตัวเลขที่ทำให้คอมพิวเตอร์สามารถรู้จักวัตถุภายในภาพได้นั้น พอจะแบ่งได้สองระดับด้วยกันคือ การประมวลผลภาพในระดับต่ำ (Low-level Image Processing) และการประมวลผลภาพในระดับสูง (High-level Image Processing)

การประมวลผลภาพระดับต่ำ เป็นการประมวลผลเชิงตัวเลขเกือบทั้งหมด เพื่อหาตัวแปรต่างๆมาอธิบายข้อมูลภาพ โดยมีจุดประสงค์เพื่อนำตัวแปรเหล่านี้ไปใช้ในการประมวลผลระดับสูงในขั้นต่อไป โดยทั่วไปแล้วการประมวลผลภาพระดับต่ำจะประกอบไปด้วย การประมวลผลภาพก่อน (Preprocessing), การกำจัดสัญญาณรบกวนหรือการทำให้ภาพคมชัด, การหาขอบภาพ เป็นต้น

การประมวลผลในระดับสูง เป็นการนำผลลัพธ์หรือสัญลักษณ์ที่ได้จากการประมวลผลระดับต่ำมาตีความหรือประมวลผลเพื่อให้คอมพิวเตอร์สามารถรู้จักและเข้าใจภาพได้

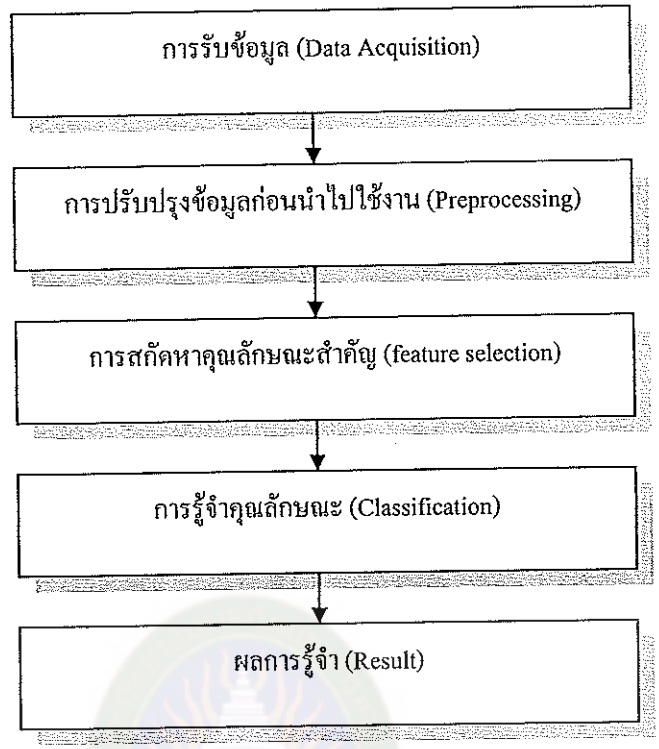
สำหรับความแตกต่างของการประมวลผลภาพระดับต่ำและการประมวลผลภาพระดับสูงนั้นก็คือ ข้อมูลที่นำมาใช้ในการประมวลผลภาพ โดยที่การประมวลผลภาพระดับต่ำจะใช้ค่าความสว่างของจุดภาพโดยตรง ส่วนการประมวลผลภาพในระดับสูงนั้นข้อมูลภาพที่จะนำมาประมวลผลจะถูกแสดงในรูปของสัญลักษณ์ ซึ่งได้จากการใช้ตัวแปรที่ได้จากการประมวลผลในระดับต่ำมาอธิบายถึงสัญลักษณ์เหล่านี้ โดยสัญลักษณ์เหล่านี้จะแสดงถึงสิ่งต่างๆ ที่อยู่ในภาพเช่น ขนาดของวัตถุ รูปร่าง และความสัมพันธ์กันระหว่างวัตถุภายในภาพ

## การรู้จำรูปแบบ (Pattern Recognition)

การรู้จำรูปแบบ คือ ความสามารถในการแยกแยะความแตกต่างของวัตถุ หรือสิ่งของต่างๆได้ เช่น มนุษย์สามารถแยกแยะมนุษย์คนอื่นๆได้ โดยอาศัยประสบการณ์จากการมองเห็นหรือ การฟังเสียง (Friedman และคณะ, 1999)

การรู้จำรูปแบบโดยทั่วไปจะประกอบด้วยขั้นตอนหลายๆขั้นตอน จะเริ่มจากเก็บข้อมูลของวัตถุด้วยเครื่องมือที่เหมาะสม และผ่านกระบวนการปรับปรุงข้อมูลบางอย่างก่อน แล้วจึงค่อยเลือกและแยกลักษณะสำคัญ (Feature Selection & Feature Extraction) ได้ผลเป็นเวกเตอร์ลักษณะสำคัญ (Feature Vectors) หลังจากนั้นจึงนำเวกเตอร์นั้นไปใช้ในการเรียนรู้ (Learning) จนสามารถจัดกลุ่ม (Classification) ตามต้องการหรือรู้จำได้ โดยได้ผลการจัดกลุ่มเป็นผลลัพธ์สุดท้าย โดยขั้นตอนต่างๆแสดงตามแผนภูมิที่ 1

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY




**ภาพที่ 2** แสดงขั้นตอนการรู้จำรูปแบบ

**1. การรับข้อมูล (Data Acquisition)**

การรับข้อมูลเพื่อนำเข้าสู่กระบวนการรู้จำ สามารถทำได้หลายวิธีขึ้นอยู่กับคุณลักษณะของระบบการรู้จำที่จะนำไปใช้งาน ส่วนใหญ่จะเป็นข้อมูลภาพถ่ายบริเวณของคุณลักษณะนั้นๆ เช่น การรู้จำลายมือจะใช้ภาพถ่ายสองมิติบริเวณนิ้วมือที่ต้องการรู้จำ และการรู้จำใบหน้าส่วนมากจะใช้ภาพถ่ายหน้าตรงของใบหน้าแบบสองมิติ เป็นต้น

**2. การปรับปรุงข้อมูลก่อนนำไปใช้งาน (Preprocessing)**

ข้อมูลส่วนใหญ่ของการรู้จำทั่วไป จะเป็นภาพถ่ายบริเวณต่างๆที่สนใจที่จะทำการรู้จำ ดังนั้นการปรับปรุงข้อมูลก่อนนำไปใช้ในการรู้จำก็คือ การประมวลผลภาพเบื้องต้นก่อนการรู้จำ ซึ่งภาพที่ได้จากการรับข้อมูลอาจมีทั้งข้อมูลที่ต้องการรวมกับฉากหลัง หรือข้อมูลนั้นไม่ชัด หรือข้อมูลนั้นมีสัญญานรบกวน ดังนั้นต้องทำการกำจัดสัญญานรบกวนนั้นออกจากภาพก่อน หรือบางครั้งอาจต้องการแบ่งภาพ หรือบางครั้งอาจต้องทำการแปลงภาพเหล่านั้นให้อยู่ในลักษณะที่เหมาะสมต่อการนำไปใช้งาน

มหาวิทยาลัยมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

**3. การสกัดหาคุณลักษณะสำคัญ (Feature Selection)**

ขั้นตอนนี้เป็นขั้นตอนสำหรับหาข้อมูลจากลักษณะที่สำคัญของวัตถุ เพื่อนำมาใช้เป็นตัวแทนของวัตถุนั้นๆ ข้อมูลลักษณะสำคัญที่ต่างกันจะมีผลต่อการเรียนรู้ที่ต่างกัน หรืออาจพูดได้ว่าข้อมูลลักษณะสำคัญแต่ละค่ามีค่าน้ำหนักที่ส่งผลต่อการเรียนรู้แตกต่างกัน ดังนั้นถ้าเลือกข้อมูลลักษณะสำคัญได้ไม่ดีพอ อาจส่งผลต่อการเรียนรู้จนทำให้ไม่สามารถแยกแยะหรือจัดกลุ่มได้ โดยทั่วไปข้อมูลของลักษณะสำคัญที่เลือกนี้จะเก็บในรูปของเวกเตอร์ลักษณะสำคัญ ตัวอย่างเช่น $\mathbf{X} = (x_1, x_2, \dots, x_n)^T$ จะมีข้อมูลลักษณะสำคัญทั้งหมด n ค่า

