สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การจำแนกภาพโดยนิวรอลเน็ตเวิร์คแบบไม่มีการสอนร่วมกับแบบมีการสอน

รศ.ดร. กิตติ ไพฑูรย์วัฒนกิจ
รัชนี สุภัคพาณิช
สุรพล คุ้นเคย

งานประกาศ ปริญญา

T034417

RCH
TK
7882
.P3
ก67ว2ก

เลขหมู่.....
เลขทะเบียน.....34417.....
วัน, เดือน, ปี- 1 พ.ย. 2542

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การจำแนกภาพโดยนิวรอลเน็ตเวิร์คแบบไม่มีการสอนร่วมกับแบบมีการสอน

บทคัดย่อ

ในรายงานนี้เป็นการนำเสนอวิธีการของการจำแนกวัตถุภายในภาพ โดยใช้หลักการของโครงข่ายประสาทเทียมแบบไม่มีการสอนร่วมกับแบบมีการสอน สำหรับนิวรอลแบบไม่มีการสอนจะใช้ Self-organization ของ Kohonen และนิวรอลแบบมีการสอนใช้โครงสร้างแบบ Recurrent โดยใช้การสอนแบบ Back propagation และนำหลักการสร้างความพอใจในขอบเขตที่มีอย่างจำกัดเข้ามารวมไว้ด้วยกัน ทำให้อัลกอริทึมนั้นสามารถทำงานในลักษณะแบบขนานได้ซึ่งจะแตกต่างจากอัลกอริทึมแบบเดิมที่อัลกอริทึมจะทำงานในลักษณะแบบลำดับ จากการทำงานที่แตกต่างนั้นทำให้วิธีการที่นำเสนอนี้ มีความสามารถในการสร้างค่าความพอใจสูงสุดโดยรวมได้ และด้วยความสามารถในการเรียนรู้แบบไม่ต้องมีฟังก์ชันเป้าหมายของโครงข่ายประสาทเทียมแบบ Kohonen's self organizing feature map ทำให้ลักษณะโครงสร้างของภาพ นั้นมีการจดจำในรูปของค่าน้ำหนัก ซึ่งค่าน้ำหนักเหล่านี้จะนำไปเป็นค่าเริ่มต้นให้แก่โครงข่ายประสาทเทียมแบบ Recurrent ที่มีความสามารถในการหาค่าที่ดีที่สุดได้ และด้วยหลักการสร้างความพอใจในขอบเขตที่จำกัด โครงข่ายประสาทเทียมแบบ Recurrent จะมีการเชื่อมต่อถึงกันทั้งหมดด้วยค่าน้ำหนักอีกชุดหนึ่งซึ่งเป็นไปตามวิธีของการสร้างความพอใจในขอบเขตที่จำกัด ดังนั้นโครงสร้างของโครงข่ายประสาทเทียมที่สร้างขึ้นมานี้จะเป็นแบบสามมิติ โดยที่ทุกๆ จุดภาพภายในภาพจะมีตัวแทนที่เป็นโครงข่ายประสาทเทียม Recurrent เท่าๆ กันทุกจุดภาพ โดยที่จำนวนของโครงข่ายประสาทเทียมจะขึ้นอยู่กับความต้องการในการจำแนก และหนึ่งใน Recurrent ของแต่ละจุดภาพที่ให้ค่าเอาท์พุทที่สูงที่สุด จะได้ว่าเป็นตัวแทนของจุดภาพนั้น จากผลการทดลอง อัลกอริทึมนี้สามารถให้ผลลัพธ์ที่น่าพอใจกับภาพที่มีลักษณะการกระจายของระดับเทาแบบต่างๆ กันได้ดี แต่ไม่สามารถทำงานได้ดีกับภาพถ่ายดาวเทียมที่มีลักษณะครุมเครือมากๆ ได้

1. บทนำ

การแบ่งวัตถุภายในภาพ คือ ขั้นตอนที่สำคัญขั้นตอนหนึ่งในขบวนการจดจำคุณลักษณะต่างๆ ของภาพ เพราะถ้าการแบ่งแยกนั้นทำได้ไม่ดีแล้ว ก็จะทำให้การจดจำนั้นมีความผิดพลาดที่มากขึ้นหรือไม่สามารถทำการจดจำได้เลย หลักของการแบ่งแยกวัตถุนั้น โดยส่วนใหญ่แล้วจะยึดหลักการคุณสมบัติที่เหมือนกันของระดับเทาหรือไม่ ก็ลักษณะของเนื้อภาพเป็นแนวในการแบ่งแยกวัตถุ ในปัจจุบันอัลกอริทึมที่ใช้อยู่นั้นจะมีแนวทางในการทำงานกับภาพได้สามแนวทาง คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การแบ่งภาพออกเป็นพื้นที่โดยพื้นที่เป็นส่วนเดียวกันหรือมีตัวแทนเป็นตัวแทนเดียวกัน จะมีคุณลักษณะที่เหมือนๆ กัน วิธีการที่สอง คือ การหาความไม่ต่อเนื่องของขอบของวัตถุภายในภาพและทำการเชื่อมต่อเข้าด้วยกันเพื่อให้ได้วัตถุที่มีความสมบูรณ์มากที่สุด และในส่วนวิธีการท้ายสุดก็คือการแบ่งแยกวัตถุในลักษณะการจำแนกตามระดับเทา ซึ่งในทุกๆ แนวความคิดที่ได้อธิบายมานี้ความต้องการคือการได้ผลลัพธ์ที่ออกมาแล้วมีความหมายต่อการตีความ ในขบวนการขั้นต่อไป แต่ด้วยข้อจำกัดของอัลกอริทึมที่ทำงานในแบบลำดับ ทำให้ผลลัพธ์ที่ได้มาไม่มีความสัมพันธ์ต่อการทำงานของอัลกอริทึม หรืออธิบายได้ว่าอัลกอริทึมไม่สามารถสร้างค่าตัวแทนต่างๆ ที่ให้แก่วัตถุในภาพนั้น เป็นตัวแทนที่ดีที่สุดของโดยรวมได้นั้นเอง

จากข้อจำกัดของวิธีการในการจำแนกภาพแบบเดิม ทำให้มีการประยุกต์ทฤษฎีต่างๆ เข้ามาแก้ไข และหนึ่งนั้นคือ วิธีการของนิวรอลเน็ตเวิร์ค หรือโครงข่ายประสาทเทียมที่มีความสามารถในการเรียนรู้ข้อมูลได้โดยการสอนให้แก่ตัวโครงข่ายนิวรอล และความสามารถในการทำงานแบบขนาน ทำให้ทฤษฎีได้รับความสนใจในการที่จะประยุกต์ทฤษฎีนี้เข้ามาในงานการจำแนกภาพ และทฤษฎีของนิวรอลเน็ตเวิร์ค์นั้นได้มีการพัฒนามาหลายโมเดลที่มีความสามารถที่แตกต่างกันทำให้การประยุกต์ในการใช้งานทางด้านนี้จึงมีมากขึ้น

