מפרט בדיקות

HOCR

2011

Maxim Drabkin

Niv Maman

*אוניברסיטת חיפה*

*החוג למדעי המחשב*

# תוכן העניינים

[בדיקות כלליות 2-3](#_Toc298612350)

[בדיקות עיבוד תמונה 3](#_Toc298612351)

[בדיקות זיהוי טקסט 5](#_Toc298612352)

# בדיקות כלליות

1. הפעלת התוכנית בסביבת Windows (כיסוי דרישה מספר 2)

תיאור הבדיקה:

הרצה של התוכנית בסביבת מערכת הפעלה של Windows

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. דפדף במערכת ההפעלה ובחר בתמונה
4. לחץ Recognize -> Get Text

תוצאות צפויות:

* המערכת צריכה להחזיר את הטקסט שבתמונה.
* המערכת שומרת בקובץ טקסט את התוצאה, תחת התיקייה Results.

תוצאה: עבר

1. הפעלת התוכנית בסביבת Linux(כיסוי דרישה מספר 3)

תיאור הבדיקה:

הרצה של התוכנית בסביבת מערכת הפעלה של Linux

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. דפדף במערכת ההפעלה ובחר בתמונה
4. לחץ Recognize -> Get Text

תוצאות צפויות:

* המערכת צריכה להחזיר את הטקסט שבתמונה.
* המערכת שומרת בקובץ טקסט את התוצאה, תחת התיקייה Results.

תוצאה: עבר

1. זיהוי קובץ PDF (כיסוי דרישה מספר 5)

תיאור הבדיקה:

בדיקת הפלט כאשר ניתן קובץ PDF כקלט ולא תמונה

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. דפדף במערכת ההפעלה ובחר בקובץ PDF
4. בחר את העמוד הרצוי בקובץ בעזרת כפתורי הדפדוף
5. לחץ Recognize -> Get Text

תוצאות צפויות:

* המערכת צריכה להחזיר את הטקסט שבקובץ ה- PDF.
* המערכת שומרת בקובץ טקסט את התוצאה, תחת התיקייה Results, כאשר לכל עמוד יש קובץ טקסט משל עצמו.

תוצאה: עבר

# בדיקות עיבוד תמונה

1. הפעלת פילטר סיבוב על התמונה (כיסוי דרישות מספר 6,7,8,9)

תיאור הבדיקה:

ביצוע פילטר סיבוב על תמונה על מנת לאפשר זיהוי מקסימלי

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. בחר בתמונה rotate.bmp אשר נמצאת בתיקיית ה- Test שסופקה יחד עם התוכנה.
4. לחץ Filtering -> Rotate
5. המתן עד שהטקסט בתמונה ייראה מיושר ביחס לציר האופקי.

תוצאות צפויות:

* המערכת תיישר את התמונה תוך חצי דקה.
* התמונה שמופיעה בחלון השמאלי של התוכנה תראה טקסט מיושר.

תוצאה: עבר

1. הפעלת פילטר ניקוי על התמונה (כיסוי דרישות מספר 6,10,11,12)

תיאור הבדיקה:

ביצוע פילטר ניקוי על תמונה על מנת לאפשר זיהוי מקסימלי

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. בחר בתמונה TimesTest2.jpg אשר נמצאת בתיקיית ה- Test שסופקה יחד עם התוכנה.
4. לחץ Filtering -> Clean
5. המתן עד שהפעולה תגמר (ראה סטטוס בסטטוס בר התחתון)

תוצאות צפויות:

* המערכת תבצע ניקוי של התמונה תוך 10 שניות.
* התמונה שמופיעה בחלון השמאלי של התוכנה תראה נקייה יותר מרעשים שהופיעו בתמונה המקורית.

תוצאה: עבר

1. הפעלת פילטר "סף" על התמונה (כיסוי דרישות מספר 6,13)

תיאור הבדיקה:

ביצוע פילטר "הפרדת סף" על תמונה על מנת לאפשר זיהוי מקסימלי

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. בחר בתמונה צבעונית.
4. לחץ Filtering -> Threshold
5. המתן עד שהפעולה תגמר (ראה סטטוס בסטטוס בר התחתון)

תוצאות צפויות:

* המערכת תבצע חלוקת סף לפיקסלים שחורים ולבנים בלבד.

תוצאה: עבר

1. הפעלת פילטר מרכוז על התמונה (כיסוי דרישות מספר 6,14)

תיאור הבדיקה:

ביצוע פילטר מרכוז על תמונה על מנת לאפשר זיהוי מהיר יותר

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. בחר בתמונה rotate.bmp אשר נמצאת בתיקיית ה- Test שסופקה יחד עם התוכנה.
4. לחץ Filtering ->Rotate
5. המתן עד שהפעולה תגמר (ראה סטטוס בסטטוס בר התחתון)
6. לחץ Filtering -> Centerize
7. המתן עד שהפעולה תגמר (ראה סטטוס בסטטוס בר התחתון)

תוצאות צפויות:

* לאחר השלב החמישי התמונה תהיה מסובבת, אך הטקסט יהיה "מרופד" בפיקסלים לבנים.
* לאחר השלב השביעי, הטקסט יהיה מרופד בסביבה מינימלית של של פיקסלים לבנים, וכל הפיקסלים הלבנים המיותרים ייעלמו.