**4. การจำแนกคุณลักษณะ (Classification)**

เป็นขั้นตอนในการจำแนกและตัดสินใจว่าข้อมูลอินพุต (Input Data) ที่เข้ามานั้นเป็นอะไร โดยในขั้นตอนนี้มีหลายวิธีเช่น การเปรียบเทียบกับข้อมูลในฐานข้อมูล การเปรียบเทียบข้อมูลอินพุตกับกฎในการตัดสินใจ การใช้โครงข่ายประสาทเทียม เป็นต้น

จะเห็นได้ว่าการรู้จำรูปแบบเป็นพื้นฐานของการรู้จำวัตถุ ซึ่งจะมีขั้นตอนหลักๆ อยู่ 4 ขั้นตอน ซึ่งขั้นตอนเหล่านี้จะถูกนำไปเป็นต้นแบบในการรู้จำอื่นๆต่อไป

## วิธีการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis : PCA)

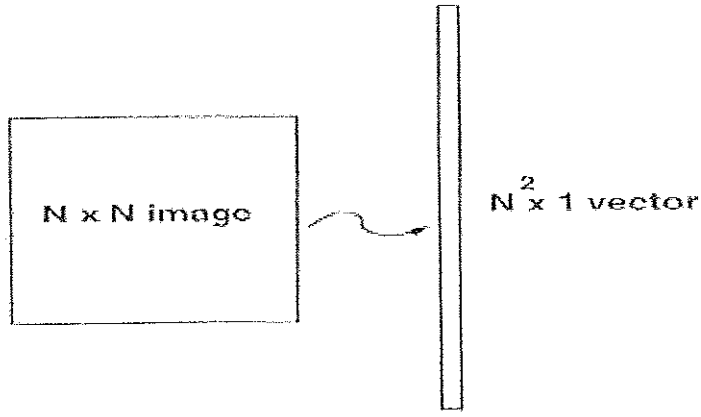
วิธีการวิเคราะห์องค์ประกอบหลัก เป็นวิธีการประมวลผลภาพเพื่อลดมิติข้อมูลภาพให้มีขนาดของค่าของข้อมูลลดลง เพื่อให้ได้ภาพใหม่ที่มีขนาดมิติลดลง และสามารถนำไปประมวลผลได้รวดเร็วกว่าภาพต้นฉบับ หลังจากใช้วิธีการวิเคราะห์องค์ประกอบหลักข้อมูลบางส่วนจะหายไปและจะเหลือเฉพาะข้อมูลที่สำคัญที่จำเป็นต่อการประมวลผลเท่านั้น

วิธีการวิเคราะห์องค์ประกอบหลัก เป็นวิธีการทางสถิติซึ่งถูกนำไปประยุกต์ใช้ในงานต่างๆ อาทิเช่น การบีบอัดข้อมูล (Data Compression) การสร้างภาพใบหน้าไอเกน เพื่อใช้ในระบบการรู้จำ เป็นต้น โดยสิโรวิชและเคอร์บี้ (Sirovich and Kirby) ได้เป็นบุคคลแรกที่ได้นำวิธีนี้มาใช้ในการรู้จำใบหน้ามนุษย์ เป็นครั้งแรกโดย ในปี 1987 และต่อมาเติร์กและเพนท์แลนด์ (Turk and Pentland) ได้สร้างระบบรู้จำใบหน้า ที่มีชื่อว่า ใบหน้าไอเกน (Eigenface) โดยการนำภาพใบหน้าไอเกนไปทำการรู้จำ

**1. การหาค่าไอเกนและเวกเตอร์ไอเกน**

กำหนดให้ I(x,y) คือ ภาพที่ใช้เป็นภาพ 2 มิติ มีขนาด 80 x 80 จุดภาพ (NxN) ชนิดสีเทา เมื่อทำเป็นรูปภาพแบบ 1 มิติ $N^2$ x 1 (1-D Vector) มีความยาวเท่ากับ 6,400 จุดต่อ 1 ภาพ เพื่อทำเป็นภาพใบหน้าชุดฝึกเรียนรู้ (Training Set Image) จากภาพใบหน้าต้นฉบับ

มหาวิทยาลัยราชภัฏมหาสารคาม RAJABHAT MAHASARAKHAM UNIVERSITY




**ภาพที่ 3** ภาพ 2 มิติขนาด (N x N) เมื่อทำเป็นรูปภาพแบบ 1 มิติ $N^2$x 1 (1-D Vector)

โดยกำหนดให้ $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_M$ เป็นชุดข้อมูลของภาพใบหน้า การสร้างชุดข้อมูลจะคำนวณค่าเฉลี่ย ค่าโควาเรียนเมตริกซ์และคำนวณค่าไอเกนเว็กเตอร์ตามลำดับ โดยมีขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 การคำนวณหาค่าเฉลี่ย ($\Psi$ =Average Face) จากทุกใบหน้า ดังสมการที่ 2.1

$$\Psi = \frac{1}{M}\sum_{1}^{M}\Gamma_i \quad \text{.....2.1}$$

เมื่อ $\Psi$ คือ ใบหน้าเฉลี่ย

$M$ คือ จำนวนใบหน้า

$\Gamma$ คือ กลุ่มข้อมูลใบหน้า

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

ขั้นตอนที่ 2 คำนวณหาผลต่างระหว่างชุดฝึกเรียนรู้และค่าเฉลี่ย หรือเรียกว่าค่าเบี่ยงเบนมาตรฐานของภาพใบหน้า ดังสมการที่ 2.2

$$\Phi_i = \Gamma_i - \Psi \quad \text{.....2.2}$$

เมื่อ $\Phi$ คือ ค่าเบี่ยงเบนมาตรฐานของภาพใบหน้า

ขั้นตอนที่ 3 คำนวณหาค่าเมตริกซ์ความแปรปรวนร่วมหรือโควาเรียนเมตริกซ์ โดยนำค่าเบี่ยงเบนมาตรฐานจากขั้นตอนที่ 2 คูณกับค่าเบี่ยงเบนมาตรฐานที่ทำการทรานโพส ดังสมการที่ 2.3

$$C = \frac{1}{M}\sum_{1}^{M} \Phi_i \Phi_i^T = AA^T \quad \text{.....2.3}$$

เมื่อ $A = [\Phi_1, \Phi_2, \ldots, \Phi_M]$ เป็นชุดข้อมูลค่าเบี่ยงเบนมาตรฐาน

$C$ คือ เมตริกซ์ความแปรปรวนร่วม

A คือ กลุ่มของเมตริกซ์ความแปรปรวนร่วม

ขั้นตอนที่ 4 คำนวณหาไอเกนเวคเตอร์ ($v$) และค่าไอเกน($\mu$) ของเมตริกซ์ความแปรปรวนร่วม เนื่องจากเมตริกซ์ C มีขนาด $N^2$ x $N^2$ มิติ นำมาสร้างเป็นไอเกนเวคเตอร์ขนาด $N^2$ และค่าไอเกน ขนาด $N^2$ ด้วยรูปภาพ ขนาด 80 x 80 พิกเซล จะไม่สามารถกระทำได้ ซึ่ง $N^2$ จะได้เมตริกซ์ที่มีขนาดใหญ่มาก ซึ่งวิธีวิเคราะห์องค์ประกอบหลักจะเป็นการลดมิติข้อมูลลง จึงทำให้ได้ผลลัพธ์มีขนาดมากเกินไป และอาจจะมีไอเกนเวกเตอร์บางค่าที่สัมพันธ์กับค่าไอเกนที่มีค่าเป็นศูนย์ จึงไม่นำมาเป็นข้อมูลในการพิจารณา เพื่อแก้ปัญหานี้เติร์กและเพนท์แลนด์ ได้ทำการปรับปรุงขั้นตอนวิธีโดยการเสนอวิธีใหม่ โดยการสร้างเมตริกซ์ $A^TA$ แทนที่จะใช้ $AA^T$ ดังนั้นค่าไอเกนเวคเตอร์ สามารถทำได้ ดังสมการที่ 2.4

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

$$A^T A v_i = \mu_i v_i \quad \text{.....2.4}$$

เมื่อ $A$ คือ กลุ่มของเมตริกซ์ความแปรปรวนร่วม

$\mu_i$ คือ ค่าไอเกน (Eigen value)

$v_i$ คือ เวกเตอร์ไอเกน (Eigen vector)