สำหรับในรายงานที่นำเสนออยู่นี้เป็นรายงานหนึ่งที่ได้มีการประยุกต์ทฤษฎีนิวรอลเน็ตเวิรค์ รวมเข้ากับแนวความคิดของการสร้างความพอใจในขอบเขตที่จำกัด เพื่อให้ได้วิธีการจำแนกภาพที่มีความสามารถในการทำงานแบบขนาน และสามารถที่จะสร้างค่าตัวแทนของแต่ละจุดภาพที่ดีที่สุดโดยรวมได้ด้วย โดยได้ใช้โมเดลของนิวรอลเน็ตเวิรค์สองโมเดล คือ Kohonen's self-organizing feature map ที่มีความสามารถในการเรียนรู้ข้อมูลได้เองโดยไม่ต้องมีฟังก์ชันเป้าหมาย ทำให้ข้อมูลของภาพนั้นได้ถูกจัดเก็บในรูปของค่าน้ำหนัก ซึ่งจะถูกนำมาเป็นค่าเริ่มต้นให้กับโมเดลของ Recurrent นิวรอลเน็ตเวิรค์ ที่อยู่ภายใต้โครงสร้างของหลักการสร้างความพอใจในขอบเขตที่จำกัด ด้วยคุณสมบัติของ Recurrent นิวรอลเน็ตเวิร์คนั้นสามารถหาค่าที่ดีที่สุดได้นั้นเมื่อทำการรวมเข้ากับการสร้างความพอใจในขอบเขตที่จำกัดที่จะสร้างค่าน้ำหนักขึ้นมาเชื่อมต่อกับทุก Recurrent นิวรอล และค่าน้ำหนักที่สร้างขึ้นมานั้นก็จะทำงานแบบสมมาตรต่อตำแหน่ง จะทำให้ได้ค่าความพอใจแบบโดยรวมได้

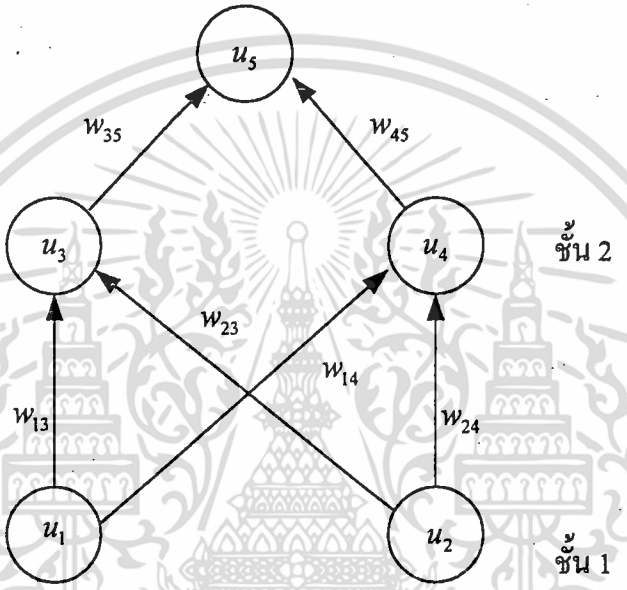
สำหรับในส่วนที่สองของรายงานนี้ เป็นการอธิบายวิธีการในการประยุกต์แนวความคิดของการสร้างค่าความพอใจสูงสุดเข้ากับนิวรอลเน็ตเวิร์ค และในส่วนที่สามจะเป็นการอธิบายถึงการสร้างค่าน้ำหนักในการเชื่อมต่อระหว่าง Recurrent นิวรอลเน็ตเวิรค์ และส่วนที่สี่เป็นการอธิบายถึงการลู่เข้าของอัลกอริทึม และส่วนท้ายสุดจะเป็นการอธิบายผลการทดลองและสรุปผลการทดลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. วิธีการประยุกต์แนวคิดการสร้างความพอใจในขอบเขตจำกัดเข้ากับนิวรอลเน็ตเวิร์คในงานการจำแนกภาพ

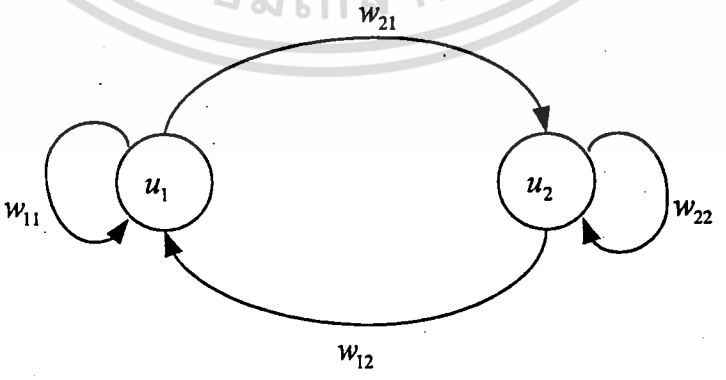
2.1 Recurrent นิวรอลเน็ตเวิร์ค

โครงสร้างของ Recurrent network นั้นจะต่างจากนิวรอลเน็ตเวิร์ค แบบทั่วไปที่เป็นการเชื่อมต่อไปข้างหน้า (Feedforward) คือ โหนดในชั้นเดียวกันจะไม่ต่อถึงกันมีแต่ต่อไปยังโหนดของชั้นถัดไปเท่านั้น ซึ่งจะมีรูปแบบดังรูปที่ 1



รูปที่ 1 โครงสร้างแบบ Feedforward

ในกรณีของ Recurrent network โหนดแต่ละโหนดจะต่อถึงกันและยังป้อนเข้าสู่โหนดของตัวเองดังแสดงในรูปที่ 2



รูปที่ 2 โครงสร้างแบบ Recurrent

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการประยุกต์แนวคิดของการสร้างความพอใจในขอบเขตที่จำกัด หรือที่เรียกย่อๆ ว่า CSP (Constraint Satisfaction Problem) ไปสู่การใช้นิวรอลเน็ตเวิร์คนั้น จะต้องรู้ก่อนว่าขบวนการ CSP คือ ขบวนการในการจัดหาตัวแทน หรือ ลาเบล ที่มีอย่างจำกัดให้กับทุกๆจุดภาพ ได้อย่างเหมาะสมกับภาวะความพอใจสูงสุดหรือการเกิดภาวะของค่าที่ดีที่สุดโดยรวม เพื่อที่ทำให้นิวรอลเน็ตเวิร์คทำงานรวมกับแนวความคิดนี้ได้จะต้องมีการประยุกต์ หรือทำการสร้างโครงสร้างขึ้นมาใหม่ให้ได้เสียก่อน ดังนั้นถ้าเรามีภาพที่มีขนาดเท่ากับ $n \times n$ และเราต้องการจำแนกภาพนี้ให้อยู่ภายในตัวแทน m ตัวแทน ดังนั้น เราจะต้องมีการสร้างนิวรอลขึ้นมาทั้งหมดเท่ากับ $n \times n \times m$ หรือสามารถมองเป็นอะเรย์ขนาดสามมิติ โดยจะมีโครงสร้างของเครือข่ายนิวรอลดังรูปที่ 3 โดยค่า (i,j) เป็นค่าที่อ้างถึงตำแหน่งของจุดภาพที่ (i,j) , $i = 1,...,n, j = 1,...n$

2.2 การแปลงค่าพารามิเตอร์ต่างๆของ CSP สู่นิวรอลเน็ตเวิร์ค

ในการอธิบายวิธีการทำงานของนิวรอลเน็ตเวิร์คให้อยู่ในรูปของ CSP จำเป็นจะต้องทำการแปลงพารามิเตอร์ ของ CSP ให้มีความสอดคล้องกันดังต่อไปนี้

(1) เซ็ตของวัตถุจะใช้สัญลักษณ์ $O = \{o_i, i = 1,...,n\}$,เมื่อ n คือ จำนวนวัตถุในความหมายของการจำแนกภาพทุกๆ จุดภาพที่จะทำการจำแนก ก็คือ วัตถุ

(2) เซ็ตของตัวแทนของแต่ละจุดภาพ หรือแต่ละวัตถุจะใช้สัญลักษณ์ $\Lambda = \{\lambda_j, j = 1,...m\}$ เมื่อ m เป็นจำนวนตัวแทนของแต่ละจุดภาพที่มีอย่างจำกัด

