

ABBYY® FineReader

เวอร์ชัน 12

คู่มือการใช้งาน Pattern Training

เนื้อหา

Pattern Training คืออะไร	3
1. การใช้ user patterns	3
2. การสร้าง และการสอนโปรแกรมด้วย user pattern	5
3. การเลือกใช้ user pattern	7
4. การแก้ไข user pattern	8
ภาคผนวก	
1. ตัวอย่างการขั้นตอนการทำ Pattern Training เอกสารภาษาไทย	10

ABBYY FineReader 12 : Pattern Training

Pattern Training เป็นเครื่องมืออย่างหนึ่งของโปรแกรม ABBYY FineReader ที่ช่วยเหลือในการสอนโปรแกรมให้สามารถอ่านและรู้จำตัวอักษร หรือตัวอักษร ที่ไม่เป็นตามมาตรฐาน เช่น ตัวอักษรที่ใช้ในออกแบบ หรือ สัญลักษณ์ที่ใช้ในทางคณิตศาสตร์ เป็นต้น

ซึ่งในการทำงาน Patten Training นั้น จะทำการสร้าง user pattern ขึ้นมา แล้วทำการนำมาใช้เพื่ออ่านและรู้จำเอกสารที่เป็นตัวอักษร

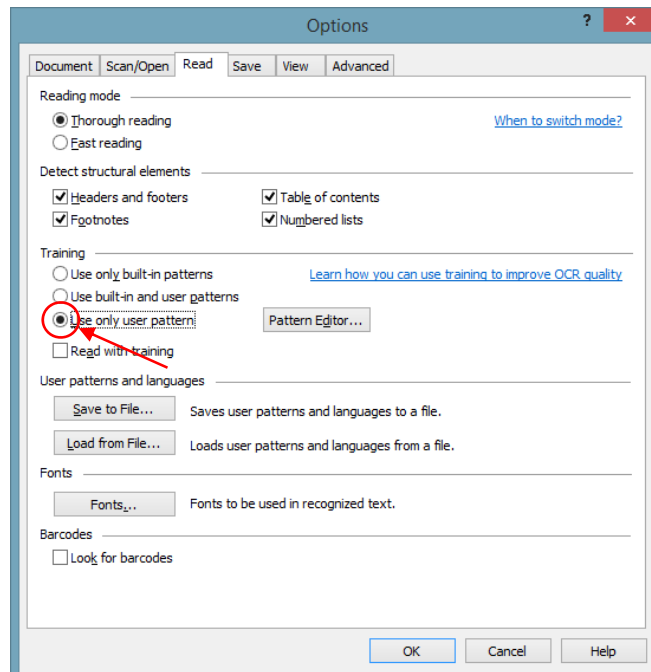
- การใช้ user patterns
- การสร้าง และการสอนโปรแกรมด้วย user pattern
- การเลือกใช้ user pattern
- การแก้ไข user pattern

1. การใช้ user patterns

การใช้ pattern ในการรู้จำเอกสาร

1.1 คลิกเมนู Tool > Option จะปรากฏหน้าต่าง Option ให้คลิกเลือกที่แท็บ Read

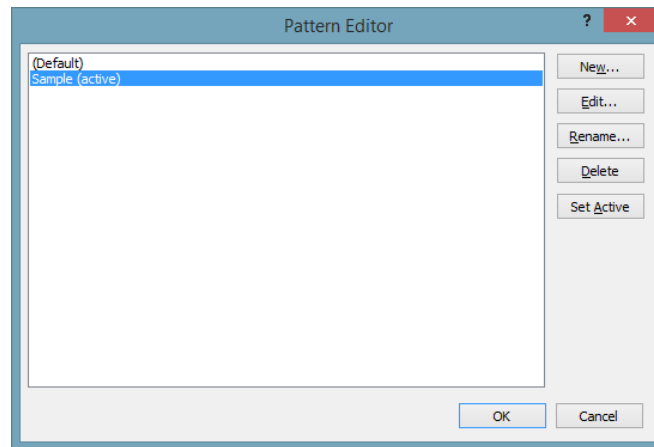
1.2 ภายใต้หัวข้อ Training ให้เลือก Use only user pattern