ขนาดของเมตริกซ์ความแปรปรวนร่วม จะมีขนาด M x M จะได้ค่าไอเกนใหญ่มาก จึงแทนค่าไอเกนเป็นขนาด $N^2$ ซึ่งจะได้สมการใหม่ ดังสมการที่ 2.5

$$AA^T A v_i = \mu_i A v_i \quad \text{.....2.5}$$

ในทั้ง 4 ขั้นตอนที่กล่าวมานี้เป็นวิธีการหาผลลัพธ์จากวิธีวิเคราะห์องค์ประกอบหลัก จะได้ค่าไอเกน ($\mu_i$) และ เวกเตอร์ไอเกน ($\nu_i$) ซึ่งข้อมูลทั้งสองมีความสมนัย (Correspondence) ซึ่งกันและกัน แล้วนำผลลัพธ์ที่ได้ไปคัดเลือกเพื่อเข้าสู่กระบวนการรู้จำต่อไป

**2. การเลือกจำนวนมิติข้อมูลเพื่อนำไปรู้จำ**

เมื่อนำค่าไอเกนที่มีความสมนัยกับเวกเตอร์ไอเกนนั้นๆเรียงข้อมูลจากมากไปหาน้อย จะได้ว่าค่าไอเกนลำดับแรกๆจะมีค่าไอเกนมากกว่าลำดับต่อๆมามาก และการลดลงของค่าไอเกนเมื่อเกินองค์ประกอบที่ 30 จะมีค่าประมาณศูนย์ทั้งหมด

ในขั้นตอนนี้จะใช้วิเคราะห์องค์ประกอบหลักโดยใช้หน้าไอเกนที่มีนัยสำคัญที่สุด ดังนั้นเพื่อลดพื้นที่ในการเก็บข้อมูล โดยการเลือกเอาเฉพาะใบหน้าไอเกนแรกๆ มาเก็บไว้เท่านั้น

ขั้นตอนที่ 5 วิธีการรู้จำ เริ่มจากการนำภาพที่ต้องการทดสอบการรู้จำ โดยนำภาพใบหน้าไอเกนไปทำการหาค่าน้ำหนัก โดยนำภาพใบหน้าไอเกนที่เลือกคูณกับภาพใบหน้า ซึ่งเป็นการฉายภาพ (Projection) ภาพใบหน้ากับภาพใบหน้าไอเกนในปริภูมิภาพใบหน้า (Face space) ดังสมการที่ 2.6

$$\varpi_k = \mu_k^T(\Gamma - \Psi) \quad \text{.....2.6}$$

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

เมื่อ $\omega$ คือค่าน้ำหนัก

$\mu_i$ คือ ค่าไอเกน

k คือจำนวนใบหน้าที่เลือก ซึ่งเป็นนำภาพใบหน้ามาคำนวณเพื่อหาค่าน้ำหนักที่จะนำไปเป็นตัวแทนจำแนกหรือรู้จำบุคคลต่อไป ซึ่งจะได้ $\Omega^T = [\omega_1, \omega_2, \omega_3, \ldots, \omega_M]$

ขั้นตอนที่ 6 คำนวนหาค่าระยะห่างระหว่างภาพใบหน้าที่นำมาทดสอบกับภาพใบหน้าไอเกน (Euclidean distance) ดังสมการที่ 2-7

$$\varepsilon_k = \left\| \Omega - \Omega_k \right\| \quad \text{.....2.7}$$

เมื่อ $\varepsilon_k$ คือ ค่าระยะห่าง

$\Omega$ คือ ใบหน้าไอเกน

$\Omega_k$ คือ ใบหน้าทดสอบ

จากสมการ 2.7 ค่า $\Omega_k$ คือระยะห่างของใบหน้าลำดับที่ k ที่สามารถค้นพบ หลังจากมีการคำนวณภาพระยะหว่างภาพใบหน้ากลุ่มเรียนรู้กับภาพใบหน้าที่นำมาทดสอบแล้ว นำค่าระยะห่างมาวิเคราะห์ โดยพิจารณาว่าข้อมูลของภาพที่นำมาทดสอบมีค่าใกล้เคียงกับภาพใบหน้าในกลุ่มเรียนรู้ภาพใด มีระยะห่างน้อยที่สุด สรุปผลออกมาได้ว่า ภาพที่นำมาทดสอบนั้นอยู่ กลุ่มเดียวกับภาพในกลุ่มเรียนรู้ภาพนั้น

จะเห็นได้ว่าวิธีการวิเคราะห์องค์ประกอบหลัก เป็นวิธีการหนึ่งในการรู้จำภาพที่มีข้อเด่นก็คือ การลดมิติข้อมูลภาพให้มีขนาดของค่าของข้อมูลลดลง และสามารถนำไปประมวลผลได้รวดเร็วกว่าภาพต้นฉบับ และเมื่อนำไปใช้ทดลองกับภาพเพื่อทำการรู้จำแล้วพบว่า ส่วนใหญ่ให้ผลการรู้จำที่ถูกต้องสูงมาก

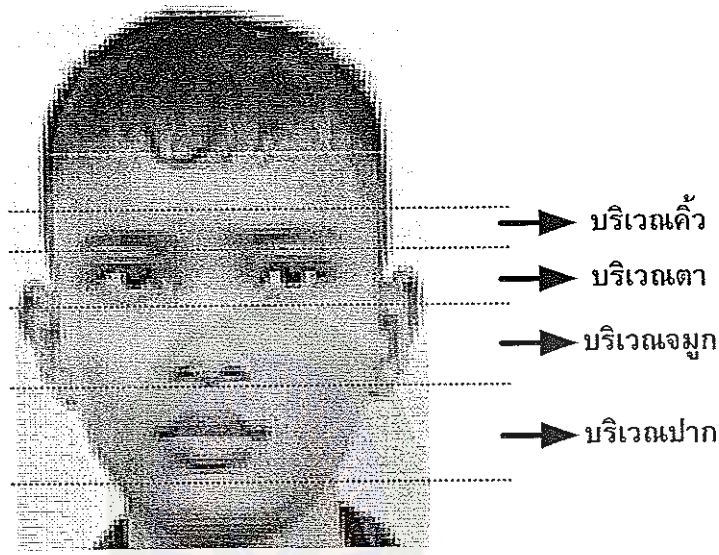
## วิธีการวิเคราะห์เชิงภูมิศาสตร์ (Geometric Feature-based Analysis)

วิธีการวิเคราะห์เชิงภูมิศาสตร์ เป็นวิธีการหนึ่งที่ใช้ในการแยกคุณลักษณะที่สำคัญออกจากวัตถุที่ต้องการนำไปสู่กระบวนการรู้จำ ซึ่งการรู้จำรูปแบบแต่ละชนิดนั้นมีความแตกต่างกันตามรูปแบบของอวัยวะที่นำมาใช้งาน และมีขั้นตอนวิธีในการสกัดหาคุณลักษณะเฉพาะเหล่านั้นหลายวิธี ซึ่งวิธีการวิเคราะห์เชิงภูมิศาสตร์เป็นตัวอย่างหนึ่งของการสกัดหาคุณลักษณะเฉพาะของใบหน้า ซึ่งโดยทั่วไปเป็นการระบุตำแหน่งของอวัยวะบนใบหน้า เช่นตำแหน่งของตา ตำแหน่งของปาก ตำแหน่งของจมูก เป็นต้น

สำหรับการวิเคราะห์เชิงภูมิศาสตร์ที่ใช้ในวิทยานิพนธ์นี้ ได้ใช้หลักการที่ว่า ใบหน้าของมนุษย์มีความแตกต่างกันโดยธรรมชาติ โดยส่วนที่สามารถสังเกตได้ชัดคือ คิ้ว ตา จมูก และปาก ซึ่งถ้าหากเราสามารถทำให้คอมพิวเตอร์สามารถจดจำ 4 ส่วนนี้ได้ ถึงแม้บุคคลนั้นจะเปลี่ยนแปลงทรงผม อารมณ์ หรือส่วนประกอบอื่นๆ คอมพิวเตอร์ก็ยังจะสามารถแยกแยะใบหน้าที่จดจำไว้ได้อย่างถูกต้อง ดังนั้นผู้วิจัยจึงได้เลือกใช้ 4 ส่วนนี้ เพื่อใช้เป็นคุณลักษณะที่สำคัญในการรู้จำใบหน้ามนุษย์ ดังนี้

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

1. บริเวณคิ้ว หมายถึงบริเวณตั้งแต่เหนือคิ้วลงมา จนถึงใต้คิ้ว
2. บริเวณตา หมายถึงบริเวณตั้งแต่เหนือตาลงมา จนถึงใต้ตา
3. บริเวณจมูก หมายถึงบริเวณตั้งแต่ใต้ตา จนถึงขอบปากด้านบน
4. บริเวณปาก หมายถึงบริเวณตั้งแต่ริมฝีปากด้านบน จนถึงริมฝีปากด้านล่าง




