

ชื่อเรื่องวิทยานิพนธ์

การพัฒนาระบบรู้จำอักษรล้านนาโดยอาศัย
เคเนียร์เรสเนเบอร์

ผู้เขียน

นายเชาวน์ ป่อแก้ว

ปริญญา

วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)

อาจารย์ที่ปรึกษาวิทยานิพนธ์

รองศาสตราจารย์ ดร. จีรยุทธ ไชยจารุณิช

บทคัดย่อ

วิทยานิพนธ์นี้นำเสนอถึงวิธีการรู้จำอักษรธรรมล้านนา ซึ่งเป็นวิธีการนำภาพของเอกสาร
ภาษาล้านนามาเปลี่ยนให้อยู่ในรูปของข้อความ โดยสกัดภาพตัวอักษรภาษาล้านนาออกจากภาพ
เอกสาร จากนั้นสกัดคุณลักษณะของภาพตัวอักษรด้วยวิธี โชนนิ่งแล้วนำมาสร้างเป็นเวกเตอร์
คุณลักษณะ จากนั้นนำเวกเตอร์ดังกล่าวไปจำแนกประเภทด้วยวิธีเคเนียร์เรสเนเบอร์ จากนั้นนำไป
สร้างเป็นข้อความภาษาล้านนา ซึ่งได้สร้างตัวแบบสำหรับการจำแนกประเภทด้วยวิธีเคเนียร์เรสเน-
เบอร์โดยนำภาพตัวอักษรและส่วนของอักษรภาษาล้านนา 113 ชนิดๆ ละ 50 ตัวอักษรมาทดสอบ
อัตราการรู้จำด้วย K-fold Cross Validation ร่วมกับวิธีการสกัดคุณลักษณะของตัวอักษรแบบ
โชนนิ่งด้วยการปรับขนาดภาพอักษรและจำนวน โชน จากผลการทดลองพบว่าเมื่อปรับขนาดของ
ภาพตัวอักษรให้มีขนาด 81×81 พิกเซล และให้ขนาดของ โชนเท่ากับ 9×9 โชน เมื่อกำหนดให้ค่า
K สำหรับกระบวนการจำแนกประเภทมีค่าเท่ากับ 3 พบว่ามีอัตราการรู้จำสูงถึง 96.6 % ซึ่งแสดงว่า
ตัวแบบดังกล่าวสามารถนำไปใช้ได้จริง

Thesis Title	Development of Lanna Characters Recognition System Using K-Nearest Neighbor
Author	Mr. Chaow Porkaew
Degree	Master of Science (Computer Science)
Thesis Advisor	Associate Professor Dr. Jeerayut Chaijaruwanich

ABSTRACT

This thesis presents a method for Lanna Characters Recognition from old Lanna machine-printed document image. The research aims to convert document image into machine readable text. The first step in the proposed method is characters segmentation from the image and extraction the feature of character image using Zoning Method, then classify the character image using K-Nearest Neighbor algorithm, after that construct a Lanna text. Using K-Nearest Neighbor algorithms, the first step is generating data training set by extract 113 types of characters image from the document images, then randomly selected 50 images for each character types for data training set. In order to evaluate the recognition rate is compare the recognition rates by adjust the normalize size of image and number of zones, which are the parameters of Zoning Method. The experimental results show that the best recognition rate is 96.6% after testing with 81×81 normalize size and 9×9 zones, which K of Nearest Neighbor is 3.