รูปภาพการเลือกใช้ Use only user pattern

1.3 คลิกที่ปุ่ม Pattern Editor

1.4 ในหน้าต่างของ Pattern Editor ให้เลือกรูปแบบต้องการ แล้วทำการคลิก Ok จากรูปตัวอย่าง ได้ทำการเลือกที่ Sample



รูปภาพตัวอย่างการเลือกใช้ pattern

1.5 ที่หน้าต่างโปรแกรมหลักของ ABBYY FineReader ให้คลิกที่ปุ่ม Read Page



รูปภาพตัวอย่างการคลิกปุ่ม Read Page

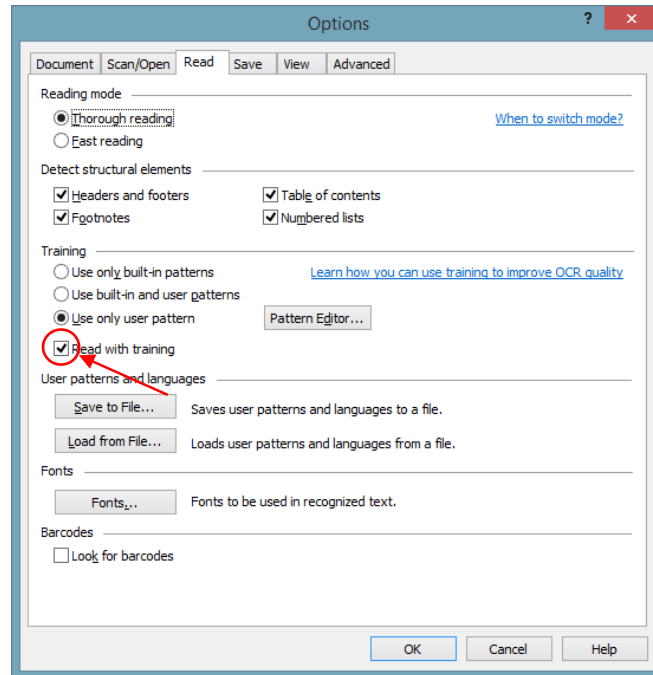
2. การสร้าง และการสอนโปรแกรมด้วย user pattern

การสอน user pattern ให้รู้จักตัวอักษร หรือตัวอักษรใหม่ กลุ่มของตัวอักษร

2.1 คลิกเมนู Tool > Option จะปรากฏหน้าต่าง Option ให้คลิกเลือกที่แท็บ Read

2.2 ภายใต้หัวข้อ Training ให้ทำการเลือก Use built-in and user pattern หรือ Use only user pattern

2.3 ให้ทำการเลือก Read with training

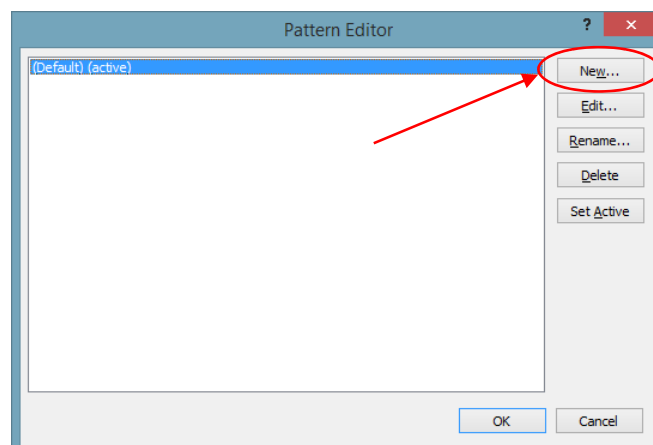


รูปภาพตัวอย่างในการเลือก Read with training

2.4 คลิกที่ปุ่ม Pattern Editor...