**ภาพที่ 4** แสดงบริเวณต่างๆบนใบหน้าที่ได้จากการแยกแยะองค์ประกอบ ตามวิธีการวิเคราะห์เชิงภูมิศาสตร์

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

**โครงข่ายประสาทเทียม (Artificial Neural Network)**

โครงข่ายประสาทเทียม (Artificial Neural Network) เป็นระบบประมวลผลข้อมูลซึ่งเป็นรูปแบบหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence) ที่พยายามเลียนแบบการทำงานของสมองมนุษย์ เป็นแนวความคิดที่ต้องการให้คอมพิวเตอร์มีความชาญฉลาดในการเรียนรู้เหมือนที่มนุษย์มีการเรียนรู้ สามารถฝึกฝนได้ และสามารถนำความรู้และทักษะไปแก้ปัญหาต่าง ๆ มีนักวิจัยจำนวนมากได้คิดค้นรูปแบบโครงข่ายประสาทเทียมแบบต่าง ๆ เพื่อนำมาประยุกต์ใช้อย่างกว้างขวาง การประยุกต์ใช้งานโครงข่ายประสาทเทียมมีตั้งแต่การใช้เพื่อตัดสินใจง่ายไปจนถึงงานที่มีความยุ่งยากซับซ้อน ตัวอย่างการประยุกต์ใช้งานบางส่วน ได้แก่ งานด้านการควบคุม งานด้านการบิน ด้านยานยนต์ ด้านการบริหารจัดการ ด้านการธนาคาร ด้านการทหาร ด้านการบันเทิง และอื่น ๆ อีกมากมาย

โครงข่ายประสาทเทียมเป็นการจำลองการทำงานโครงข่ายประสาทของมนุษย์ (Biological Neurons) ซึ่งประกอบด้วยส่วนของการประมวลผลที่เรียกว่า นิวรอน (Neuron) ทุกๆ นิวรอนสามารถมีอินพุตได้หลายอินพุตแต่มีเอาต์พุตเพียงเอาต์พุตเดียว และทุก ๆ เอาต์พุตจะแยกไปยังอินพุตของนิวรอนอื่น ๆ ภายในโครงข่าย การติดต่อกัน ภายในระหว่างนิวรอนไม่ใช่ลักษณะการต่อแบบธรรมดาทุก ๆ อินพุตจะมีน้ำหนักเป็นตัวกำหนดกำลังของการติดต่อภายในและช่วยในการตัดสินใจ การทำงานของนิวรอนในบางโครงข่ายจะถูกกำหนดไว้ตายตัว แต่บางโครงข่ายสามารถที่จะปรับแต่งได้ซึ่งอาจจะเป็นการปรับแต่งจากภายนอกโครงข่ายหรือนิวรอนสามารถปรับได้ด้วยตัวของมันเอง ในจุดนี้แสดงถึงความสามารถในการเรียนรู้และจดจำของโครงข่ายประสาทเทียม

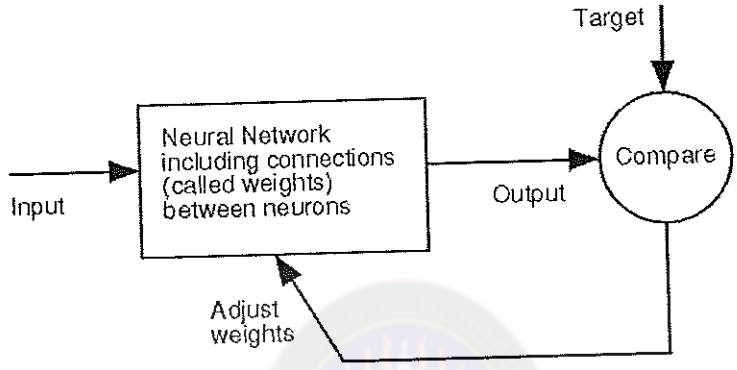
โครงข่ายประสาทเทียมเป็นแบบจำลองการทำงานของระบบประสาทส่วนกลางที่มีโครงสร้างเป็นลักษณะของโครงข่ายเชื่อมโยงกันระหว่างหน่วย ซึ่งสามารถที่จะรับรู้ข้อมูลและปรับตัวเข้ากับสถานการณ์หรือสิ่งแวดล้อมที่กำลังเผชิญอยู่ นักวิจัยหลายท่านเชื่อว่า แบบจำลองโครงข่ายประสาทเทียมเป็นเครื่องมือชนิดหนึ่งที่ใช้ในการสร้างระบบคอมพิวเตอร์อัจฉริยะอย่างได้ผล (Intelligent Computer System) นอกจากนี้ โครงข่ายประสาทเทียม ยังเป็นเครื่องมือที่มีประสิทธิภาพในการประยุกต์ใช้ในงานต่างๆ ที่เกี่ยวข้องกับการคำนวณและการจดจำ เช่น การจำแนกข้อมูล (Data classification) การทำนายเหตุการณ์ (Forecasting) การบีบอัดข้อมูล (Data Compression) การกรองสัญญาณรบกวน (Noise filter) เป็นต้น ในทางคณิตศาสตร์โครงข่ายประสาทเทียม อาจมองในแง่ของการเป็น Universal approximator เนื่องจากความสามารถในการกำหนดความสัมพันธ์ระหว่างรูปแบบของข้อมูล Input-Output ซึ่งทำให้สามารถแก้ปัญหาที่ยากและสลับซับซ้อนได้

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

**1. แนวคิดเบื้องต้นโครงข่ายประสาทเทียม**

สำหรับหลักการทำงานเบื้องต้นของโครงข่ายประสาทเทียม ได้แสดงไว้ดังภาพที่ 5 ซึ่งสามารถอธิบายได้ว่า ในการทำงานของโครงข่ายประสาทเทียมจะออกแบบตามแบบจำลองทางคณิตศาสตร์ และทำงานโดยสร้างการเชื่อมต่อระหว่างหน่วยที่ใช้ประมวลผล (Processing Element) ซึ่งทำหน้าที่เสมือน เซลล์ประสาท (Neuron) การประมวลผลข้อมูลของโครงข่ายประสาทเทียมจะเกิดขึ้นที่หน่วยพื้นฐานจำนวนมากที่เรียกว่า Neuron, Cell, Node, หรือ Unit และสัญญาณข้อมูลจะส่งผ่านกันระหว่างเซลล์ประสาทได้โดยเส้นเชื่อมต่อ (Connection Links) ระหว่างเซลล์ประสาท ซึ่งแต่ละเส้นเชื่อมต่อจะได้รับการกำหนดค่าน้ำหนัก (Weight) ไว้เพื่อจะนำไปใช้คูณกับสัญญาณข้อมูลที่ส่งผ่านมาตามเส้นเชื่อมต่อนั้นๆ จากนั้นแต่ละเซลล์ประสาทจะ

ใช้ ฟังก์ชันกระตุ้น (Activation Function) หรือเรียกว่า ฟังก์ชันถ่ายโอน (Transfer Function) กับสัญญาณข้อมูลที่ส่งเข้ามาเพื่อคำนวณค่าสัญญาณผลลัพธ์ออกไป เมื่อโครงข่ายประสาทเทียมเริ่มเรียนรู้แล้ว โครงข่ายจะพยายามปรับหรือฝึกฝนเพื่อลดค่าความผิดพลาดระหว่างผลลัพธ์ที่ได้กับค่าเป้าหมาย (Target) ลงเรื่อยๆ




**ภาพที่ 5** หลักการทำงานของโครงข่ายประสาทเทียม

**2. แบบจำลองเซลล์ประสาท (Neuron Model)**

โครงข่ายประสาทเทียมที่มีค่าข้อมูลแบบหลายข้อมูลอินพุตในรูปเวกเตอร์ $p = [p_1\ p_2 \cdots p_R]^T$ มีข้อมูลอินพุต R ค่า ดังแสดงในภาพที่ 6 จากนั้นจะนำแต่ละค่าข้อมูลอินพุตแต่ละตัวถูกคูณด้วยเวกเตอร์น้ำหนัก $W = [w_{11}\ w_{12} \cdots w_{1R}]$ แล้วรวมค่าทั้งหมดกับค่าไบแอสเป็นข้อมูลเอาต์พุต n

สำหรับค่าไบแอสเป็นการเพิ่มค่าให้กับผลรวมของผลคูณค่าป้อนเข้ากับค่าน้ำหนัก หรือเป็นการเลื่อนฟังก์ชัน $f$ ไปทางซ้ายเป็นจำนวน $b$ ค่าไบแอสนี้เป็นเหมือนค่าน้ำหนักโดยมีค่าป้อนเข้าเป็น 1 คงที่เสมอ

