

**פרויקט מסכם לתואר בוגר במדעים (B.Sc)
במתמטיקה שימושית**

מקדם מתאם מנורמל

נדב וגנר

Normalized Cross-Correlation

Nadav Wagner

**פרויקט מסכם לתואר בוגר במדעים (B.Sc)
במתמטיקה שימושית**

מקדם מתאם מנורמל

נדב וגנר

Normalized Cross-Correlation

Nadav Wagner

Advisor:

Dr. Samuel Kosolapov

מנחה:

ד"ר שמואל קוסולאפוב

Karmiel

2014

כרמיאל

תקציר

הפרויקט בוצע במסגרת פרויקט גמר במחלקת מתמטיקה שימושית במכללת אורט בראודה. הפרויקט הינו פרויקט הרחבה לפרויקט גמר שעשיתי במחלקת הנדסת חשמל. הפרויקט עוסק במודל מקדם מתאם מנורמל (Normalized Cross Correlation) ושימושיו בעיבוד תמונה. בהתחלה נציג את ההיסטוריה של המודל והפיתוחים המתמטיים. לאחר מכן נציג 3 שימושים עיקריים של המודל בעיבוד תמונה, מציאת אובייקט, עקיבה אחר אובייקט והפרויקט בהנדסת חשמל שעסק בזיהוי ספרות. הפרויקט של זיהוי ספרות עסק בעיצוב ובנייה של מערכת משוב מהירה (FFS – Fast Feedback System) אשר תפעל כיישום אנדרואיד לשליחת תמונה של מבחן אל שרת, אשר יבצע עיבוד תמונה ויפענח את תשובת המבחן.

בפרויקט פותח יישום אנדרואיד בשפת תכנות Java + XML. היישום כולל 3 חלקים :

1. לקיחת תמונה - אחראית לפנות להפעיל את המצלמה ולשמור את התמונה בשם ספציפי.
2. כיווץ תמונה - אחראית על המרת התמונה לגודל סטנדרטי 1024x768.
3. שליחת תמונה - אחראית על המרת הקובץ לקבצים ושליחתם דרך DATA/WIFI אל שרת PHP והמתנה לתשובה כדי להציג את התשובה.

בפרויקט פותח שרת PHP, אשר מקבל את התמונה מהטלפון החכם. השרת מותקן על מחשב אישי על-ידי תוכנת WAMP Server. פיתוח סקריפט בשפת תכנות PHP אשר מקבל את התמונה ומפעילה את תוכנת Matlab אשר מבצע עיבוד תמונה.

בפרויקט נעשה עיבוד תמונה על-ידי תוכנת Matlab. הסקריפט הראשון מבצע הזזה ומתקן את הקנה מידה של התמונה. הסקריפט השני חותך את התמונה ספרות לצורך זיהוי בסיום משווה מול תבנית של המרצה לחישוב הצלחה כישלון של התשובה בסיום מחזיר תשובה לטלפון החכם.

הקדשות ותודות

ברצוני להודות למנחה, ד"ר שמואל קוסאלפוב, שהנחה אותי בכל שלבי הפרויקט ועל כך שליווה את עבודתי במסירות ומקצועיות רבה, היה תמיד זמין לשאלות וידע לכווני בדרך הטובה ביותר.

תוכן עניינים:

III	תקציר
III	הקדשות ותודות
IV	תוכן עניינים
VI	רשימת איורים
VI	רשימת טבלאות
1	1. מבוא
1	1.1. עיבוד תמונה
2	1.2. פיקסלים
2	1.3. זיהוי מטרות אוטומטי
3	1.4. מקדם מתאם של פירסון
3	1.5. מקדם מתאם מנורמל
3	1.6. הסבר על אנדרויד (Android)
3	1.7. הסבר על שרת-לקוח (Client-Server)
4	1.8. הסבר על מחשב בענן (Cloud computing)
4	1.9. הסבר על מחלקה (Class)
5	2. סקר ספרות
6	3. הצגת משוואת מקדם מתאם
7	3.1. חסרונות בשימוש במקדם מתאם דיסקרטי
7	3.2. הצגת מקדם מתאם מנורמל
7	3.3. מימוש אלגוריתם מקדם מתאם מנורמל
8	4. שימושים במקדם מתאם מנורמל
8	4.1. מציאת אובייקט
8	4.2. עקיבה אחר אובייקט
9	4.3. זיהוי ספרות
9	4.3.1. מפרט פונקציונאלי
9	4.3.2. תרשים מלבנים כולל של המערכת
10	4.3.3. אפליקציית אנדרויד
10	4.3.3.1. שלב תכנון בניית האפליקציה
10	4.3.3.2. לקיחת תמונה
11	4.3.3.3. הורדת קנה מידה לתמונה
11	4.3.3.4. העלאת תמונה לשרת והמתנה לתשובה
12	4.3.3.5. דיאגרמת מחלקה – Class Diagram
12	4.3.4. טופס תשובות
12	4.3.4.1. תכנון טופס תשובות
13	4.3.4.2. הוראות צילום ומילוי טופס בחינה
13	4.3.4.3. טופס הבחינה