(3) เซ็ตของนิวรอลใน CSP ในที่อยู่ในบริเวณที่ใกล้ชิดกันและมีการเชื่อมต่อถึงกันด้วยค่าน้ำหนักเราจะใช้สัญญลักษณ์ Ω โดยที่ $\Omega_i = \{o_i\} \cup \{o_{i1}, o_{i2}, ..., o_{ip}\}$ เมื่อ $1 \le i_1,...,i_p \le n, i \ne i_1$ สำหรับทุกๆ $1 \le l \le p$ ซึ่งก็มีวัตถุบริเวณใกล้เคียง ที่มีจุดศูนย์อยู่ที่มีการเชื่อมต่อด้วยค่าน้ำหนักระหว่างส่วนประกอบเหล่านี้ $\{o_{i1}, o_{i2}, ..., o_{ip}$ และ o_i

(4) เซ็ตของค่าน้ำหนักของสองวัตถุในบริเวณใกล้เคียง ที่เชื่อมกันด้วยค่าน้ำหนัก $\{o_i, o_{i1}\} \subset \Omega_i$ ที่อยู่บนเซ็ตของตัวแทน $\{\lambda_j, \lambda_{j1}\} \subset \Lambda$ ได้นิยามไว้ดังนี้

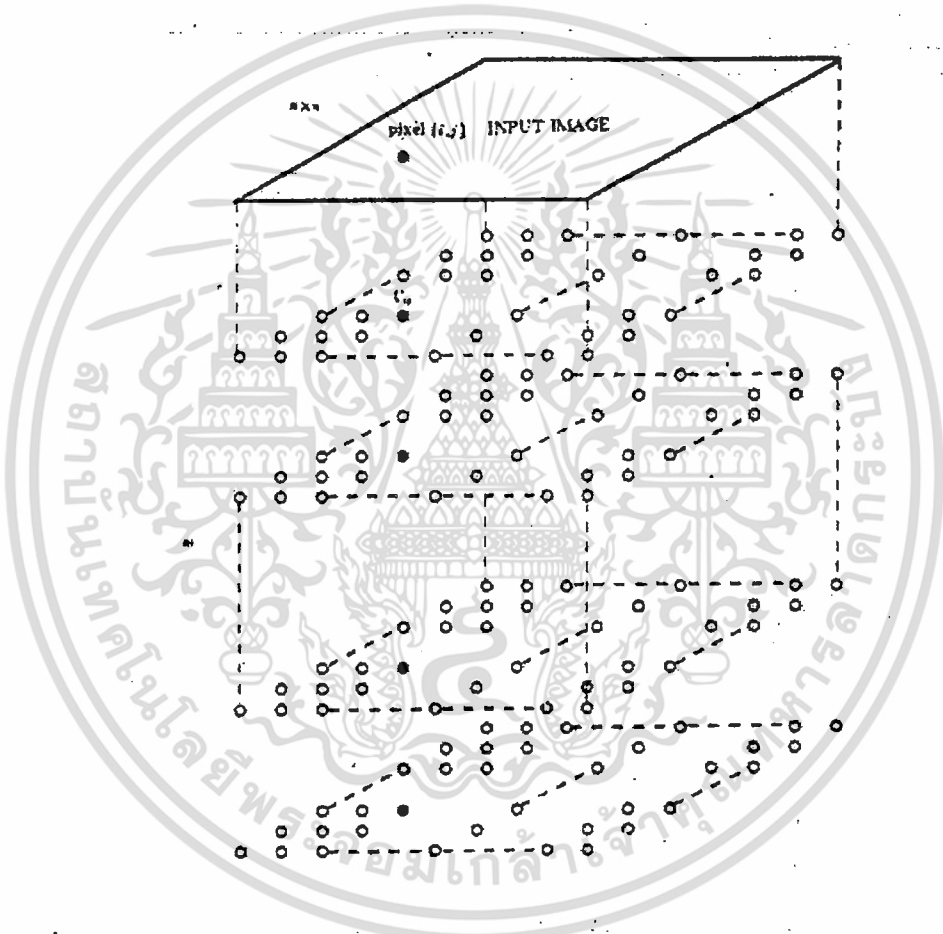
$$\Lambda_{ii1} : \Lambda_i \times \Lambda_{i_1} \to [min, max]$$

เมื่อ $\Omega_i = \{o_i\} \cup \{o_{i1}, o_{i2}, ..., o_{ip}\}, 1 \le i_1,...,i_p \le n, \Lambda_i = \Lambda_{i_2} = \Lambda$ ค่า min คือ จำนวนจริงที่เป็นลบ และค่า max นั้นคือจำนวนจริงที่มีค่าเป็นบวก $\Lambda_{ii1}(\lambda_j, \lambda_{j1})$ สามารถที่จะพิจารณา เป็นมาตราวัดที่สะท้อนถึงความเหมาะสมในการที่จะทำการกำหนดตัวแทน λ_j ให้กับวัตถุ o_i และ λ_{j1} ให้กับ o_{i_1} ที่มีกฎเกณฑ์ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\Lambda_{ii_1}(\lambda_j,\lambda_{j1})=\begin{cases}\text{มีค่ามาก ถ้าตัวแทน } \lambda_j \text{ บนวัตถุ } o_{jl} \\ \quad \text{สนับสนุนตัวแทน } \lambda \text{บนวัตถุ } o_i \\ \text{มีค่าน้อย ถ้าตัวแทน } \lambda_j \text{ บนวัตถุ } o_{jl} \text{ และตัวแทนของ} \\ \quad \lambda \text{บนวัตถุ } o_i \text{ ไม่สอดคล้องกัน} \\ 0 \quad \text{ถ้า } o_i \text{ และ } o_{jl} \text{ อิสระต่อกัน}\end{cases}$$

เมื่อคำว่า มี *ค่ามาก* คือ จำนวนจริงที่มีค่าเป็นบวก และคำว่ามี *ค่าน้อย* นั้น คือ จำนวนจริงที่มีค่าเป็นลบ



รูปที่ 3 เป็นโครงสร้างของ Recurrent Neural Network ในการทำการแบ่งแยกวัตถุจากภาพ โดยทุก ๆ คอลัมน์มีดัชนีที่ (i, j) จะประกอบด้วยจำนวนนิวรอลเท่ากับ m นิวรอล ทำหน้าที่เป็นตัวแทนในขอบเขตที่จำกัดของแต่ละจุดภาพ

2.3 การกำหนดสัญลักษณ์

เพื่อให้ง่ายต่อการอธิบายในการแบ่งแยกวัตถุในภาพ เราจะต้องกำหนดสัญลักษณ์เสียก่อน คือ

$n \times n$ ขนาดของภาพและจำนวนของวัตถุใน CSP

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

m จำนวนของตัวแทนที่มีอย่างจำกัด

$U...$ $n \times n \times m$ เป็นอะเรย์ขนาดสามมิติของนิวรอลเน็ตเวิร์ค

U_{ij} $m \times 1$ เป็นกลุ่มของเวกเตอร์นิวรอลที่จุดภาพตำแหน่ง (i, j)

U_{ijk} เป็นส่วนประกอบลำดับที่ K ของ U_{ij} คือ นิวรอลที่เป็นตัวแทนของเครื่องหมายต่อภาพที่จุดภาพ (i, j)

$O...$ $n \times n \times m$ เป็นอะเรย์ขนาดสามมิติที่แสดงเป็นเอาท์พุทของนิวรอล $U...$

O_{ij} $m \times 1$ เป็นเวกเตอร์ที่แสดงเป็นเอาท์พุทของนิวรอล U_{ij}

O_{ijk} เป็นส่วนประกอบที่ลำดับที่ K ที่แสดงเป็นค่าความน่าจะเป็นของจุดภาพ (i, j) ที่จะมีเครื่องหมายลำดับที่ K