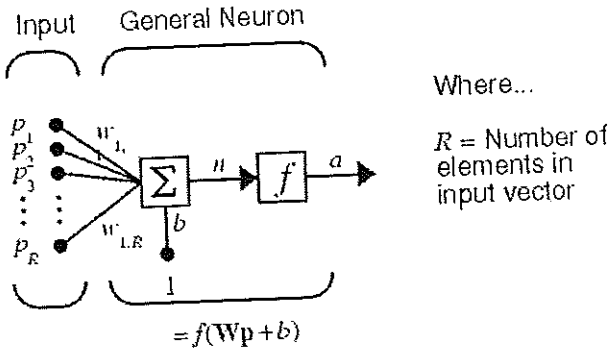
$$n = w_{1,1}p_1 + w_{1,2}p_2 + \ldots + w_{1,R}p_R + b$$

หลังจากนั้นค่า $n$ จะถูกป้อนเข้าฟังก์ชันถ่ายโอน $f$ แล้วให้ค่าข้อมูลเอาท์พุตออกมา

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

สำนักวิทยบริการฯ มหาวิทยาลัยราชภัฏมหาสารคาม





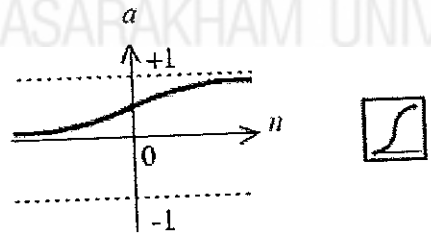

**ภาพที่ 6** แบบจำลองเซลล์ประสาท

**3. ฟังก์ชันกระตุ้น (Activation Function) หรือ ฟังก์ชันถ่ายโอน (Transfer Function)**

เป็นฟังก์ชันที่เซลล์ประสาทใช้กำหนดค่าข้อมูลอินพุตไปเป็นค่าข้อมูลเอาท์พุต ซึ่งค่าข้อมูลเอาท์พุตที่ได้นี้ อาจกลายเป็นค่าข้อมูลอินพุตของเซลล์ประสาทตัวอื่น หรือเป็นค่าข้อมูลเอาท์พุตของระบบโครงข่าย ฟังก์ชันถ่ายโอนนั้นมีใช้กันอยู่หลายชนิด แต่ในที่นี้จะขอกล่าวถึงเพียงแค่ 3 ชนิด คือ

3.1 ฟังก์ชันถ่ายโอนแบบล็อกซิกมอยด์ (Log-Sigmoid Transfer Function) จะรับค่าข้อมูลอินพุตเป็นจำนวนจริงใดๆ แล้วให้ค่าข้อมูลเอาท์พุตเป็นค่าตั้งแต่ 0 ถึง 1

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY



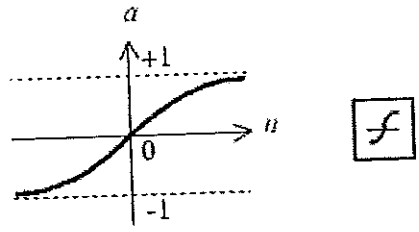

$a = logsig(n)$

Log-Sigmoid Transfer Function

**ภาพที่ 7** ฟังก์ชันถ่ายโอนแบบล็อกซิกมอยด์

196525

2.2 ฟังก์ชันถ่ายโอนแบบแทนซิกมอยด์ (Hyperbolic tangent Sigmoid Transfer Function :Tan- Sigmoid Transfer Function) จะรับค่าข้อมูลอินพุตเป็นจำนวนจริงใดๆ แล้วให้ค่าข้อมูลเอาท์พุตเป็นค่าตั้งแต่ -1 ถึง 1

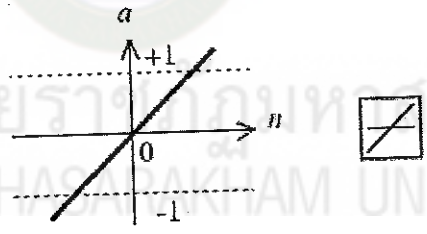



$a = tansig(n)$

Tan-Sigmoid Transfer Function

**ภาพที่ 8** ฟังก์ชันถ่ายโอนแบบแทนซิกมอยด์

2.3 ฟังก์ชันถ่ายโอนแบบเชิงเส้น (Linear Transfer Function) จะรับค่าข้อมูลอินพุตเป็นจำนวนจริงใดๆ แล้วให้ค่าข้อมูลเอาท์พุตเท่ากับค่าข้อมูลอินพุต




$a = purelin(n)$

Linear Transfer Function

**ภาพที่ 9** ฟังก์ชันถ่ายโอนแบบเชิงเส้น

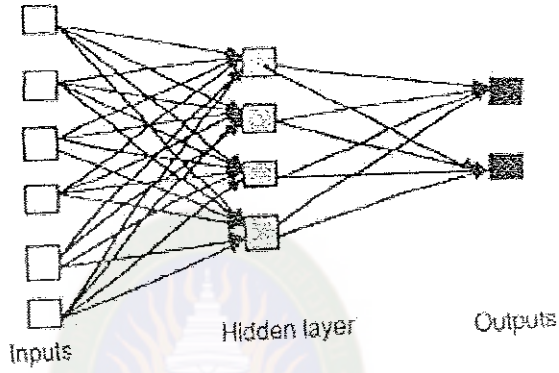
มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

**4. สถาปัตยกรรมของโครงข่ายประสาทเทียม (Architecture of Neural Networks)**

โครงข่ายประสาทเทียมมีโครงสร้างที่ประกอบด้วยเซลล์ประสาทจำนวนมาก รูปแบบการเชื่อมต่อกันระหว่างเซลล์ประสาทนั้นขึ้นอยู่กับสถาปัตยกรรมที่เลือกใช้ โดยสถาปัตยกรรมของโครงข่ายสามารถแบ่งได้เป็น 2 ประเภท คือ

4.1 โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า (Feedforward Neural Network)

สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้านั้นสัญญาณข้อมูลจะวิ่งผ่านไปได้แค่ทางเดียวเท่านั้น คือ จากทิศที่ข้อมูลถูกป้อนไปสู่ผลลัพธ์ โดยไม่มีการวนย้อนกลับและผลลัพธ์ของชั้นใดๆจะไม่มีส่งผลกระทบภายในชั้นเดียวกัน

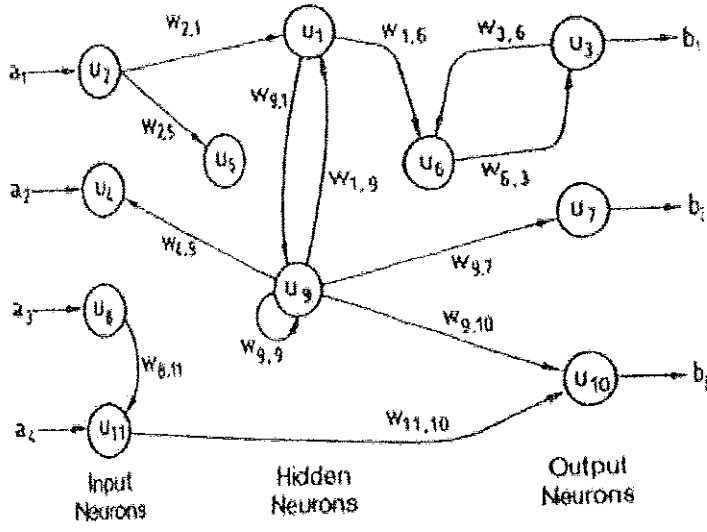



**ภาพที่ 10** สถาปัตยกรรมโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า

4.2 โครงข่ายประสาทเทียมแบบป้อนกลับ (Feedback Neural Network)

สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบป้อนกลับนั้นสัญญาณข้อมูลสามารถวิ่งผ่านได้ทั้งสองทิศทาง ซึ่งทำให้ในโครงข่ายประสาทเทียมสามารถเกิดวงย้อนกลับ (Loop) ได้ สถาปัตยกรรมโครงข่ายประสาทเทียมแบบนี้มีลักษณะเป็นแบบพลวัต (Dynamic) โดยสถานะ (State) ของโครงข่ายจะเปลี่ยนแปลงตลอดจนกว่าจะถึงจุดสมดุล และคงอยู่ที่จุดสมดุลนั้นจนกว่าข้อมูลที่ป้อนเข้ามาจะเปลี่ยนไป แล้วจึงปรับเข้าหาสมดุลใหม่(มีลักษณะเป็นแบบ Interactive หรือ Recurrent) ดังนั้นสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบนี้จึงมีสมรรถภาพสูงและสามารถรองรับระบบที่ซับซ้อนได้ดี