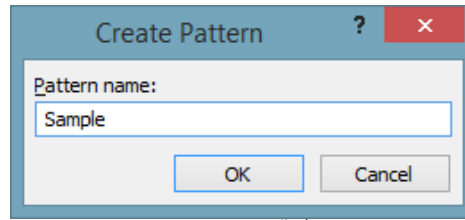
หมายเหตุ : Pattern Training ไม่รองรับภาษาแถบอาเซียน

2.5 ที่หน้าต่าง Pattern Editor ให้คลิกที่ปุ่ม New



รูปภาพตัวอย่างการสร้าง new pattern

2.6 จะปรากฏหน้าต่าง Create Pattern ให้ทำการตั้งชื่อของ user pattern แล้วคลิกปุ่ม Ok จากรูปตัวอย่างตั้งชื่อ Pattern name ว่า Sample

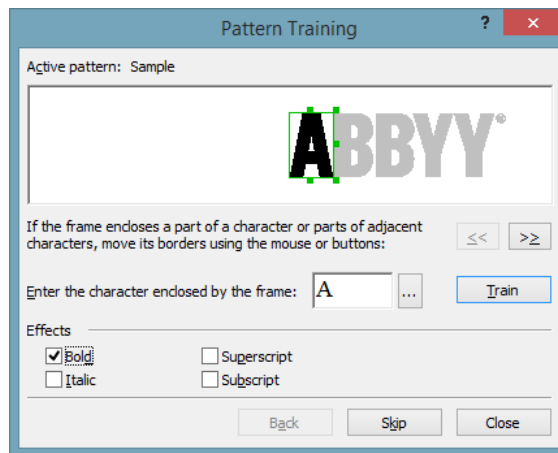


รูปภาพตัวอย่างในการตั้งชื่อ pattern

2.7 ทำการปิดหน้าต่าง Pattern Editor และที่หน้าต่าง Option ให้ทำการคลิกปุ่ม Ok

2.8 ที่แถบเครื่องมือของหน้าต่าง Image คลิกที่ Read และถ้าโปรแกรม ABBYY FineReader พบตัวอักษร หรือตัวอักษรที่ไม่รู้จัก ตัวอักษรตัวนั้นจะถูกแสดงในหน้าต่างของ Pattern Training

- ป้อนตัวอักษรที่สอดคล้องกับตัวอักษรที่แสดงอยู่ในช่องที่ใส่ตัวอักษร จากรูปภาพตัวอย่าง คือตัวอักษร "A"
- สามารถเลือกรูปแบบหรือลักษณะของตัวอักษรได้ที่ภายใต้หัวข้อ Effect จากรูปภาพตัวอย่าง ได้ทำการเลือกที่ Bold



รูปภาพตัวอย่างที่ป้อนอักษร และเลือก effect

2.9 ทำการสอนโปรแกรมให้อ่านตัวอักษร หรือตัวอักษรใหม่ และกลุ่มของตัวอักษร


กลุ่มของตัวอักษร คือ การรวมกันของสองหรือสามตัวอักษรที่ "เขียนติดๆ กัน รวมเข้าด้วยกัน" (ตัวอย่างเช่น fi, fl, ffi, etc.) และเป็นเรื่องยากสำหรับโปรแกรมที่จะแยก ในความเป็นจริงแล้ว หากจะให้ได้ผลลัพธ์ที่ดีกว่า โดยการจัดการรวมเป็นกลุ่มตัวอักษรเข้าด้วยกัน

หมายเหตุ : คำที่เป็นตัวหนา หรือตัวเอียง ในข้อความ หรือคำมีที่เขียนแบบยกขึ้นหรือเขียนข้างใต้ คุณสามารถเลือกใช้งาน effect ที่ต้องการ ภายใต้หัวข้อ Effect

การย้อนกลับไปยังตัวอักษรที่ได้ทำการสอนแล้ว ทำได้โดยการคลิกที่ปุ่ม Back โดยกรอบหน้าต่างจะข้ามไปยังตำแหน่งก่อนหน้านี้ และการสอนล่าสุด จะไม่ถูกบันทึกลงใน user pattern

ข้อสำคัญ!