$A...$ $n \times n \times m$ เป็นอะเรย์ขนาดสามมิติที่แสดงเป็นอินพุทของนิวรอลที่ได้มาจากการป้อนกลับของ $O...$

A_{ij} $m \times 1$ เป็นเวกเตอร์ที่แสดงเป็นเอาท์พุทของนิวรอล U_{ij}

A_{ijk} เป็นส่วนประกอบที่ลำดับที่ K ที่แสดงเป็นค่าความน่าจะเป็นของจุดภาพ (i, j) ที่จะมีเครื่องหมายลำดับที่ K

Ω_{ij} เป็นอะเรย์ขนาดสองมิติที่แสดงเป็นค่าน้ำหนักของกลุ่มเพื่อนบ้านของจุดภาพ (i, j)

p เป็นขนาดของอะเรย์สองมิติของน้ำหนักของเพื่อนบ้าน โดยมีจุดศูนย์กลางอยู่ที่ตำแหน่ง ดังนั้น $p = |\Omega_{ij}| - 1$

N_{ij} เป็นอะเรย์ขนาดสามมิติของกลุ่มเพื่อนบ้านที่อยู่รอบ ๆ U_{ij} โดยมีข้อสังเกต คือ (1) $\{U_{ijk}, k = 1, ..., m\} \subset N_{ij}$ และ (2) $\{U_{qrk}, k = 1, ..., m\} \subset N_{ij}$ ถ้า $(q, r) \neq (i, j)$ และจุดภาพ (q, r) ในรูปที่ 4 เป็นการแสดงเพื่อนบ้านขนาด (p = 3 x 3 –1 = 8 และ m = 8) , เมื่อ

$N_{ij} = \{U_{(i-1)(j-1)}, U_{(i-1)j}, U_{(i-1)(j+1)}, U_{i(j+1)}, U_{(i+1)(j-1)}, U_{(i+1)j}, U_{(i+1)(j+1)}\}$ และ

$|N_{ij}| = p * m$

$H...$ $n \times n \times m$ เป็นอะเรย์ขนาดสามมิติที่แสดงเป็นผลบวกของของอินพุทจากเพื่อนบ้าน

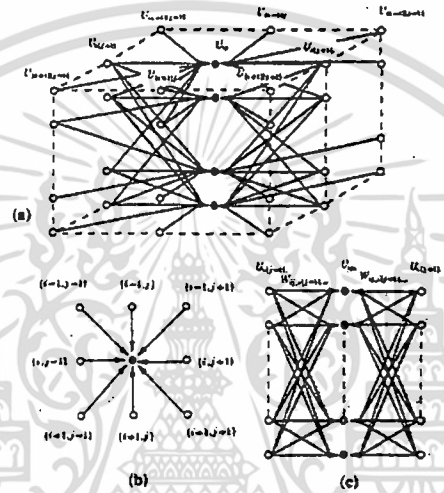
H_{ij} $m \times 1$ เป็นเวกเตอร์ที่แสดงเป็นผลบวกของอินพุทจาก N_{ij} ถึง U_{ijk} ในรูปที่ 5 เป็นการแสดงความสัมพันธ์ของ $O_{ijk}, A_{ijk}, H_{ijk}$ $W_{ij,qr,k,l}$ เป็นค่าน้ำหนักระหว่างนิวรอล $U_{qrl} \in N_{ij}$ และ U_{ijk} ซึ่งค่าน้ำหนักนี้จะมีลักษณะที่สมมาตรกัน คือ $W_{ij,qr,k,l} = W_{qr,ij,l,k}$

$f(i, j)$ เป็นค่าระดับเทาของจุดภาพ (i, j)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

d	จำนวนของนิวรอลใน Kohonen's self-organizing feature map. โดยปกติแล้ว m = d
I_i	เป็นนิวรอลลำดับที่ I ใน feature map เมื่อ $i = 1,\ldots\ldots,d$
$N(I_k,r)$	เป็นเพื่อนบ้านของ I_k โดย r คือ รัศมีของเพื่อนบ้าน
α	เป็นอัตราการเรียนรู้ของ Feature map
w_i	เป็นค่าน้ำหนักของนิวรอล I_i, ใน feature map, $i = 1,\ldots\ldots,d$



รูปที่ 4 เป็นการแสดงถึงความหมายของบริเวณใกล้เคียงเท่ากับ (3 x 3 – 1) x 8 ในรูป (a) เป็นการแสดงความเชื่อมต่อของบริเวณที่ใกล้เคียงกันในรูปสี่เหลี่ยมโดยที่ p = 8, m = 8 ในทุก ๆ นิวรอล U_{ij} จะรับข้อมูลจาก $U_{(i-1)(j-1)}, U_{(i-1)j}, U_{(i-1)(j+1)}, U_{i(j+1)}, U_{(i+1)(j-1)}, U_{(i+1)j}, U_{(i+1)(j+1)}$ (b) เป็นการมองจากด้านบน (c) เป็นการมองแบบตัดขวาง

2.4 ขบวนการทำงานของอัลกอริทึม

ขั้นตอนของการทำงานของอัลกอริทึม เริ่มต้นโดยการสร้างค่าเริ่มต้นให้กับระบบโดยใช้นิวรอลเน็ตเวิร์คโมเดลของ Kohonen's self-organizing feature map ที่มีความสามารถในการเรียนรู้ข้อมูลโดยไม่ต้องใช้ฟังก์ชันเป้าหมาย ผลลัพธ์ที่ได้จะเป็นค่าน้ำหนัก ซึ่งเราจะนำมาใช้เป็นค่าเริ่มต้นให้กับระบบ โดยกำหนดให้เป็นค่าเอาท์พุทของ Recurrent นิวรอลเน็ตเวิร์คทุก ๆ ตัวซึ่งทำหน้าที่เป็นตัวแทนให้แก่จุดภาพทุก ๆ จุดภาพ ต่อจากนั้น Recurrent นิวรอลเน็ตเวิร์คจะทำการหาค่าที่ดีที่สุดในขณะเดียวกันค่าทุก ๆ นิวรอลที่เป็นบริเวณใกล้เคียงซึ่งกันและกันก็จะมีการส่งค่าน้ำหนักให้กันโดยที่ขั้นตอนเหล่านี้จะถูกทำซ้ำ ๆ ไปเรื่อยจนกระทั่งระบบมีการลู่เข้าหาคำตอบหรือมีค่าผิดพลาดน้อยจนรับได้ระบบก็จะหยุดทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

34417

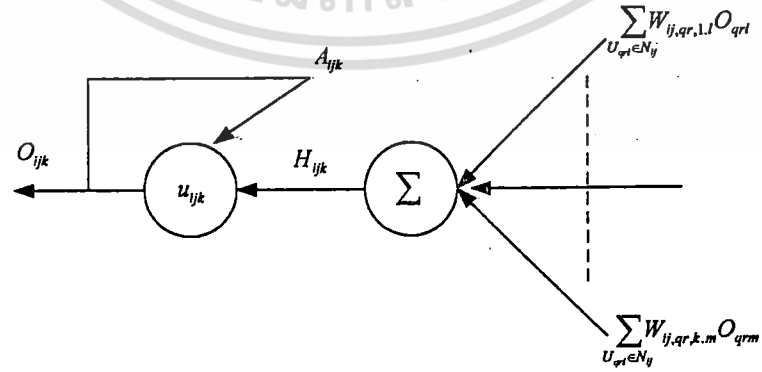
จากขั้นตอนที่กล่าวไว้ข้างบนนั้น เมื่อเรานำมาเขียนในทางคณิตศาสตร์ เพื่อที่จะนำไปสร้างเป็นโปรแกรมขึ้นมาได้ก็สามารถที่จะแสดงได้ ดังนี้

$$H\ldots = \begin{bmatrix} H_{11} & H_{12} & \ldots & H_{1n} \\ H_{21} & H_{22} & \ldots & H_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ H_{n1} & H_{n2} & \ldots & H_{nn} \end{bmatrix} \qquad (1)$$