13		4.3.5 שרת PHP
13		4.3.5.1 תכנון שרת PHP
14		4.3.6 עיבוד תמונה
14		4.3.6.1 סיבוב ותיקון קנה מידה
15		4.3.6.2 חיתוך וזיהוי
17		5. סיכום ודיון
17		5.1 סיכום
17		5.2 מסקנות
18		6. סימוכין

רשימת איורים

2	איור 1 – תמונה בגווני אפור ומטריצה דו-מימדית
2	איור 2 – תמונה צבעונית ומטריצה תלת-מימדית
2	איור 3 – מערכת ATR טיפוסית
6	איור 4 – מציאת מרחק על ידי משוואת מרחק אוקלידית בריבוע
6	איור 5 – התאמת תבנית לתמונה
8	איור 6 – מציאת אובייקט בתוך תמונה
9	איור 7 – עקיבה אחר אובייקט
10	איור 8 – תרשים זרימה של מכלולי הפרויקט
10	איור 9 – תרשים זרימה של האפליקציה
11	איור 10 – מסך ראשון
11	איור 11 – אפשר צילום/בחירה
11	איור 12 – הודעת שגיאה
11	איור 13 – מסך שני
12	איור 14 – מסך שלישי
12	איור 15 – העלאת תמונה
12	איור 16 – תשובה מהשרת
12	איור 17 – דיאגרמת מחלקות
13	איור 18 – אפשרויות הכתיבה
13	איור 19 – טופס הבחינה
14	איור 20 – תרשים זרימה של השרת
14	איור 21 – אלגוריתם SIFT לתיקון סיבוב ויחס קנה מידה
15	איור 22 – דרך פעולה של Median פילטר 3x3
15	איור 23 – אלגוריתם מציאת ספרה

רשימת טבלאות

16	טבלה 1 – תוצאות זיהוי ספרת
16	טבלה 2 – טבלת זיהוי הסתברות והטעויות הנפוצות לכל ספרה

1. מבוא

פרויקט זה הוא פרויקט הרחבה מתמטית לפרויקט גמר בהנדסת חשמל. כיום קיימות שיטות רבות לבדיקה אוטומטית של טפסי תשובות ומתן ציון כאשר הדרך הפופולרית ביותר היא השאלון האמריקאי (Multiple Choice Question). אציג כאן שיטה חדשה של מערכת משוב מהירה (Fast Feedback System - FFS), אשר תפעל על עקרון זיהוי ספרות מהצורה של Seven Segment.

כיום בעולם הדיגיטלי עם הנגישות הגבוהה לטלפונים חכמים שכולם בעלי גישה לאינטרנט, ניתן לנצל זאת כדי לפתח מערכת אוטומטית ללקיחת תמונה של טופס התשובות ובכך לקבל באופן אוטומטי את תוצאות המבחן.

מטרת הפרויקט בהנדסת חשמל היא לפתח אפליקציה אנדרויד ללקיחת תמונה ושליחתה לשרת חיצוני אשר יבצע עיבוד תמונה. העיבוד תמונה יתבסס על Normalized 2-D Cross-Correlation ויחזיר תשובה לטלפון החכם בצורת ציון. העבודה נעשתה בכמה שפות תכנות:

- אפליקציית אנדרויד : Java ו-XML.
- שרת PHP : PHP.
- עיבוד תמונה : Matlab.

מטרת הפרויקט במתמטיקה שימושית היא להראות את הפיתוח המתמטי של המודל מקדם מתאם מנורמל ואיך ניתן להשתמש בו בעיבוד תמונה בדגש למערכת משוב מהירה FFS.

1.1. עיבוד תמונה

עיבוד תמונה הינו תחום המשלב מתמטיקה, הנדסת חשמל ואלקטרוניקה והנדסת תוכנה. התחום עוסק בעיבוד של תמונות ובמניפולציות שונות שניתן לבצע עליהם. התמונה נדגמת והופכת מאות רציף לאות דיגיטלי (בדיד) ועליו נעשה העיבוד.