- คุณสามารถสอนโปรแกรม ABBYY FineReader ให้สามารถอ่านตัวอักษรโดยใช้ตัวอักษรที่อยู่กลุ่มของตัวอักษรที่โปรแกรมรองรับเท่านั้น ในกรณีที่ต้องการสอนโปรแกรมให้อ่านอักษร ที่ไม่ปรากฏอยู่บนคีย์บอร์ดนั้น สามารถสร้างโดยการรวมตัวอักษรมากกว่าหนึ่งเข้าด้วยกัน เช่น ตัวอักษรญี่ปุ่น หรือการตัดลอกตัวอักษรจาก

Insert Character (โดยสามารถคลิกได้ที่ "  ")

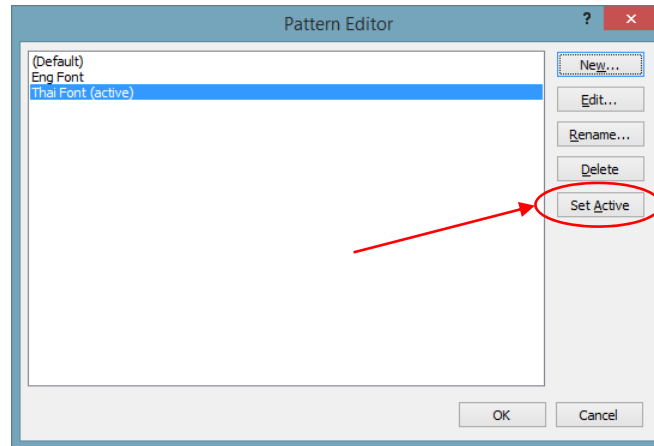
- Pattern ของแต่ละแบบ อาจจะมีมีความยาวของตัวอักษรใหม่ได้ถึง 1000 อักษร อย่างไรก็ตาม ควรที่จะหลีกเลี่ยงในการสร้างกลุ่มของตัวอักษรมากเกินไป เพราะอาจส่งผลกระทบต่อคุณภาพในการ OCR ได้

3. การเลือกใช้ user pattern

โปรแกรม ABBYY FineReader อนุญาตให้คุณสามารถใช้ pattern เพื่อปรับปรุงคุณภาพในการ OCR

3.1 เลือกที่เมนู Tool คลิกเลือกที่ Pattern Editor...

3.2 ที่หน้าต่าง Pattern Editor ให้เลือก pattern จากการรายการที่แสดงอยู่ และคลิกที่ปุ่ม Set Active



รูปภาพตัวอย่างการเลือกใช้ pattern และ Set Active

จุดสำคัญที่ต้องจดจำ :