เมื่อ

$$H_{ij} = \begin{bmatrix} H_{ij1} \\ H_{ij2} \\ \vdots \\ H_{ijm} \end{bmatrix} = \begin{bmatrix} \sum_{U_{qrl} \in N_{ij}} W_{ij,qr,1,l} O_{qrl} \\ \sum_{U_{qrl} \in N_{ij}} W_{ij,qr,2,l} O_{qrl} \\ \vdots \\ \sum_{U_{qrl} \in N_{ij}} W_{ij,qr,m,l} O_{qrl} \end{bmatrix}$$

จากการสนับสนุนของเพื่อนบ้านจะมีคุณสมบัติเป็นดังนี้ : (1) ถ้าทั้ง O'_{qrl} และ $W_{ij,ar,k,l}$ มีค่าเป็นบวกแล้ว $O'_{qrl} W_{ij,ar,k,l}$ จะช่วยให้นิวรอล U_{ijk} มีค่าเป็นบวกมากขึ้น (2) ถ้า O'_{qrl} มีค่าสูง แต่ค่า $W_{ij,ar,k,l}$ มีค่าต่ำแล้ว $O'_{qrl} W_{ij,ar,k,l}$ จะช่วยให้นิวรอลที่ U_{ijk} นั้นมีค่าเป็นลบมากขึ้น (3) ถ้า O'_{qrl} มีค่าสูงและ $W_{ij,ar,k,l}$ มีค่าเท่ากับศูนย์แล้ว $O'_{qrl} W_{ij,ar,k,l}$ จะไม่มีผลต่อ U_{ijk} โดยที่ H_{ijk} จะเป็นตัวแทนของการสนับสนุนจากบริเวณที่ใกล้เคียงจะมีส่วนร่วมในการปรับค่าระดับการกระตุ้นของนิวรอล U_{ijk} หลังจากที่ได้รับสัญญาณจากบริเวณใกล้เคียง N_{ij} ทุก O_{ijk} จะมีการปรับค่า (Update) ให้เป็นไปตาม H_{ijk} และ A_{ijk} ในรายงานนี้ A_{ijk} คือสัญญาณที่ป้อนกลับจาก O_{ijk} ดังนั้น $A_{ijk} = O'_{ijk}$ โดยที่ขบวนการปรับค่าของ O... จะเป็นตามขั้นตอนดังนี้



รูปที่ 5 เป็นการแสดงความสัมพันธ์ระหว่าง O_{ijk}, A_{ijk}, H_{ijk} ได้มาจากการป้อนกลับของสัญญาณ

O_{ijk} และ $H_{ijk} = \sum_{U_{qrl} \in N_{ij}} W_{ij,qr,m,l} O_{qrl}, l = 1, \ldots, m$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$O^{t+1} = \begin{bmatrix} O_{11}^{t+1} & O_{12}^{t+1} & \cdots & O_{1n}^{t+1} \\ O_{21}^{t+1} & O_{22}^{t+1} & \cdots & O_{2n}^{t+1} \\ \vdots & \vdots & \cdots & \vdots \\ O_{n1}^{t+1} & O_{n2}^{t+1} & \cdots & O_{nn}^{t+1} \end{bmatrix} \quad (2)$$

$$= G\begin{bmatrix} A_{11}^{t} & A_{12}^{t} & \cdots & A_{1n}^{t} \\ A_{21}^{t} & A_{22}^{t} & \cdots & A_{2n}^{t} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1}^{t} & A_{n2}^{t} & \cdots & A_{nn}^{t} \end{bmatrix}, \begin{bmatrix} H_{11}^{t} & H_{12}^{t} & \cdots & H_{1n}^{t} \\ H_{21}^{t} & H_{22}^{t} & \cdots & H_{2n}^{t} \\ \vdots & \vdots & \cdots & \vdots \\ H_{n1}^{t} & H_{n2}^{t} & \cdots & H_{nn}^{t} \end{bmatrix}$$

เมื่อ

$$A_{ij} = \begin{bmatrix} A_{ij1} \\ A_{ij2} \\ \vdots \\ A_{ijm} \end{bmatrix}, O_{ij} = \begin{bmatrix} O_{ij1} \\ O_{ij2} \\ \vdots \\ O_{ijm} \end{bmatrix}$$

และ G นิยามดังต่อไปนี้

$$O_{ijk}^{t+1} = \frac{Pos\left(A_{ijk}^{t} + \Delta A_{ijk}^{t}\right)}{\sum_{l=1}^{m} Pos\left(A_{ijl}^{t} + \Delta A_{ijl}^{t}\right)} \quad (3)$$

เมื่อ

$$\Delta A_{ijk}^{t} = \begin{cases} \delta & if\ H_{ijk}^{t} = \max_{l}\left\{H_{ijk}^{t}\right. \\ -\delta & \text{ในกรณีอื่น} \end{cases}$$

$$Pos(X) = \begin{cases} X & if\ X \geq 0 \\ 0 & \text{ในกรณีอื่น} \end{cases}$$

เมื่อ δ จำนวนจริงที่มีค่าเป็นบวกน้อย ๆ ขบวนการจะมีการปรับค่า (Update) ของเอาท์พุท ทุกๆ นิวรอลบนฐานของเอาท์พุทของนิวรอลบริเวณใกล้เคียง และค่าเอาท์พุทเก่าที่ผ่านมา ในการออกแบบของกฎการปรับค่านี้คือ ยึดหลักการดังนี้ (1) ถ้า H_{ijk} มีค่าสูงสุดแล้วจำนวนที่เป็นบวกจะมีการกระจายผลไปยัง U_{ijk} ด้วย และ A_{ijk}^{t} มีค่าที่เพิ่มขึ้นด้วย (2) ถ้า H_{ijk} มีค่าไม่สูงสุดแล้วจะมีการกระจายผลไปยัง U_{ijk} และ A_{ijk}^{t} ให้มีค่าที่ลดลงในสมการที่ (3) ในส่วนของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$\frac{1}{\sum_{l=1}^{m} Pos\left(A'_{ijl} + \Delta A'_{ijl}\right)}$ เป็นกระบวนการปรับค่าไม่ให้มีค่าเกินกว่าหนึ่ง ดังนั้น ค่าสูงสุดของเอาท์พุท คือ 1 นั่นเอง

2.5 เงื่อนไขเริ่มต้น

วิธีการในการกำหนดเงื่อนไขเริ่มต้นนั้นเราจะยึดหลักการแบ่งโดยใช้การกระจายของระดับเทาของภาพที่เป็นอินพุท โดยใช้ Kohonen's self-organizing feature map เป็นแบบจำลองวิเคราะห์การกระจายของระดับเทาของภาพทางอินพุท หรือพูดได้อีกอย่างว่า โครงข่ายจะทำการแยกระดับเทาออกเป็นหลาย ๆ ระดับ ในการทำงานของแบบจำลองนี้จะมีอยู่สองส่วนด้วยกัน คือ ส่วนของการเรียนรู้ (Leaning Phase) และส่วนของการจัดแบ่ง (Categorization Phase) ในส่วนของการเรียนรู้เมื่อค่าระดับเทาของจุดภาพนั้นเข้ามาสู่เน็ตเวิร์คที่เป็นแบบมิติเดียว จะมีนิวรอลอยู่เพียงตัวเดียวเท่านั้นที่จะถือว่าตอบสนองต่ออินพุทที่เข้ามาได้อย่างดีที่สุด ดังนั้น นิวรอลตัวนั้นและบริเวณใกล้เคียงของมันจะมีการปรับค่าน้ำหนักในการคัดเลือกนิวรอลที่ชนะจะใช้หลักของการวัดระยะทางระหว่างค่าของอินพุทกับค่าของน้ำหนักโดยที่สูตรการหาระยะทางมีดังนี้

$$D_{I_1} = \| wl - f(i,j) \| \quad l = 1,\ldots d$$

โดยสมมุติให้

$$D_{I_K} = \min_{l=1,\ldots d} D_{I_l} \qquad (4)$$

เมื่อ I_k เป็นนิวรอลที่ชนะ หลังจากที่ค้นพบนิวรอลที่ชนะแล้ว ค่าน้ำหนักก็จะมีการปรับโดยมีสูตรคำนวณ ดังนี้

$$w_k'^{t+1} = w_k'^{t} + \alpha'^{t}\left(\mathrm{f}(\mathrm{i},\mathrm{j}) - w_k'^{t}\right) \qquad (5)$$