มหาวิทยาลัยราชภัฏพระนครศรีอยุธยา
PHRANAKHON SI AYUTTHAYA RAJABHAT UNIVERSITY




ภาพที่ 11 สถาปัตยกรรมโครงข่ายประสาทเทียมแบบป้อนกลับ

**5. กฎการเรียนรู้(Learning Rule)**

กฎการเรียนรู้เป็นกระบวนการที่ใช้ในการเปลี่ยนแปลงค่าน้ำหนักของการเชื่อมต่อ ทำให้โครงข่ายประสาทเทียมสามารถเรียนรู้วิธีการแก้ปัญหาได้ โดยค่าน้ำหนักการเชื่อมต่อจะถูกปรับเมื่อมีการเรียนรู้ความรู้ใหม่เกิดขึ้น กระบวนนี้ยังอาจเรียกอีกอย่างว่า อัลกอริทึมในการฝึกฝน (Training Algorithm) กฎการเรียนรู้นี้สามารถแบ่งได้เป็น 2 ประเภท คือ

5.1 การเรียนรู้แบบมีการสอน (Supervised Learning)

เป็นการเรียนแบบที่มีการตรวจคำตอบเพื่อให้โครงข่ายประสาทเทียมปรับตัว ชุดข้อมูลที่ใช้สอนโครงข่ายประสาทเทียมจะมีคำตอบไว้คอยตรวจดูว่าโครงข่ายประสาทเทียมให้คำตอบที่ถูกหรือไม่ ถ้าตอบไม่ถูก โครงข่ายประสาทเทียมก็จะปรับตัวเองเพื่อให้ได้คำตอบที่ดีขึ้น (เปรียบเทียบกับคน เหมือนกับการ สอนนักเรียนโดยมีครูผู้สอนคอยแนะนำ) โดยกฎการเรียนรู้แบบนี้จะใช้กลุ่มข้อมูลตัวอย่างมาใช้ในการฝึกฝนโครงข่าย โดยกลุ่มข้อมูลที่ใช้ในการฝึกเรียนรู้ (Training Set) นี้ จะประกอบด้วยข้อมูลอินพุตและข้อมูลเป้าหมาย (Target) ที่ต้องการ เมื่อโครงข่ายได้รับข้อมูลอินพุตและให้ข้อมูลเอาท์พุตแล้ว โครงข่ายจะเปรียบเทียบข้อมูลเอาท์พุตกับข้อมูลเป้าหมายแล้วปรับค่าน้ำหนักใหม่เพื่อให้เอาท์พุตในครั้งต่อไปเข้าใกล้ค่าเป้าหมายมากขึ้น

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

5.2 การเรียนรู้แบบไม่มีการสอน (Unsupervised Learning)

เป็นการเรียนแบบไม่มีผู้แนะนำ ไม่มีการตรวจคำตอบว่าถูกหรือผิด โครงข่ายประสาทเทียมจะจัดเรียงโครงสร้างด้วยตัวเองตามลักษณะของข้อมูล ผลลัพธ์ที่ได้ โครงข่ายประสาทเทียมจะสามารถจัดหมวดหมู่ของข้อมูลได้ (เปรียบเทียบกับคน เช่น การที่เราสามารถแยกแยะพันธุ์พืช พันธุ์สัตว์ตามลักษณะ รูปร่างของมันได้เองโดยไม่มีใครสอน) โดยกฎการเรียนรู้แบบนี้จะใช้แต่ข้อมูลอินพุต ไม่ใช้ข้อมูลเป้าหมาย โครงข่ายจะปรับค่าน้ำหนักเพื่อให้ข้อมูลอินพุตที่ใกล้เคียงกันหรือมีรูปแบบคล้ายกัน แสดงค่าข้อมูลเอาท์พุตออกมาเหมือนกัน

**6. โครงข่ายประสาทเทียมหลายชั้นแบบป้อนไปข้างหน้า (Multilayer Feedforward Neural Network)**

โครงข่ายนี้มีสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า และประกอบด้วยเซลล์ประสาทเรียงตัวกันเป็นชั้น(layer) โดยในแต่ละโครงข่ายจะประกอบด้วยชั้นประเภทต่างๆ 3 ประเภท คือ

6.1 ชั้นอินพุต(Input Layer)

ประกอบด้วยโหนดที่มีจำนวนเท่ากับจำนวนของข้อมูลอินพุต ทำหน้าที่กระจายสัญญาณข้อมูลอินพุตของโครงข่ายไปสู่โหนดหรือเซลล์ประสาทในระดับชั้นถัดไป โดยไม่มีการเปลี่ยนแปลงใดเกิดขึ้นกับสัญญาณที่ผ่านเข้ามา

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

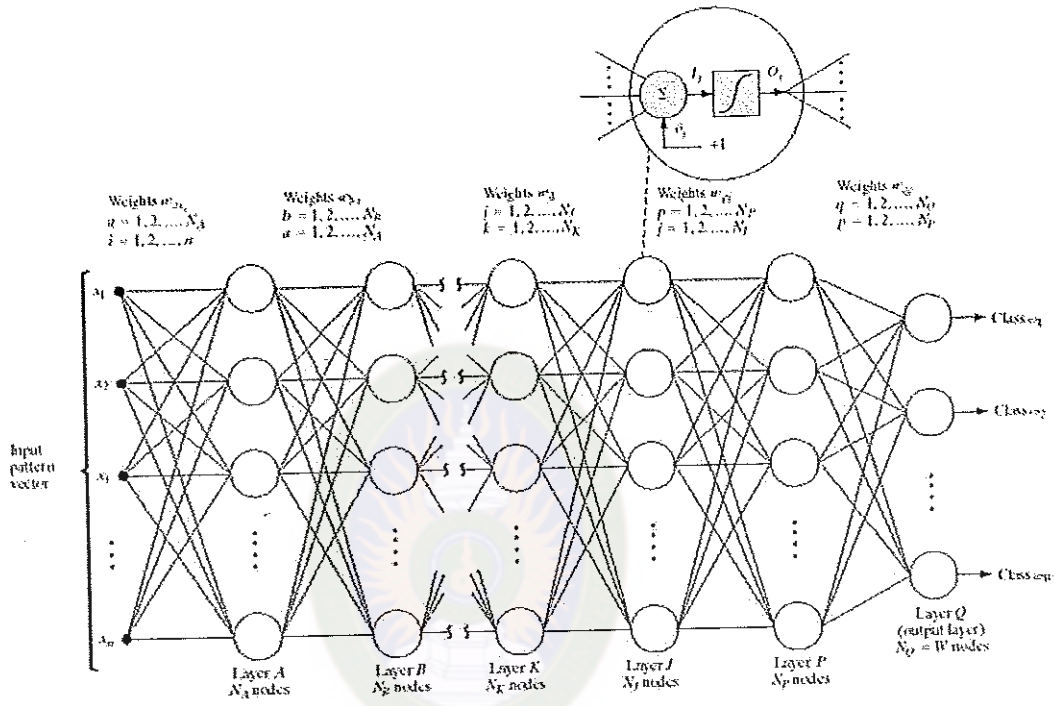
6.2 ชั้นแฝง(Hidden Layer)

ประกอบด้วยโหนดที่ทำหน้าที่เป็นเซลล์ประสาท ซึ่งจะรับสัญญาณข้อมูลที่มาจากโหนดในชั้นอินพุตหรือจากโหนดในชั้นแฝงอื่น แล้วคำนวณค่าสัญญาณใหม่ส่งเป็นเอาท์พุตไปสู่โหนดในชั้นเอาท์พุตหรือโหนดในชั้นแฝงอื่นซึ่งอยู่ในระดับชั้นถัดไป

6.3 ชั้นเอาท์พุต(Output Layer)

ประกอบด้วยโหนดที่ทำหน้าที่เป็นเซลล์ประสาท ซึ่งจะทำหน้าที่เหมือนเซลล์ประสาทในชั้นแฝงทุกประการ เอาท์พุตที่ได้ในชั้นนี้จะเป็นเอาท์พุตของโครงข่ายทั้งระบบ

ในการระบุลักษณะของโครงข่ายว่ามีจำนวนชั้นเท่าไหร่ ให้ระบุตามจำนวนชั้นรวมของชั้นแฝงกับชั้นเอาท์พุตเท่านั้น ตัวอย่างเช่น โครงข่ายแบบ 3 ชั้น จะประกอบด้วยชั้นแฝง 2 ชั้นและชั้นเอาท์พุต