עיבוד תמונה נעשה לשתי מטרות עיקריות:

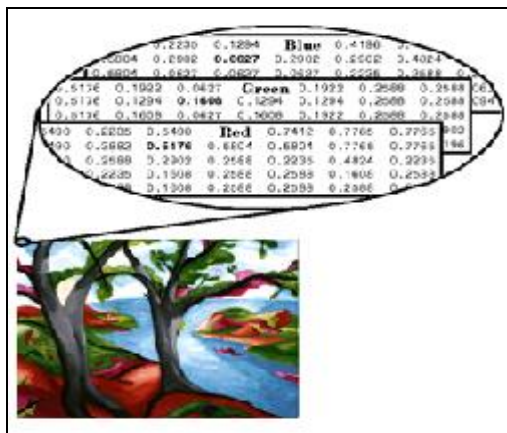
- שיפור התמונה: הורדת "רעשים", הדגשת פרטים, חידוד, תיקון רמות אפור, תיקון צבעים וכדומה.
- הסקת מסקנות: מדידות, זיהוי פגמים, זיהוי צורות.

עיבוד תמונה זהו פעולה אשר לוקחת תמונה כפלט ומוציאה כפלט תמונה. בדרך כלל תמונה תוגדר

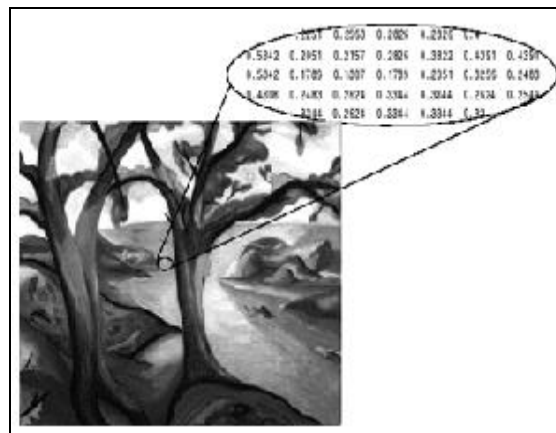
$$\{f: \mathbb{R}^2 \rightarrow \mathbb{R}^2\}$$

- כאשר $f(x, y)$ מגדיר לנו את העוצמה בנקודה (x, y) .
- מוגדר באזור מרובע עם גודל סופי $f: [a, b] \times [c, d] \rightarrow [0, 1]$.
- תמונה צבעונית מוגדרת $f(x, y) = (f_r(x, y), f_g(x, y), f_b(x, y))$.

כל תמונה לאחר ההמרה לאות דיגיטלי מורכבת מנקודות זעירות הקרויות פיקסלים. כל תמונה מוגדרת כמטריצה תלת מימדית במקרה של תמונה צבעונית או מטריצה דו מימדית במקרה של תמונה בגוון אפור או שחור לבן.



איור 2 – תמונה צבעונית ומטריצה תלת-מימדית



איור 1 – תמונה בגוון אפור ומטריצה דו-מימדית

1.2. פיקסלים

בעיבוד תמונה פיקסל או אלמנט בתמונה זהו גודל פיסי או עוצמה של הגודל הקטן ביותר בתמונה. הכתובת של הפיקסל מציינת את המיקום הפיזי בתמונה. כל פיקסל זהו דגימה של האות המקורי כאשר כל פיקסל משנה את עוצמת הפיקסל בהתאם לאות המקורי.

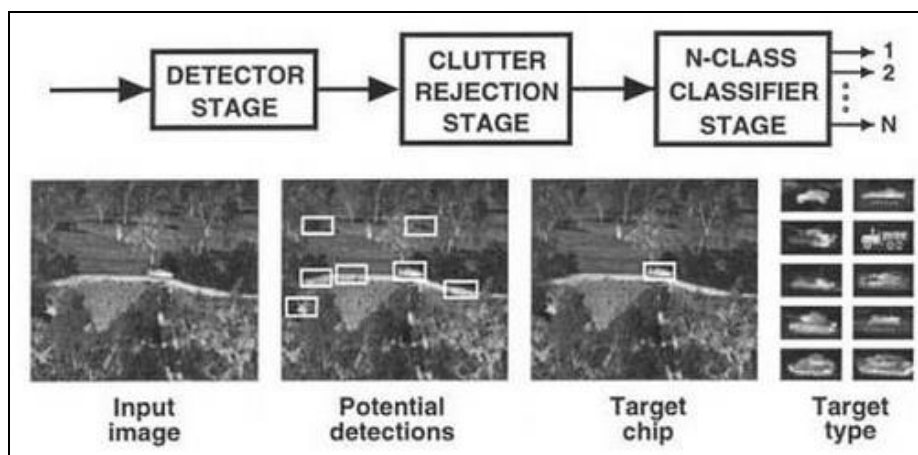
נתייחס לעוצמת הפיקסלים בצורה הבאה :

בתמונה צבעונית או אפורה $[0, 255]$, בתמונה שחור לבן $[0, 1]$.