เมื่อ $w_k'^{t}$ คือ ค่าน้ำหนักของนิวรอล I_k ลำดับที่ k ก่อนที่จะมีการปรับค่าใหม่

$w_k'^{t+1}$ เป็นค่าน้ำหนักหลังจากมีการปรับเรียบร้อยแล้วและ α'^{t} (มีค่าระหว่าง 0 กับ 1) เป็นค่าอัตราการเรียนรู้ที่เวลา t_k ในการปรับค่าน้ำหนักนั้นนอกจากนิวรอลที่ชนะแล้ว นิวรอลในบริเวณใกล้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เคียงก็จะมีการปรับค่าน้ำหนักใหม่เช่นกัน แต่จะน้อยกว่านิวรอลที่ชนะ ซึ่งสูตรในการปรับค่าน้ำหนักของนิวรอลในบริเวณใกล้เคียงมีดังต่อไปนี้

$$w_q^{t_q+1} = w_q^{t_q} + \alpha^{t_q}\left(f(i,j) - w_q^{t_q}\right) \qquad (6)$$

เมื่อ $I_q \in N(I_k, r)$ โดยที่เพื่อนบ้านของนิวรอลลำดับที่ k สามารถนิยามได้ ดังนี้

$$N(I_k, r) = \begin{Bmatrix} I_q \mid q = (k-r), (k-r+1), \ldots, (k-1), (k+1), \ldots, (k+r-1), (k+r), \\ if\ I_q\ exist\ and\ I_q \neq I_k \end{Bmatrix}$$

และสูตรในการปรับกฎของการเรียนรู้ มีดังนี้

$$\alpha^{t_k+1} = K\alpha^{t_k} \qquad (7)$$

เมื่อ K เป็นค่าพารามิเตอร์ที่มีขนาดไม่เกิน 1

โดยสรุปขั้นตอนของการปรับค่าน้ำหนักมีดังนี้

(1) ทำการเริ่มต้นค่าน้ำหนักของ feature map

(2) สำหรับทุก ๆ จุดภาพ *(i, j)* ในภาพที่เป็นอินพุททำตามขั้นตอน ต่อไปนี้

(2.1) หานิวรอลที่ชนะ I_k โดยใช้สูตรหาระยะทางในสมการที่ (4)

(2.2) ปรับค่าน้ำหนักของนิวรอลที่ชนะ I_k โดยใช้สมการที่ (5)

(2.3) ปรับค่าน้ำหนักของเพื่อนบ้านของนิวรอลที่ชนะ $N(I_k, r)$ โดยใช้สมการที่ (6)

(3) ปรับค่า α โดยใช้สมการที่ (7)

(4) ถ้า $\alpha > \varepsilon$ (εเป็นค่าตัวเลขจำนวนน้อย ๆ) กลับไปเริ่มที่ขั้นตอนที่สองใหม่ จนกว่าเงื่อนไขเป็นจริง

ต่อไปเป็นการอธิบายในส่วนของการจัดแบ่ง (Categorization Phase) เป็นขั้นตอนในการกำหนดค่าความน่าจะเป็นแบบไม่มีประสบการณ์ (Prior Probability) โดยยึดหลักการจัดแบ่งตามระดับเทาของจุดภาพที่ทำให้สอดคล้องกับนิวรอลที่มีเครื่องหมาย L นั้นเกิด excited โดยการกำหนดจะให้ค่าความน่าจะเป็นมากสำหรับนิวรอลที่มีเครื่องหมาย L มากที่สุดและจะน้อยลงไปบริเวณนิวรอลที่ใกล้กัน โดยขั้นตอนสามารถที่จะทำการอธิบายได้ ดังนี้

(1) สำหรับทุก ๆ จุดภาพนั้นทำตามขั้นตอนที่สอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(2) กำหนดค่าเริ่มต้นให้กับอินพุทของ A_{ijk}^{0} กับทุก ๆ นิวรอลโดยให้สอดคล้องกับการ feature map ที่ได้เกิด excited โดยวิธีการกำหนดค่า มีดังนี้

$$\text{ถ้า } \|f(i,j)-w_l\| = \min_l \|f(i,j)-w_l\| \text{, แล้ว}$$

$$\begin{cases} A_{ij1}^{0} = \dfrac{v_1}{m} \\ A_{ij2}^{0} = \dfrac{v_2}{m} \\ \vdots \\ A_{ijm}^{0} = \dfrac{v_m}{m} \end{cases} \qquad (8)$$

เมื่อ

$v_1 \geq v_2 \geq \ldots\ldots \geq v_m \geq 0$ และ $\sum_{i=l}^{m} v_i = m$

จากการกำหนดอย่างนี้จะมีเพียงจำนวน v_i จำนวนน้อยที่ไม่เท่ากับ 0 จำนวนที่ v_i ที่ไม่เท่ากับ 0 เราตั้งชื่อว่า s_1

$$\text{ถ้า } \|f(i,j)-w_k\| = \min_l \|f(i,j)-w_l\| \text{, แล้ว}$$

$$\begin{cases} A_{ij1}^{0} = \dfrac{y_1}{m} \\ \vdots \\ A_{ij(k-1)}^{0} = \dfrac{y_{k-1}}{m} \\ A_{ijk}^{0} = \dfrac{y_k}{m} \\ A_{ij(k+1)}^{0} = \dfrac{y_{k+1}}{m} \\ \vdots \\ A_{ijm}^{0} = \dfrac{y_m}{m} \end{cases} \qquad (9)$$

เมื่อ

$y_k \geq y_{(k-1)} \geq \ldots\ldots \geq y_k \geq 0, y_k \geq y_{(k+1)} \geq \ldots\ldots \geq y_m \geq 0$ และ $\sum_{i=l}^{m} y_i = m$

จากการกำหนดอย่างนี้ จะมีเพียงจำนวน y_i จำนวนน้อยที่ไม่เท่ากับ 0 จำนวนที่ y_i ที่ไม่เท่ากับ 0 เราตั้งชื่อว่า s_2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้า $\|f(i,j)-w_d\| = \min_l \|f(i,j)-w_l\|$, แล้ว

$$\begin{cases} A^0_{ij1} = \frac{z_1}{m} \\ A^0_{ij2} = \frac{z_2}{m} \\ \vdots \\ A^0_{ij(m-1)} = \frac{z_{(m-1)}}{m} \\ A^0_{ijm} = \frac{z_m}{m} \end{cases} \qquad (10)$$

เมื่อ

$z_m \geq z_{m-1} \geq \ldots\ldots \geq z_1 \geq 0$ และ $\sum_{i=l}^{m} z_i = m$