תוצאה: עבר

# בדיקות זיהוי טקסט

1. זיהוי אותו טקסט בפונטים שונים (כיסוי דרישות מספר 18,15)

תיאור הבדיקה:

הפעלת התוכנה על אותו הטקסט, רק בפונטים שונים וקבלת אותו קלט

שלב מקדים:

* למד את התוכנה את הפונט Ariel
* למד את התוכנה את הפונט David
* צור תמונה עם הטקסט "זיהוי טקסט בפונט שונה" בפונט Ariel (נקרא לתמונה FindArial.jpeg)
* צור תמונה עם הטקסט "זיהוי טקסט בפונט שונה" בפונט David (נקרא לתמונה FindDavid.jpeg)

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. בחר בתמונה FindAriel.jpeg אשר יצרת בשלב המקדים
4. לחץ Recognize -> Get Text
5. לחץ Recognize -> Select Image
6. בחר בתמונה FindDavid.jpeg אשר יצרת בשלב המקדים
7. לחץ Recognize -> Get Text

תוצאות צפויות:

* לאחר השלב הרביעי, על המערכת להחזיר את הטקסט "זיהוי טקסט בפונט שונה"
* לאחר השלב השישי, על המערכת להחזיר את הטקסט "זיהוי טקסט בפונט שונה".

תוצאה: עבר

1. קביעת הפונט הנכון בצורה ידנית (כיסוי דרישות מספר 16,18,19)

תיאור הבדיקה:

בדיקה חיובית- המשתמש קובע את הפונט הנכון ואז מריץ זיהוי. הזיהוי אמור להצליח

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. בחר בתמונה ArialTest.bmp אשר נמצאת בתיקיית ה- Test שסופקה יחד עם התוכנה.
4. בחר בפונט Ariel מתיבת הפונטים הקיימים.
5. לחץ Recognize -> Get Text

תוצאות צפויות:

* המערכת מזהה את הטקסט בצורה נכונה בתמונה.

תוצאה: עבר

1. קביעת הפונט הלא נכון בצורה ידנית (כיסוי דרישות מספר 16,18,19)

תיאור הבדיקה:

בדיקה שלילית- המשתמש קובע את הפונט הלא נכון ואז מריץ זיהוי. הזיהוי אמור להיות שגוי

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. בחר בתמונה ArialTest.bmp אשר נמצאת בתיקיית ה- Test שסופקה יחד עם התוכנה.
4. בחר בפונט David מתיבת הפונטים הקיימים.
5. לחץ Recognize -> Get Text

תוצאות צפויות:

* המערכת מזהה את הטקסט אך הזיהוי לא תמיד נכון, מכיוון שהשתמשנו ברשת ניורונים אחרת.

תוצאה: עבר

1. קביעת הפונט ע"י המערכת (כיסוי דרישות מספר 17,18,19,20,21)

תיאור הבדיקה:

המשתמש לא קובע את הפונט ואז מריץ זיהוי. הזיהוי אמור להצליח במידה והפונט קיים

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. בחר בתמונה ArialTest.bmp אשר נמצאת בתיקיית ה- Test שסופקה יחד עם התוכנה.
4. בחר באפשרות Auto מתיבת הפונטים הקיימים.
5. לחץ Recognize -> Get Text
6. חזור על התהליך מסעיף ב' עבור תמונה אחרת.

תוצאות צפויות:

* המערכת מזהה את הטקסט בצורה נכונה בתמונה.
* המערכת מציגה שהפונט בתמונה הוא מסוג Arial.
* התיקיה Results יופיע שם התמונה בתצורת txt ובה טקסט שזוהה.

תוצאה: עבר

1. תיקון פונט (כיסוי דרישות מספר 24,26)

תיאור הבדיקה:

נתקן פונט שגוי ונבדוק שלאחר תיקון ואימון התוכנה לא טועה עבור האות שתיקנו.

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Recognize -> Select Image
3. בחר בתמונה חדשה שבה אותיות שטרם נלמדו.
4. בחר בפונט שבו הטקסט נכתב.
5. לחץ Recognize -> Get Text
6. כעת סמן אות שבה טעות, הקלד את האות הנכונה ובחר ב Repair.
7. בחר ב Training->Train Network
8. התחל אימון ע"י לחיצה על Start והמתן שאימון יסתיים.
9. כעת חזור על סעיף ה'.

תוצאות צפויות:

* המערכת מזהה את הטקסט בצורה נכונה בתמונה.

תוצאה: עבר

1. יצירת פונט חדש וטעינה במחשב אחר (כיסוי דרישות מספר 25,27)

תיאור הבדיקה:

ניצור פונט חדש, נאמן אותו, נשמור ונטען במחשב אחר.

שלבי הבדיקה:

1. הפעל את התוכנה.
2. לחץ Training -> Create Font -> Empty
3. ניתן לו שם TestFont
4. נבחר את הפונט שיצרנו.
5. לחץ Recognize -> Select Image
6. כעת נטען טקסט שכתוב בפונט שיצרנו.
7. נבחר Recognize->Get Text
8. נתקן ידנית את הפונט ע"י סימון האות בטקסט התוצאה, הכנסת האות הנכונה ובחירה ב Repair
9. לאחר שסיימנו לתקן את הפונט נבחר ב Training->Train Network
10. נלחץ START ונמתין שהאימון יסתיים.
11. כעת הפונט נשמר לקובץ .net שנמצא בתיקייה Fonts. נעתיק אותו למחשב אחר ונטען בו את התוכנה מחדש כאשר הקובץ נמצא בתיקייה ה Fonts של התוכנה.
12. נבחר בפונט.
13. לחץ Recognize -> Select Image
14. נטען טקסט שכתוב בפונט שיצרנו.
15. נבחר Recognize->Get Text
16. נבדוק האם טקסט התוצאה תואם את התמונה שטענו.

תוצאות צפויות:

* המערכת מזהה את הטקסט בצורה נכונה בתמונה.

תוצאה: עבר