1.3. זיהוי מטרות אוטומטי

הפיתוח של מערכות לזיהוי מטרות אוטומטי (ATR-automatic target recognition) צריך להתמודד עם אתגרים רבים כגון : זיהוי מטרות מרובה, אזורים גאוגרפים שונים, תנאי מזג האוויר, סיבוב, קנה מידה ועוד בעיות רבות.

כל מערכת צריכה להיות מותאמת לזיהוי מטרות ספציפי כדי למנוע זיהוי שגוי. כל מערכת ATR כוללת בתוכה עיבוד מקדים, חלוקה, הבלטת תוכנות, סיווג, תעדוף ומעקב. מסך כל המכלולים האלו אנו נתמקד בעיקר במודלים של סיווג וזיהוי שמוצגים באיור 3 במקרה הכללי.



איור 3 - מערכת ATR טיפוסית

1.4. מקדם מתאם של פירסון

מקדם מתאם של פירסון הוא מדד למתאם לינארי בין שתי קבוצות X ו- Y של מספרים ונותן תוצאה בין $[-1, 1]$:

• כאשר $+1$ קיים קשר חיובי בין 2 המשתנים.

• כאשר 0 לא קיים קשר בין שני המשתנים.

• כאשר -1 קיים קשר שלילי בין 2 המשתנים.

השימוש העיקרי הוא להראות קשר לינארי בין 2 משתנים. לאחר קבלת המקדם של פירסון נקבל מידע על עוצמת הקשר, ככל שנתקרב לערכים $+1$, -1 העוצמה חזקה יותר ובנוסף נקבל מידע על כיוון הקשר חיובי או שלילי.

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)} \sqrt{E(Y^2) - E^2(Y)}}$$

במקרה ויש לנו דגימה, כאשר ידועים לנו את כל הנתונים ניתן להשתמש בנוסחה באופן הבא :

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

1.5. מקדם מתאם מנורמל

בעולם העיבוד תמונה מקדם מתאם פירסון נקרא מקדם מתאם מנורמל (NCC-Normalized Cross-Correlation), משתמשים בשיטה זו למציאת דמיון או חוסר דמיון בין 2 תמונות או אותות. Normalized Cross Correlation הוא כלי סטטיסטי למציאת דמיון בין 2 סיגנלים. כיום בעולם הדיגיטלי יש צורך ממשי באלגוריתמים מדויקים ומהירים להתאמת תבניות למימוש מגוון אפליקציות בעיבוד תמונה. כמו מקדם פירסון הערך שנקבל יקבע את רמת הלינאריות בין 2 תמונות כאשר קרוב ל-1 מראה על קשר לינארי וקרוב ל-0 מראה על חוסר קשר.

1.6. הסבר על אנדרויד (Android) :

מערכת אנדרויד היא מערכת הפעלה המופצת על ידי חברת Google בשיתוף פעולה עם Open Handset Alliance. מערכת הפעלה זו מעוצבת במיוחד לשימוש בטלפונים חכמים מבוססי מסך מגע וטאבלטים. לקראת סוף שנת 2010, מערכת הפעלה אנדרואיד הפכה למערכת ההפעלה הנפוצה ביותר בעולם לטלפונים חכמים. בשנת 2012 היא שלטה ב- 70% מכלל שוק הטלפונים החכמים העולמי.

1.7. הסבר על שרת-לקוח (Client-Server) :

מודל שרת-לקוח הוא אחד מתצורות התקשורת הנפוצות שיש ברשתות מחשבים. עקרון הפעולה זה שהוא מחלק את המשאבים (השרת), לבין מבקש השירות (הלקוח). הלקוח הוא תוכנה

אקטיבית שבה המשתמש צריך לפנות לשרת. השרת הוא תוכנה פסיבית, המאזינה לרשת ומחכה לפעול בהתאם לבקשות.

1.8. הסבר על מחשוב בענן (Cloud computing):

מחשוב בענן הוא שירותי מחשוב הניתנים למשתמש באמצעות מחשב מרוחק אליו מתחבר המשתמש דרך רשת האינטרנט. היתרון של השיטה זה שאנו משתמשים במשאבים של מחשב מרוחק במקום במשאבים של המחשב או טלפון חכם שלנו.

1.9. הסבר על מחלקה (Class):

מחלקה היא אחד המאפיינים של תכנות מונחה-עצמים. מחלקה מתארת אובייקט על מאפייניו ופעולותיו שכוללות משתנים, הגדרות ופונקציות המאוגדים למבנה יחיד ופועלים יחדיו. שימוש של מחלקה מופיע בכתיבת האפליקציית אנדרויד.