จากการกำหนดอย่างนี้ จะมีเพียงจำนวน z_i จำนวนน้อยที่ไม่เท่ากับ 0 จำนวนที่ z_i ที่ไม่เท่ากับ 0 เราตั้งชื่อว่า s_3

โดยปกติแล้ว ถ้าต้องการให้เน็ตเวิร์คลู่เข้าอย่างมีเสถียรภาพจะต้องเป็นตามเงื่อนไขสองข้อ คือ ข้อแรก การกำหนดค่าน้ำหนักเริ่มต้นจะต้องไม่เท่ากัน เพราะว่าการเรียนรู้ของของนิวรอลใช้ระยะทางระหว่างค่าน้ำหนักกับค่าระดับเทาของภาพทางอินพุท ดังนั้น จะต้องมีเพียงนิวรอลตัวใดตัวหนึ่งเท่านั้นที่จะเกิด Excited ได้ ข้อที่สอง คือ ค่าของ α จะต้องลดลงตามเวลาในทางปฏิบัติแล้วค่า α สามารถลดลงอย่างเชิงเส้นได้ โดยที่ขบวนการในการเรียนรู้จะสิ้นสุดโดยอัตโนมัติเมื่อค่า α เข้าใกล้ศูนย์

3. การกำหนดค่าน้ำหนัก

การกำหนดค่าน้ำหนักในระบบก็คือ การสร้างการถ่วงค่าระหว่างนิวรอลในบริเวณที่ใกล้เคียงซึ่งกันและกัน วิธีที่ใช้นี้เป็นวิธีการที่ได้จากการทดลองว่ามีความสมเหตุสมผล ก็คือ นิวรอลที่เป็นตัวแทนที่เหมือนกันนั้นย่อมจะแสดงผลในทางบวกต่อนิวรอลในบริเวณใกล้เคียงที่เป็นตัวแทนตัวเดียวกันของจุดภาพในบริเวณที่ใกล้เคียง และในทางที่กลับกันนั้นย่อมจะมีผลในทางลบกับตัวแทนอื่น ๆ ต่อบริเวณใกล้เคียงในทางคณิตศาสตร์เขียนได้เป็น

$$W_{ij,qr,k,l} = \frac{1}{p}\left(1-2\frac{2|k-1|}{m}\right) \qquad (11)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ $W_{ij,qr,k,l}$ เป็นค่าน้ำหนักระหว่างนิวรอลที่ทำหน้าที่เป็นตัวแทนของจุดภาพที่ลำดับ k ของจุดภาพ (i, j) กับนิวรอลที่ทำหน้าที่เป็นตัวแทนลำดับที่ 1 ของจุดภาพ $(q,r), U_{qrl} \in N_{ij}$

4. การลู่เข้าของ Recurrent Neural Network

หลังจากขั้นตอนของอัลกอริทึมผ่านเงื่อนไขเริ่มต้นแล้ว ก็จะถึงขบวนการทำซ้ำของ Recurrent นิวรอลเน็ตเวิร์ค และอัลกอริทึมจะสิ้นสุดลงตามเงื่อนไขต่อไปนี้

$$CE = \left\| O^{t+1} - O^{t} \right\|^2 < \varepsilon \qquad (12)$$

โดยค่า CE จะลดลงเรื่อยในกรณีอัลกอริทึมมีการทำซ้ำมากขึ้นจนกระทั่งมีค่าน้อยกว่า ε ซึ่งเป็นตัวเลขจำนวนน้อยที่ทำหน้าที่ตรวจเช็คเงื่อนไขในการจบการทำงานของอัลกอริทึม

5. สรุปขั้นตอนการแบ่งแยกวัตถุในภาพ

ขั้นตอนการแบ่งแยกวัตถุจากภาพมีขั้นตอนโดยสรุปดังนี้

(1) กำหนดเงื่อนไขเริ่มต้นที่ได้อธิบายไว้ในส่วนที่ 2.5 และกำหนดให้ H^0_{ijk} มีค่าเท่ากับศูนย์

(2) หาค่าน้ำหนักระหว่าง Recurrent นิวรอลโดยใช้สมการที่ (8)

(3) ตั้งค่า t = 0

(4) ทำการคำนวณ H^t ... โดยใช้สมการที่ (1)

(5) หา O^{t+1} โดยใช้สมการที่ (2)

(6) คำนวณค่า CE ในสมการที่ (12)

(7) ถ้าค่า $CE > \varepsilon$ แล้วตั้งค่า t = t + 1 และกลับไปทำงานในขั้นตอนที่ 4 ใหม่ จนกระทั่งเงื่อนไขเป็นจริง

เมื่ออัลกอริทึมสิ้นสุดลง ค่าเอาท์พุทของนิวรอลในแต่ละคอลัมน์จะมีตัวใดตัวหนึ่ง ที่มีค่าสูงสุดและเราก็จะได้ตัวนั้นเป็นตัวแทนให้แก่จุดภาพจุดนั้น

6. ผลลัพธ์ที่ได้จากการทดลอง

เพื่อแสดงความสามารถวิธีการนี้เราได้ทดลองโดยใช้ภาพขนาด 256*256 พิกเซล โดยมีข้อมูลขนาด 8 บิต แบบระดับเทาและตั้งค่าพารามิเตอร์ของขบวนการไว้ ดังนี้

d = m = 8, p = 8

α = 0.9, r = 1, k = 0.9, = 0.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$s_1 = s_2 = s_3 = 3$

$v_1 = 5,\ v_2 = 4,\ v_3 = 1,\ v_r = 0$ สำหรับ $3 < r \leq m$

$y_k = 6, y_{(k+1)} = 2, y_{(k-1)} = 2,\ y_r = 0$ สำหรับ $r \neq k,\ (k+1),\ (k-1)$

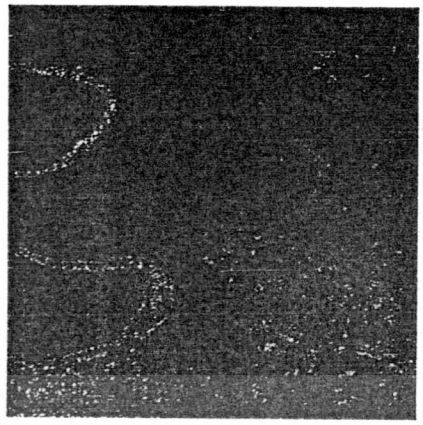
$z_m = 5, z_{(m-1)} = 4,\ z_{(m-2)} = 1, z_r = 0$ สำหรับ $1 \leq r < (m-2)$

โดยที่ Kohonen's self-organizing feature map จะประกอบด้วยนิวรอลจำนวน 8 นิวรอล แบบมิติเดียว ดังนั้น จำนวนนิวรอลที่เป็นแบบ recurrent ทั้งหมดจะเท่ากับ 256*256*8 ตัว ภาพที่แสดงไว้นี้เป็นภาพในรูปแบบต่าง ๆ โดยการคัดเลือกภาพที่นำมาแสดงคำนึงถึงการกระจายของระดับเทาว่ามีความแตกต่างกันอย่างไร ภาพที่ทดสอบจะเป็นภาพ ขนาดข้อมูล 8 บิต ในการทดลองจะใช้ภาพที่เป็นแบบสีเทียมหรือ False colour โดยใช้ภาพทั้งจากแถบสีแดง เขียวและน้ำเงิน มารวมกันก่อนที่จะทำการจำแนกภาพ วิธีนี้จะทำให้ขั้นตอนการทำงานทั้งหมดเร็วขึ้นมากแทนที่จะทำทีละภาพ รูปที่6(a) เป็นภาพต้นแบบที่มีการกระจายของความเข้มตั้งแต่ 0-255 รูปที่ 6(b)-(h) เป็นผลลัพธ์จากการจำแนกภาพ ปัญหาที่พบก็คือพิกเซลบางค่าจะอยู่ระหว่างกลุ่มซึ่งยากต่อการแยกแยะให้เด็ดขาดว่าอยู่ในกลุ่มไหน ทั้งนี้เปอร์เซ็นต์ของการจำแนกผิดกลุ่มคิดเป็น 10-15% วิธีที่จะลดเปอร์เซ็นต์การผิดพลาดก็ต้องใช้ผู้เชี่ยวชาญในการแยกกลุ่มที่ได้จากผลลัพธ์ที่ได้จากนิวรอลของ Kohonen ก่อนเพื่อขจัดจำนวนกลุ่มที่ไม่ควรมีให้น้อยลงและขจัดการ overlap ของแต่ละกลุ่มให้น้อยที่สุดก่อนที่จะป้อนต้นแบบในการสอนให้แก่ นิวรอลแบบ Recurrent เพื่อว่าประสิทธิภาพของอัลกอริทึมทั้งระบบจะดีขึ้นในการจำแนกภาพ โดยเฉพาะเปอร์เซ็นต์การจำแนกผิดพลาด