**ภาพที่ 12** โครงข่ายประสาทเทียมหลายชั้นแบบป้อนไปข้างหน้า

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

### 7. การฝึกฝนแบบแพร่กระจายกลับ (Training by Backpropagation)

วิธีการเรียนรู้แบบแพร่กระจายกลับเป็นการเรียนรู้แบบมีครูสอนสามารถประยุกต์ใช้กับโครงข่ายประสาทเทียมหลายชั้นแบบป้อนไปข้างหน้า ทุกรูปแบบโดยการใช้ฟังก์ชันของการกระตุ้น (Activation function) ที่สามารถหาอนุพันธ์ได้ การหาค่าถ่วงน้ำหนักที่เหมาะสมจะใช้วิธีการ Optimization ที่เรียกว่า Gradient descent ซึ่งจะทำการปรับค่าถ่วงน้ำหนักเพื่อที่จะลดความผิดพลาดของการคำนวณในแต่ละรอบในระหว่างการเรียนรู้ รูปแบบของอินพุตจะถูกส่งผ่านไปข้างหน้าตามลำดับชั้น จนได้ผลการคำนวณของเอาท์พุต จากนั้นผลการคำนวณจะถูกนำมาเปรียบเทียบกับข้อมูลจริงหรือข้อมูลเป้าหมายเพื่อหาค่าคลาดเคลื่อนที่เกิดขึ้น ค่าความคลาดเคลื่อนดังกล่าวจะนำไปใช้เป็นข้อมูลอินพุตย้อนกลับ ในขณะเดียวกันจะทำการ

ปรับค่าถ่วงน้ำหนักที่เชื่อมโยงระหว่างหน่วยในแต่ละชั้นในทิศทางย้อนกลับ เพื่อให้ข้อมูลเอาท์พุตใหม่ใกล้เคียงหรือตรงตามค่าเป้าหมายที่ต้องการ

สำหรับอัลกอริทึมแบบแพร่กระจายกลับนี้ จะคำนวณหาค่าความผิดพลาดเป็นแบบค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean square error, MSE) ซึ่งแสดงเป็นสมการได้ดังนี้

$$MSE = \frac{1}{N}\sum_{i}^{N}(t_i - z_i)^2 \quad \text{..... (2.8)}$$

เมื่อ MSE คือ ค่าความผิดพลาดเป็นแบบค่าคลาดเคลื่อนกำลังสองเฉลี่ย

N คือ จำนวนข้อมูล

t คือ เวกเตอร์เป้าหมาย

z คือ แทนเวกเตอร์เอาต์พุท

สำหรับกระบวนการทำงานของอัลกอริทึมแบบแพร่กระจายกลับมีขั้นตอนการทำงานออกเป็น 4 ขั้นตอน ดังนี้

ขั้นที่ 1 การคำนวณค่าน้ำหนักอินพุต

ตัวแปรที่อยู่ในชั้นอินพุต (Input Layer) แต่ละตัวจะถูกส่งไปยังชั้นซ่อน (Hidden Layer) โดยจะมีค่าของน้ำหนัก (Weight) เฉพาะตัว ค่าน้ำหนักดังกล่าวเป็นค่าที่แสดงถึงความสัมพันธ์ของตัวแปรนั้นกับเซลล์ที่ส่งไปในชั้นซ่อนในแต่ละครั้ง ค่าน้ำหนักดังกล่าวจึงเป็นค่าตั้งต้นของกระบวนการฝึกของระบบเครือข่ายประสาท ค่าน้ำหนักที่สามารถทำให้เกิดการส่งสัญญาณออกไปได้เป็นค่าน้ำหนักเชิงบวก (Positive Weight) ส่วนค่าน้ำหนักที่ไม่สามารถทำให้เกิดการส่งสัญญาณออกไปได้เป็นค่าน้ำหนักเชิงลบ (Negative Weight) โดยที่ค่าน้ำหนักทุกค่าที่ถูกส่งมายังแต่ละหน่วย (Unit) ของชั้นซ่อนจะถูกคำนวณด้วยฟังก์ชันผลรวม

ขั้นที่ 2 การส่งถ่ายค่าผลรวมน้ำหนักไปยังชั้นต่อไป (Transfer Weight)

ค่าผลรวมที่ได้จากการคำนวณจากฟังก์ชันในกระบวนการก่อนหน้านี้ ผลรวมนั้นจะถูกส่งออกไปยังชั้นต่อไป ซึ่งอาจยังเป็นชั้นซ่อนอยู่ ในกรณีที่การออกแบบโครงข่ายนั้นมีชั้นซ่อนมากกว่า 1 ชั้น ส่วนในกรณีที่มีชั้นซ่อนเพียง 1 ชั้น ค่าผลรวมน้ำหนักดังกล่าวจะถูกส่งไปยังชั้นเอาต์พุต โดยการส่งถ่ายดังกล่าวต้องใช้ฟังก์ชันที่เรียกว่า แอกติเวชันฟังก์ชัน (Activation Function) ซึ่งมักจะอยู่ในรูปฟังก์ชันแบบไม่เป็นเชิงเส้น (Nonlinear Function) โดย

มหาวิทยาลัยราชภัฏมหาสารคาม RAJABHAT MAHASARAKHAM UNIVERSITY

หากค่าที่คำนวณได้จากฟังก์ชันดังกล่าวมีค่ามากกว่าค่าที่กำหนดไว้ (Threshold Value) จึงจะเกิดผลลัพธ์

ขั้นที่ 3 การเปรียบเทียบค่าผลลัพธ์

ขั้นตอนนี้จะเป็นการเปรียบเทียบค่าผลลัพธ์ที่ได้จาก แอกติเวชันฟังก์ชัน กับค่าที่กำหนดไว้ หากค่าความแตกต่างดังกล่าวอยู่ในเกณฑ์ที่ยอมรับได้ ก็หยุดการทำงานของโครงข่ายประสาทหรือเรียกว่าการสิ้นสุดการฝึกข่ายงานโครงข่ายประสาทเทียม ซึ่งข้อกำหนดความแตกต่างที่สามารถยอมรับได้สำหรับโครงข่ายประสาทเทียม นิยมใช้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean square error, MSE) ที่มีค่าน้อยที่สุด

ขั้นที่ 4 การปรับปรุงค่าน้ำหนัก

ขั้นตอนนี้ของการปรับปรุงค่าน้ำหนักนี้จะทำก็ต่อเมื่อค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าอยู่ในเกณฑ์ที่ยอมรับไม่ได้ ผู้วิจัยก็จะต้องทำการปรับค่าน้ำหนัก ซึ่งก็คือ เริ่มขั้นตอนที่ 1 ถึง 3 ใหม่ และต้องทำเช่นนี้ไปเรื่อยๆ จนกว่าจะได้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่พอใจ (ในบางกรณีจะเรียกว่า ค่าผิดพลาดที่ยอมรับได้)

จะเห็นได้ว่าโครงข่ายประสาทเทียม ก็คือโมเดลทางคณิตศาสตร์ สำหรับประมวลผลสารสนเทศ เพื่อจำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำรูปแบบ (Pattern Recognition) และการอุปมานความรู้ ( Knowledge deduction) เช่นเดียวกับความสามารถที่มีในสมองมนุษย์

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

## งานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษา ค้นคว้างานวิจัยที่เกี่ยวข้องกับการรู้จำใบหน้าด้วยวิธีการต่าง ๆ เพื่อใช้เป็นพื้นฐานในการวิจัย การรู้จำใบหน้าในปัจจุบันนี้ สามารถนำไปใช้ได้จริง และมีความเร็วเป็นแบบเวลาจริง (Real Time) ซึ่งเป็นเรื่องที่ท้าทายอย่างมาก ที่เราสามารถทำให้คอมพิวเตอร์มีความสามารถรู้จำ ซึ่งเลียนแบบการมองเห็นของมนุษย์ ที่มีการพัฒนาความสามารถขึ้นตามลำดับ

### 1. งานวิจัยในประเทศ

เชิดชัย พิมพา (2550 : บทคัดย่อ) ได้ทำการวิจัยการรู้จำใบหน้ามนุษย์โดยใช้ภาพใบหน้าไอเกนจากภาพใบหน้าที่หันหน้าตรง ในงานวิจัยนี้จะนำภาพในกลุ่มเรียนรู้ไปทำการสร้างใบหน้าไอเกนด้วยทฤษฎีการวิเคราะห์องค์ประกอบหลักของภาพใบหน้าตรง การวิจัยนี้ใช้