2. סקר ספרות

טכניקות רבות בעיבוד תמונה פותחו ב-1960 במעבדת Jet Propulsion, MIT, מעבדת Bell, אוניברסיטת מרילנד ועוד, כאשר השימושים העקריים היו צילומי לוויין, צילומים רפואיים, טלפון וידאו, זיהוי אובייקטים ושיפור תמונות. באותה תקופה העלויות והטכנולוגיה לא אפשרו את פתרונם של רוב הבעיות. ב-1970 הטכנולוגיה של המחשבים אפשרה את התחלת תקופת עיבוד התמונה. כיום התחום מאוד רחב ומאפשר אפליקציות רבות בעיבוד תמונה.

עיבוד תמונה דיגטלי מאפשר שימוש באלגוריתמים מרוכבים וכך מאפשר שימוש בשיטות שיהיו בלתי אפשריות בעולם האנלוגי. בפרט הטכנולוגיות הבאות יכולות להיפתר רק באמצעות עיבוד תמונה דיגטלי: סיווג, הבלטת תכונות, זיהוי תבניות, הקרנה וניתוח אותות בקנה מידה. חלק מהטכניקות שאיתם מבצעים עיבוד תמונה דיגטלי: פיקסליזציה, סינון לינארי, ניתוח גורמים ראשיים, ניתוח גורמים עצמאים, מודל מרקוב חבוי, משוואות דיפרנציאליות חלקיות, רשת קוהונן, רשת עצבית מלאכותית, Waveletes.

המחקר הראשון שעסק בנושא היה בשנת 1995 ע"י J.P.Lewis. באמצעות המחקר שלו פותח האלגוריתם למקדם מתאם מנורמל לזיהוי אובייקטים בתוך תמונה בתוכנה מטלב. במאמר Lewis מראה איך שימוש במקדם מתאם מנורמל פותר כמה בעיות ואיך ניתן בעזרת אלגוריתם פשוט לבצע זיהוי בתוך תמונה.

מחקר נוסף שעליו הסתמכתי היה של סארביה, פאטניק ובומביוולה (Sarvaiya, Patnaik & Bombaywala 2009) שעסק בזיהוי תבניות בתוך תמונה כללית ע"י שימוש במקדם מתאם. מאמר זה נותן גישה יותר אלגוריתמית לפתרון הבעיה.

מחקר נוסף שעליו הסתמכתי היה של ראו, פראתפאניל ונאגאבושאנס (Rao, Prathapanil & Nagabhooshanam 2014) שעסק באפליקציות של מקדם מתאם מנורמל על זיהוי אובייקטים בתמונה. מאמר זה נתן הסבר לפתרון הבעיה ודוגמא לזיהוי תבנית בתוך תמונה.

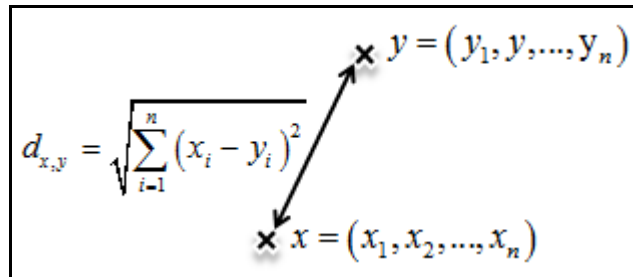
בעבודה זו אציג את המתמטיקה מאחורי המודל של מקדם מתאם מנורמל ואיך המודל מציג כמה פתרונות מעשיים בעולם עיבוד תמונה. מחקרים רבים פורסמו בנושא וכולם מראים שיטות רבות לשימוש במקדם מתאם מנורמל. במהלך העבודה אציג כמה שימושים בולטים בעולם העיבוד תמונה ואיך המודל בא לידי ביטוי.

3. הצגת משוואת מקדם מתאם

השימוש במקדם מתאם להתאמת תבניות מונעת על ידי מדידת מרחקים בשימוש של משוואת מרחק אוקלידית בריבוע.