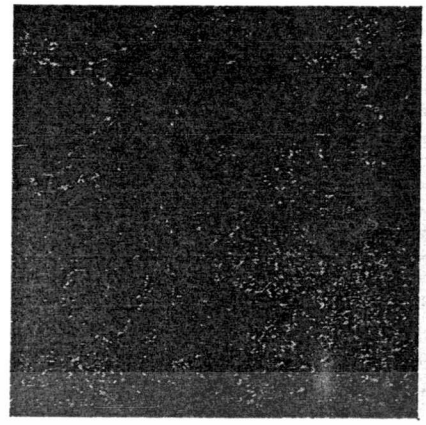
6(a)

6(b)

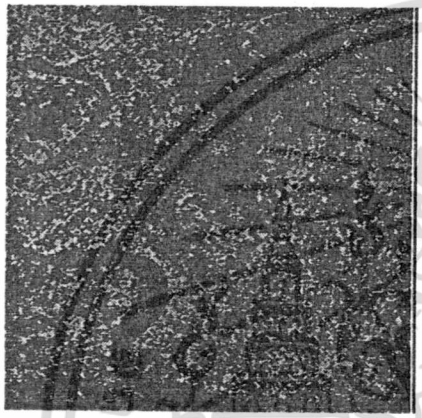
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



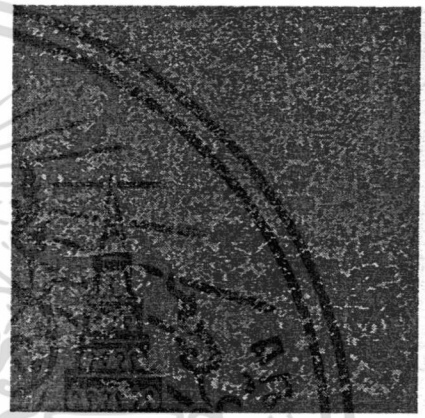
6(c)



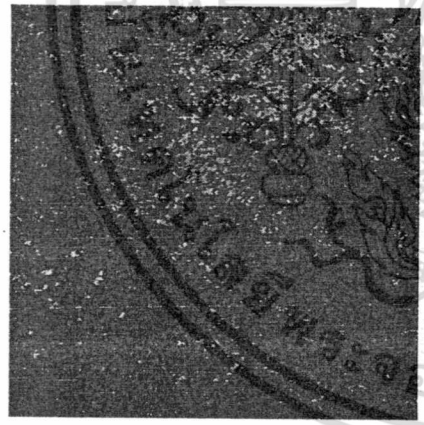
6(d)



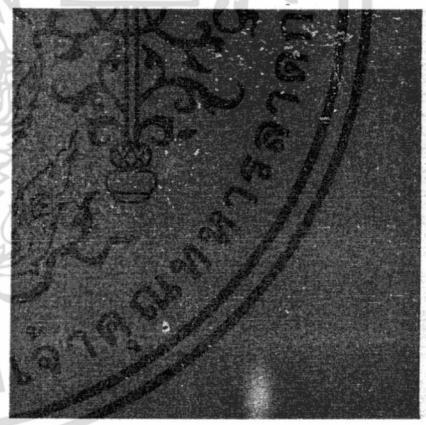
6(e)



6(f)



6(g)



6(h)

รูปที่ 6(a) เป็นภาพต้นแบบ และ (b)-(h) เป็นผลลัพธ์จากการจำแนกภาพ

7. บทสรุป

ในรายงานนี้เป็นการนำเสนอวิธีการในการนำเอาวิธีทฤษฎีนิวรอลเน็ตเวิร์คมาใช้ในงานด้านการแบ่งแยกวัตถุออกจากภาพโดยอัตโนมัติ ซึ่งข้อดีของการใช้นิวรอลเน็ตเวิร์คคือการทำงานแบบขนานทำให้ระบบสามารถแบ่งแยกวัตถุในภาพได้ดีมากขึ้นเพราะเมื่อมีการทำงานอย่างขนานแล้วการตัดสินใจของนิวรอลแต่ละตัวสามารถที่จะหาค่าที่ดีที่สุดได้ (Global Satisfaction) ด้วยการเชื่อม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่อทุกๆนิวรอลด้วยน้ำหนักที่ต่างๆกันซึ่งมีความเป็นเหตุผล ในแง่ของการจำแนกวัตถุ เงื่อนไขการเริ่มต้นนั้นได้ใช้โมเดล Kohonen's elf-organizing feature map ทำหน้าที่ในการคัสเตอร์ริ่ง ซึ่งจำนวนของนิวรอลมากน้อยมีผลต่อ การแบ่งแยกภาพมากพอสมควรหลังจากนั้นเราจะใช้หลักการของ CSP คือทุกๆพิกเซลคือวัตถุและต้องการสิ่งที่มาแทนวัตถุได้อย่างเหมาะสม เรียกสิ่งที่แทนว่าลาเบล ซึ่งด้วยเหตุนี้เน็ตเวิร์คที่ถูกสร้างขึ้นจาก Recurrent neural network ให้แก่ทุกๆวัตถุจะเท่ากับจำนวนของลาเบลที่ระบบนั้นต้องการและทุกนิวรอลจะถูกเชื่อมต่อเข้าด้วยกันโดยใช้ค่าน้ำหนัก ข้อดีของการใช้น้ำหนักของบริเวณใกล้เคียงมาช่วยในการตัดสินใจของนิวรอลคือระบบสามารถที่จะทนต่อสัญญาณรบกวนได้สูงขึ้นและมีโอกาสเกิดค่าที่ดีที่สุด (ลาเบล) ของทั้งภาพได้ (Global Optimization) ได้ง่ายกว่าอัลกอริทึมที่ทำงานในแบบลำดับ ข้อเสียของระบบนี้คือการใช้เวลาในการคำนวณที่นาน และการเก็บรายละเอียดไม่ดีพอ สาเหตุเกิดจากการที่นิวรอลบริเวณใกล้เคียงส่งผลมากเกินไป

เอกสารอ้างอิง

[1] Wei-Chung Lin, "Constraint Satisfaction Neural Networks for Image Network," Pattern Recognition, Vol. 25, No. 7, pp. 679-693, 1992.

[2] F. Rosenblatt, "The perceptron : a probabilistic model for information storage and organization in the brain," Psychological Review, Vol. 65, pp. 513-408, 1986.

[3] Teuvo Kohonen, "Self-orgainzed formation of topologically correct feature maps," Biological Cybernetics, Vol. 43, pp. 59-69, 1982.

[4] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational ability," Proceedings of the National Acedemy of Sciences, Vol. 79, pp. 2554-2558, 1998.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้