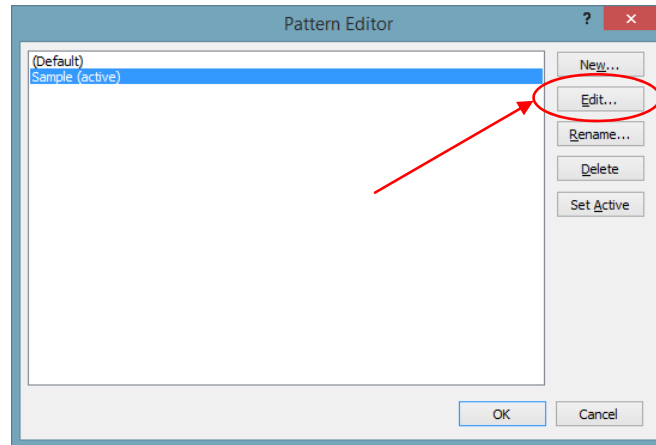
- 1) แทนที่จะแสดงความแตกต่างกันในบางตัวอักษรที่แตกต่างกัน แต่กลับกลายเป็นว่า ABBYY FineReader อาจตระหนักว่าตัวอักษรเหล่านั้นเป็นตัวอักษรเดียวกัน ตัวอย่างเช่น เครื่องหมายเส้นตรง (|), เครื่องหมายคำพูดซ้าย (‘) และขวา (’) จะถูกเก็บไว้ในรูปแบบที่เป็นตัวอักษรตัวเดียว ซึ่งหมายความว่า เครื่องหมายคำพูดซ้าย และขวาจะไม่ถูกนำมาใช้ในข้อความที่ได้อ่านแล้ว แม้ว่า คุณจะพยายามที่จะสอนโปรแกรม
- 2) สำหรับบางภาพตัวอักษรนั้น ABBYY FineReader จะทำการเลือกตัวอักษรที่สอดคล้องตามบริบทโดยรอบ ตัวอย่างเช่น ภาพของวงกลมขนาดเล็ก จะถูกรู้จำเป็นเป็นตัวอักษร "o" ถ้ามีตัวอักษรอยู่เป็นตัวอยู่ถัดไป และเป็นตัวเลข "0" ถ้ามีตัวเลขเป็นตัวอยู่ถัดไป
- 3) Pattern สามารถใช้ได้เพียงเอกสารที่มีตัวอักษร, ขนาดตัวอักษร และความละเอียดของเอกสารที่ใช้เหมือนกัน ที่ใช้ในการสร้าง pattern
- 4) เพื่อให้สามารถใช้งาน pattern ได้ หลังจากที่ได้ทำการจัดเก็บแล้ว สามารถดูข้อมูลเพิ่มเติมได้ที่หัวข้อ "What Is a FineReader Document?"
- 5) ถ้าต้องการกลับใช้ค่า default ของโปรแกรม คือ "Use only built-in pattern" สามารถทำได้โดยการไปที่เมนู Tool > Option > แท็บ Read และคลิกเลือกที่ Use only built-in pattern

4. การแก้ไข user pattern

ถ้าคุณต้องการแก้ไข pattern ของคุณได้ทำการสร้างขึ้นมาก่อนที่จะนำไปใช้ในการประมวลผล OCR นั้น สามารถทำได้โดยการ

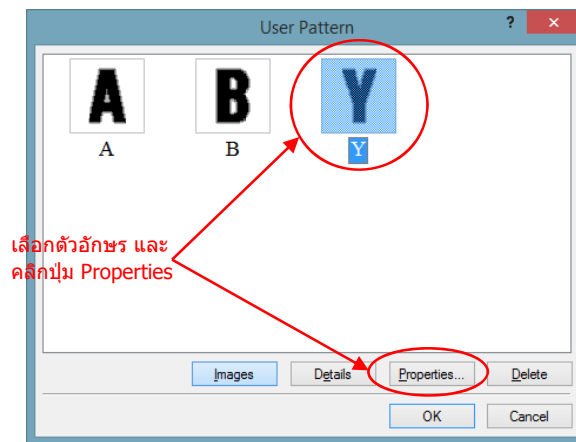
4.1 เลือกที่เมนู Tool แล้วคลิกเลือกที่ Pattern Editor

4.2 เมื่อหน้าต่าง Pattern Editor ปรากฏ ให้เลือก pattern จากรายการที่แสดงอยู่ และคลิกที่ปุ่ม Edit



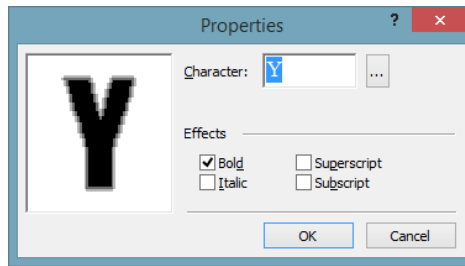
รูปภาพตัวอย่างการเลือก Edit

4.3 เมื่อหน้าต่าง User Pattern ปรากฏ ให้เลือกตัวอักษร และคลิกที่ปุ่ม Properties ดังภาพตัวอย่าง เลือกตัวอักษร "Y"