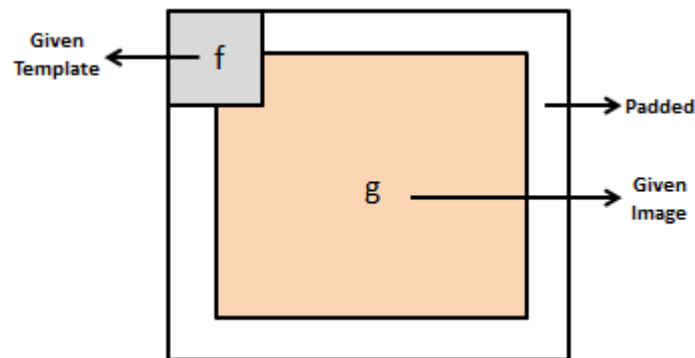
אם האותות הם דסקרטיים ונרצה לדעת מרחק בין הנקודה $x = (x_1, x_2, \dots, x_n)$ והנקודה

$$y = (y_1, y_2, \dots, y_n) \text{ נקבל את המשוואה } d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \text{ כמתואר באיור 4.}$$



איור 4 - מציאת מרחק על ידי משוואת מרחק אוקלידית בריבוע

באם נתון תמונה מקורית g ואנו רוצים להתאים את התבנית f כמו שמתואר באיור 5.



איור 5 - התאמת תבנית לתמונה

$$d_{f,g}^2 = \sum_{x,y} [f(x,y) - g(x-u, y-v)]^2 \text{ נשתמש במשוואת מרחק אוקלידית בריבוע}$$

f , התבנית שנרצה לחפש והסכום הוא על x, y תחת החלון g , שמכוון במיקום u, v .

$$d_{f,g}^2 = \sum_{x,y} [f^2(x,y) - 2 \cdot f(x,y) \cdot g(x-u, y-v) + g^2(x-u, y-v)]$$

$$\sum f^2(x,y) \approx \text{Const} \text{ והביטוי } \sum g^2(x-u, y-v) = \text{Const}$$

$$c(u,v) = \sum_{x,y} f(x,y) g(x-u, y-v) \text{ מה שמשאיר אותנו עם הנוסחא של מקדם מתאם}$$

ניתן גם לפתור לאותות רציפים למרות שמקרה של עיבוד תמונה זה לא בא לידי ביטוי.

$$\varphi_{x,y}(t) = \int_{-\infty}^{\infty} x(\tau-t) y(\tau) d\tau \text{ מוגדר:}$$

$$\varphi_{x,y}(t) = \int_{-\infty}^{\infty} x(\tau-t) y(\tau) d\tau \text{ אם נשווה זאת אם קובולוציה}$$

ההבדל היחיד שמקדם מתאם שאחד מהפונקציות לא הפיכה לכן: $\varphi_{x,y}(t) = x(-t) * y(t)$.
 אם נבצע פעולה זו במישור התדר פעולה זו מוגדרת: $\phi_{x,y} = FT[\varphi_{x,y}(t)] = X(f)Y(f)$.
 מקדם מקדם אינו פעולה קוממוטטיבית אך קונבולוציה כן לכן ניתן לרשום $\varphi_{x,y}(t) = \varphi_{x,y}(-t)$.

3.1. חסרונות בשימוש במקדם מתאם דיסקרטי

ישנם חסרונות לשימוש במקדם מתאם $c(u, v)$:

- אם האנרגיה של התבנית $\sum f^2(x, y)$ משתנה שימוש במקדם מתאם יכול להיכשל (דוגמא: הערך של תמונה מקורית יהיה קטן מערך של תמונה עם אור בזהב).
- ערך הביטוי $c(u, v)$ משתנה לפי ערך התבנית.
- מקדם מתאם אינו נותן מענה לשינויים של אור בתמונה.

3.2. הצגת מקדם מתאם מנורמל

מקדם מתאם מנורמל מתגבר על חסרונות אלה על ידי נורמליזציה של הוקטור תבנית ותמונה לגודל יחידה. נשתמש במקדם מתאם מנורמל של פירסון:

$$\gamma(u, v) = \frac{\sum_{x,y} [(f(x, y) - \overline{f_{u,v}})] [g(x - u, y - v) - \overline{g}]}{\sqrt{\sum_{x,y} [f(x, y) - \overline{f_{u,v}}]^2 \cdot \sum_{x,y} [g(x - u, y - v) - \overline{g}]^2}}$$

- $f(x, y)$ - התבנית, $\overline{f_{u,v}}$ - הממוצע של התבנית.
 - $g(x - u, y - v)$ - התמונה, $\overline{g}$ - הממוצע של התמונה בגודל התבנית.
- משוואה זו נקראת מקדם מתאם מנורמל.

3.3. מימוש אלגוריתם מקדם מתאם מנורמל

נסביר פה את עקרון הפעולה של האלגוריתם מקדם מתאם מנורמל במטלב כמתואר באיור 5.