ข้อมูลทั้งหมด 100 ภาพ แบ่งเป็นกลุ่มเรียนรู้ 90 ภาพ และกลุ่มทดสอบ 10 ภาพปรากฏว่ามีความถูกต้องเป็นร้อยละ 100

ธราพงษ์ เพียซ้าย (2543 : บทคัดย่อ) ได้ทำการวิจัยเรื่องการรู้จำภาพใบหน้าคนด้วยการวิเคราะห์โดยรวมและการวิเคราะห์เฉพาะที่ โดยในการวิเคราะห์โดยรวมจะใช้วิธีการวิเคราะห์องค์ประกอบหลัก โดยภาพหน้าคนจะถูกแปลงเป็นเวกเตอร์น้ำหนักแล้วนำมาเปรียบเทียบกับฐานข้อมูลภาพหน้าคน หากพบว่าไม่สามารถจับคู่เปรียบเทียบได้ จะถูกนำไปวิเคราะห์อีกครั้งโดยการวิเคราะห์พาะที่ด้วยวิธีอินแวเรียนท์โมเมนต์ การทดลองได้ทดลองกับภาพหน้าคนจากฐานข้อมูลภาพ ORL (Olivetti Research Lab) ผลการทดลองปรากฏว่าวิธีการที่นำเสนอ ให้ผลความถูกต้องประมาณร้อยละ 76 ซึ่งถูกต้องเพิ่มขึ้นร้อยละ 14 เมื่อเปรียบเทียบกับวิธีการวิเคราะห์องค์ประกอบหลักเพียงอย่างเดียว

พรรณฤมล เต็มดี (2541 : บทคัดย่อ) ได้นำเสนอวิธีการรู้จำหน้าคนโดยใช้รหัสข้อมูลแบบเศษส่วน (fractal) ของขอบเขตภาพหน้าและโครงข่ายประสาทเทียมแบบแพร่กลับ การวิจัยนี้ใช้เพียงพื้นที่ที่สนใจบนใบหน้าเท่านั้นในการรู้จำ โดยกำหนดให้เป็นบริเวณที่ครอบคลุม คิ้ว ตา และจมูก วิธีการที่นำเสนอเริ่มต้นด้วยการนำภาพขอบของพื้นที่ที่สนใจดังกล่าวไปเข้ารหัสข้อมูลแบบเศษส่วน เพื่อทำการลดปริมาณของข้อมูลที่ใช้ในการระบุใบหน้า จากนั้นรหัสข้อมูลแบบเศษส่วนถูกป้อนเป็นอินพุตของโครงข่ายประสาทเทียมแบบแพร่กลับ 4 ชั้นเพื่อดำเนินการรู้จำและระบุตัวบุคคล ทดสอบกับภาพหน้าของคนจำนวน 50 คน โดยใช้คนละ 15 ภาพ สำหรับการเรียนรู้ของโครงข่ายประสาทเทียมโดยการใช้เวลาในการเรียนรู้ 1.5 ชั่วโมง และทดสอบกับภาพหน้าคนจำนวน 50 ภาพ ปรากฏว่าการประมวลผลเพื่อการรู้จำตามวิธีการที่นำเสนอใช้เวลา 10 วินาทีต่อภาพและสามารถให้อัตราการรู้จำเฉลี่ยถูกต้องร้อยละ 80

**2. งานวิจัยต่างประเทศ**

บรูเนลลีและพอกกิโอ (Brunelli and Poggio. 1993) ใช้วิธีการเทียบเทมเพลต ทำงานโดยการหาสหสัมพันธ์ของภาพสองภาพโดยตรง ได้ผลอย่างมีประสิทธิภาพ โดยมีอัตราในการรู้จำเป็น 100% เมื่อภาพมีขนาดเดียวกัน มีการวางตรงกัน และมีการส่องสว่างของแสงเดียวกัน ถ้าภาพที่เข้ามาทดสอบ มีการเปลี่ยนแปลงไปจากเดิม ต้องมีการประมวลผลภาพก่อน เพื่อให้ได้ภาพที่มีการส่องสว่างใกล้เคียงกัน

เตริกและเพนท์แลนด์ (Turk and Pentland. 1991) ได้ริเริ่มและเสนอวิธี การในการรู้จำใบหน้าโดยการฉาย (Projection) ภาพใบหน้าไปยังองค์ประกอบหลัก (Principle Component)

มหาวิทยาลัยราชภัฏพระนคร
PA JARPHAT MAHASARAKHAM UNIVERSITY

โดยเรียกภาพดังกล่าวว่า เป็นภาพใบหน้าไอเกน (Eigenface) โดยนำภาพใบหน้าไอเกนดังกล่าวไปทำการค้นหาในฐานข้อมูล ใบหน้าไอเกน โดยแสดงให้เห็นการทดสอบกับฐานข้อมูลใบหน้าไอเกน 16 คน ที่มีการเลื่อน และวางตำแหน่งศีรษะหลายๆ ลักษณะ มีการย่อขยาย และมีการเปลี่ยนแปลงทางแสง และเปลี่ยนแปลงเทคนิคเพียงเล็กน้อย โดยระบบมีความสามารถในการรู้จำ เมื่อมีการเปลี่ยนแปลงของแสง มีความสามารถในการรู้จำ 85 %

คานาเด (Kanade. อ้างถึงในเชิดชัย พิมพา. 2550 : 14) ได้เสนอวิธีการแยกลักษณะเด่นอัตโนมัติ (Automatic Feature) โดยอัตราระยะทาง (Ratio of Distance) โดยมีการรู้จำอยู่ระหว่าง 45-75% ของฐานข้อมูลภาพ 20 ภาพ

บรูเนลลีและพอกกิโอ (Brunelli and Poggio. 1992) ได้คำนวณเซตของลักษณะเรขาคณิต เช่น ความกว้างของจมูก ตำแหน่งของปากบนใบหน้า รูปร่างของคาง โดยมีอัตราการรู้จำ 90% ของฐานข้อมูล 47 คน และพวกเขายังใช้วิธีเทียบเทมเพล็ต อย่างง่ายให้ผลในการรู้จำ 100 % โดยใช้ฐานข้อมูลเดียวกัน

เทมดีและคณะ (Temde. อ้างถึงในเชิดชัย พิมพา. 2550 : 15) ได้ทำการหาเส้นขอบของภาพใบหน้าและนำมาตัดเป็นส่วนๆขององค์ประกอบของใบหน้าเช่น ตา จมูก ผม แล้วนำมาเข้ารหัสภาพแต่ละส่วนเพื่อใช้เป็นอินพุตให้กับโครงข่ายประสาทเทียมแบบแพร่กระจายย้อนกลับ (Backpropagation Neural Network) สำหรับใช้ในการรู้จำใบหน้า ซึ่งได้ทำการทดลองกับฐานข้อมูล ORL ซึ่งได้ผลความถูกต้องในการรู้จำ 85%

เดเมอร์ (Demer. อ้างถึงในเชิดชัย พิมพา. 2550 : 15) ใช้สัมประสิทธิ์ 50 ตัวแรกจากส่วนประกอบหลักที่ดึงมาจากภาพแล้วนำมาใช้โครงข่ายประสาทเทียมแบบมาตรฐานของเปอร์เซฟตรอนหลายชั้น (Standard Multilayer Perceptron) ในการรู้จำ ผลการจำค่อนข้างดี แต่ฐานข้อมูลที่ใช้ค่อนข้างง่ายเกินไป และภาพไม่มีการเปลี่ยนแปลงของแสง ไม่มีการเปลี่ยนแปลงจากการหมุน โดยใช้ฐานข้อมูลแค่ 20 คนเท่านั้น

ลอว์เรนซ์และคณะ (Lawrence and others. 1997) ใช้โครงข่ายประสาทเทียมแบบผสม (Hybrid Neural Network) โดยใช้การผสมของโครงข่ายประสาทเทียมแบบ Self-Organizing Map (SOM) ในการลดมิติ (Quantization) ของข้อมูลภาพใบหน้าไปในโทโปโลยีสเปส (Topology Space) และโครงข่ายประสาทเทียมแบบคอนโวลูชั่น (Convolution Neural Network : CNN) ในการรู้จำ และได้เปรียบเทียบผลการรู้จำได้ ดังนี้

1. ใช้ขั้นตอนวิธี SOM ในการลดมิติข้อมูลและใช้วิธี CNN ในการรู้จำ มีความถูกต้องของการการรู้จำ 96.2%

มหาวิทยาลัยราชภัฏพระนครศรีอยุธยา
RAJABHAT MAHASARAKHAM UNIVERSITY