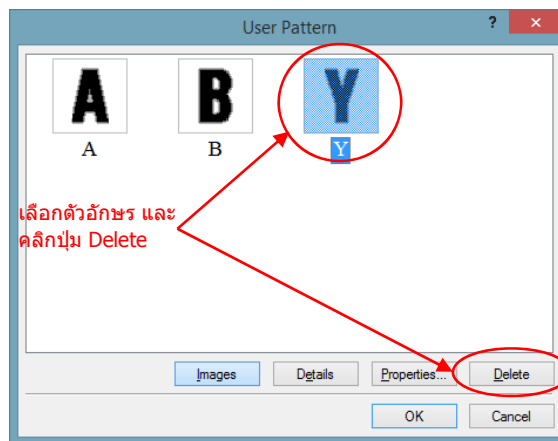
รูปภาพตัวอย่างในการเลือกตัวอักษร เพื่อแก้ไข

ก็จะปรากฏหน้าต่าง Properties ของตัวอักษรได้ทำการเลือก เพื่อสำหรับในการแก้ไข



รูปภาพตัวอย่างของหน้าต่าง properties


ที่หน้าต่าง User Pattern สามารถลบตัวอักษรที่ทำการสอนไปแล้วได้ โดยทำการเลือกที่ตัวอักษร และคลิกที่ปุ่ม Delete

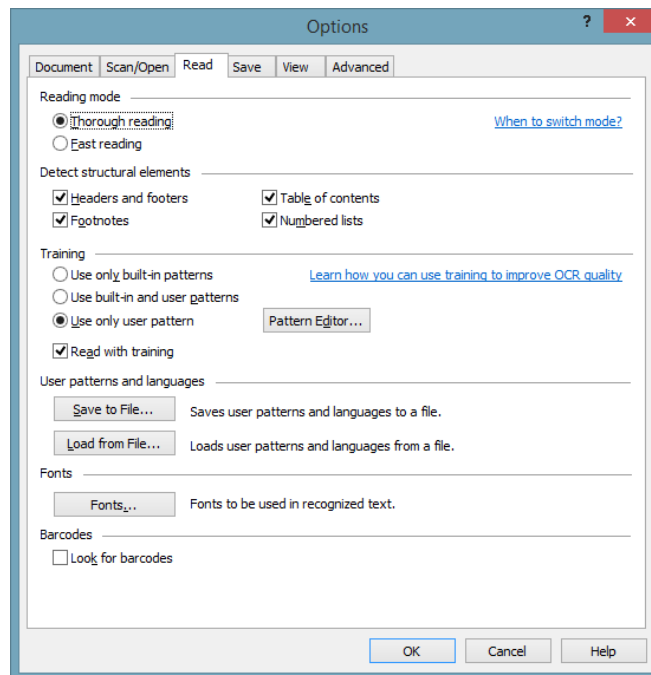


รูปภาพตัวอย่างในการเลือกตัวอักษร เพื่อลบ

ภาคผนวก

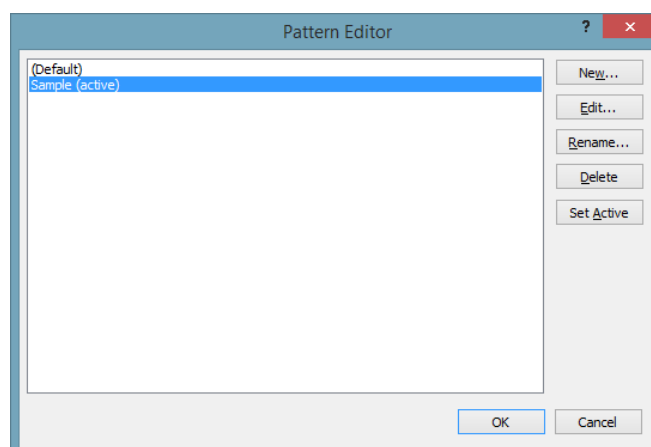
ตัวอย่างการขั้นตอนการทำ Pattern Training เอกสารภาษาไทย