- חישוב גודל התבנית $f(x, y)$.
- ריפוד התמונה $g(x, y)$ באפסים, נוסף חצי מאורך התבנית בחלק העליון והתחתון של התמונה ונוסיף חצי מרוחב התבנית בחלק הימני והשמאלי של התמונה.
- נזיז את התבנית על גבי כל התמונה ותוך כדי חישוב הסכום של הערכים בתמונה.
- בנוסף נחשב הערך בריבוע, הסכום והשורש של התבנית והתמונה כמתואר בנוסחא.
- נחלק את הביטוי שקיבלנו מונה חלקי מכנה.
- נמצא את הערך המקסימלי של הביטוי ושם נמצא ההתאמה הקרובה ביותר.
- נשחזר את התבנית בתוך התמונה המקורית ונראה אם יש התאמה.

4. שימושים במקדם מתאם מנורמל

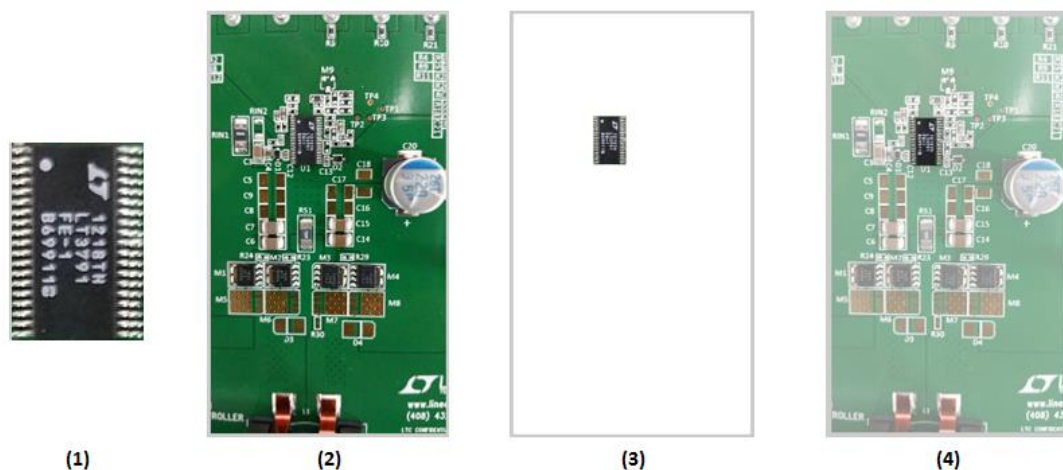
לאחר הצגת המודל מקדם מתאם מנורמל, בעולם עיבוד התמונה התפתחו לאורך השנים כמה שימושים מרחקי לכת.

השימוש הנפוץ במקדם מתאם מנורמל היה כמדד להערכת מידת הדמיון (או השוני) בין שתי תמונות. היתרון בשימוש במקדם מתאם מנורמל כנגד מקדם מתאם הוצג מקודם, החסרון היחיד הוא שלא ניתן לחשב את המודל בשיטת ה-Fast Fourier Transform במישור התדר דבר המקשה על יעילות החישוב. למרות זאת היתרונות והיעילות עולים על החסרונות לכן עדיף להשתמש במודל מקדם מתאם מנורמל.

אציג כאן כמה מהשימושים הבולטים ואת הפרויקט בהנדסת חשמל שמשתמש במודל.

4.1. מציאת אובייקט

אחד מהשימושים הבולטים הוא מציאת אובייקט בתוך תמונה. מציאת אובייקט למחשב זו מטלה קשה אם משווים לבני אדם שמבצעים פעולה זו ללא מאמץ ובאופן מיידי. אלגוריתם של מחשב צריך להתייחס לצורה, זווית, הצלה, יחס מידה ועוד לכן מטלה זו אינה פשוטה. שימוש במקדם מתאם מנורמל פותר לנו חלק מהבעיות הללו אך לא ב-100%. נצטרך להגדיר תבנית שנרצה לחפש בתוך תמונה גדולה. שימוש עיקרי במציאת אובייקט נשתמש בעורכי מעגלים חשמליים לודא שהרכיבים קיימים. מעגלים חשמליים עוברים הלחמה והרכבה אוטומטיים, צריך לודא בסיום העבודה שכל הרכיבים קיימים ובעייה נוספת הגדלים הקטנים של הרכיבים מקשה על בדיקה ויזואלית לכן ניתן לתכנת אלגוריתם שיאתר את הרכיבים בשיטה אוטומטית כמוצג באיור 6.