1. เปิดโปรแกรม ABBYY FineReader 12 ทำการคลิกที่ Open "  " เพื่อเลือกไฟล์เอกสารที่ต้องการจะทำการ OCR
2. เลือกที่เมนู Tool > Option > และคลิกเลือกที่ แท็บ Read
3. คลิกเมนู Tool > Option จะปรากฏหน้าต่างต่าง Option ให้คลิกเลือกที่แท็บ Read
4. ภายใต้วินิจฉัย Training ให้เลือก Use only user pattern และ Read with pattern



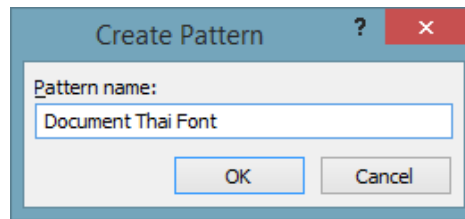
รูปภาพตัวอย่างในการเลือก Read with training

5. แล้วคลิกที่ปุ่ม Pattern Editor...
6. ที่หน้าต่าง Pattern Editor ให้คลิกที่ปุ่ม New...



รูปภาพตัวอย่างการสร้าง new pattern

7. จะปรากฏหน้าต่าง Create Pattern ให้ทำการตั้งชื่อของ user pattern แล้วคลิกปุ่ม Ok จากรูปตัวอย่างตั้งชื่อ Pattern name ว่า "Document Thai Font"



รูปภาพตัวอย่างในการตั้งชื่อ pattern

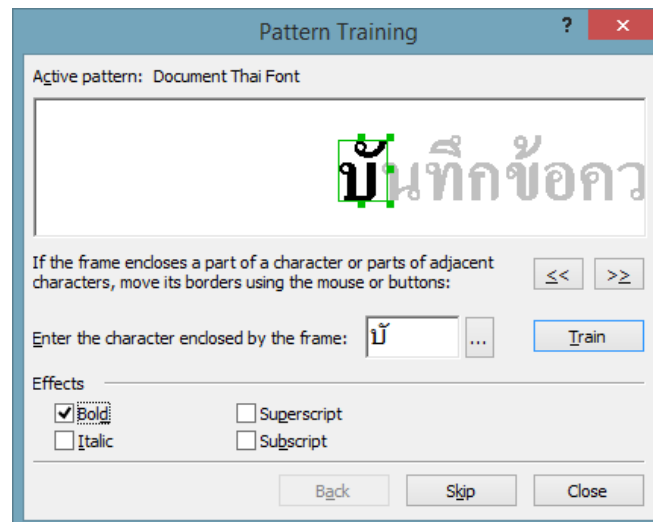
8. ทำการปิดหน้าต่าง Pattern Editor และที่หน้าต่าง Option ให้ทำการคลิกปุ่ม Ok
9. ที่แถบเครื่องมือของหน้าต่าง Image คลิกที่ Read Page



รูปภาพตัวอย่างการคลิกปุ่ม Read Page

10. และถ้าโปรแกรม ABBYY FineReader พบตัวอักษร หรือตัวอักษรๆ ที่ไม่รู้จัก ตัวอักษรตัวนี้จะถูกแสดงในหน้าต่างของ Pattern Training
- ✓ ป้อนตัวอักษรที่สอดคล้องกับตัวอักษรที่แสดงอยู่ ในช่องที่ใส่ตัวอักษร จากรูปภาพตัวอย่าง คือ ตัวอักษร "บ้"
 - ✓ สามารถเลือกรูปแบบหรือลักษณะของตัวอักษรได้ที่ภายใต้หัวข้อ Effect จากรูปภาพตัวอย่าง ได้ทำการเลือกที่ Bold

หลังจากที่ได้ทำการใส่ข้อมูลและเลือก effect เรียบร้อยแล้ว ให้คลิกที่ปุ่ม Train



รูปภาพตัวอย่างที่ป้อนอักษร และเลือก effect

11. จากนั้นทำการ train อักษร หรือตัวอักษรจนครบตามความต้องการ โดยโปรแกรมจะทำการบันทึกการ train ทั้งหมดเอาไว้ภายใต้ชื่อของ pattern ที่ได้สร้างและเลือกใช้งานอยู่