איור 6 - מציאת אובייקט בתוך תמונה

(1) Chip LT3791 (2) מעגל מודפס (3) מציאת LT3791 בתמונה (4) הצגת המיקום בלוח

4.2. עקיבה אחר אובייקט

היום הולך וגדל השימושים של עקיבה לאחר אובייקטים בתחום הצבאי והאזרחי. הדבר הראשון שצריך לעשות הוא לבצע רכישה של המטרה בתמונה ואז להגדיר כל כמה זמן לחשב את מיקום האובייקט. גם כאן נעשה שימוש במקדם מתאם מנורמל על מנת לעקוב לאחר אובייקט ולסמן את המסלול שלו כמתואר באיור 7.



(3)



(4)



(1)



(2)

איור 7 - עקיבה אחר אובייקט
(1) רכישת מטרה (2) עקיבה 1 (3) עקיבה 2 (4) עקיבה 3

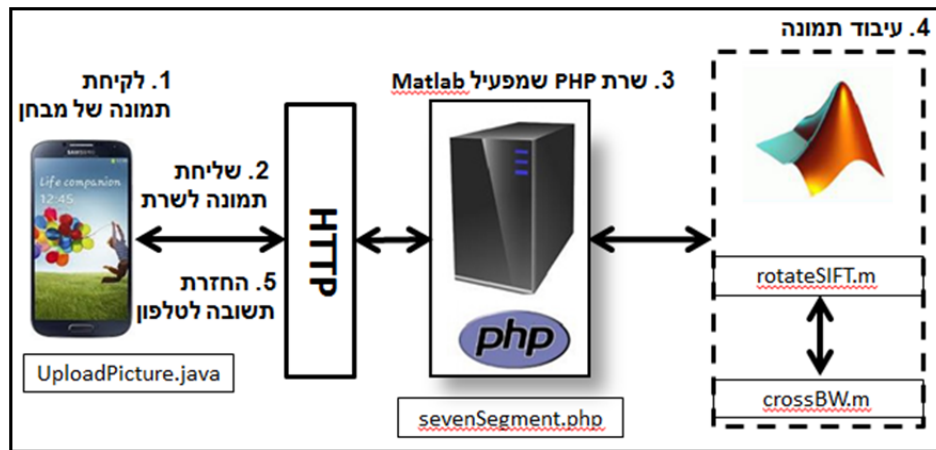
4.3. זיהוי ספרות

4.3.1. מפרט פונקציונאלי

המערכת מצלמת טופס מבחן בעזרת טלפון חכם שרץ עליו מערכת הפעלה אנדרויד 2.2 ומעלה, בעלת מצלמה עם רזולוציה של מינימום 0.8 מגה פיקסל (1024x768). בעזרת אפליקצית אנדרויד נכניס מספר תעודת זהות ונצלם/נבחר תמונה של המבחן. התמונה תעבור דחיסה לתמונה בגודל 1024x768 לכל תמונה שנבחרת שגדולה מהגודל הנ"ל. התמונה תשלח דרך אינטרנט סלולרי או אינטרנט מקומי אל השרת בכתובת קבועה. השרת יבצע עיבוד תמונה על מנת להמיר את התמונה לתשובה מספרית ולהשוות אותה עם תשובה שתשמר בשרת. השרת יחזיר מחרוזת המכילה את התשובה דרך האינטרנט אל האפליקציה אשר תציג את התוצאה.

4.3.2. תרשים מלבנים כולל של המערכת

המערכת בנויה מ-3 חלקים עיקריים תוך שימוש ברשת האינטרנט להעברת מידע. באיור 8 ניתן לראות את תרשים המלבנים של המערכת וסדר שלבי המעבר בין המכלולים השונים.

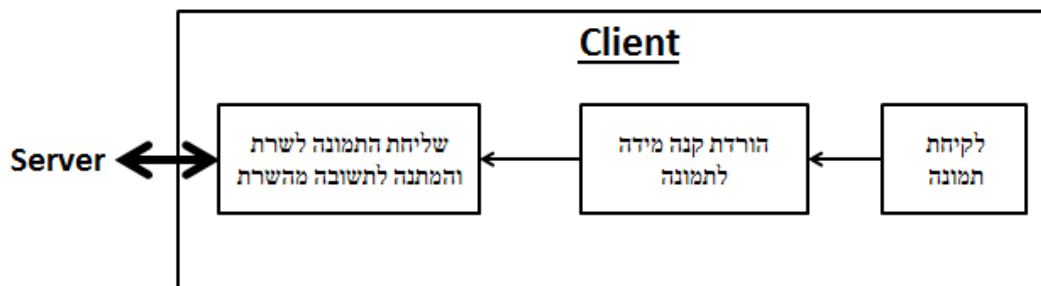


איור 8 - תרשים זרימה של מכלולי הפרויקט

4.3.3. אפליקציית אנדרויד.

4.3.3.1. שלב תכנון בניית האפליקציה

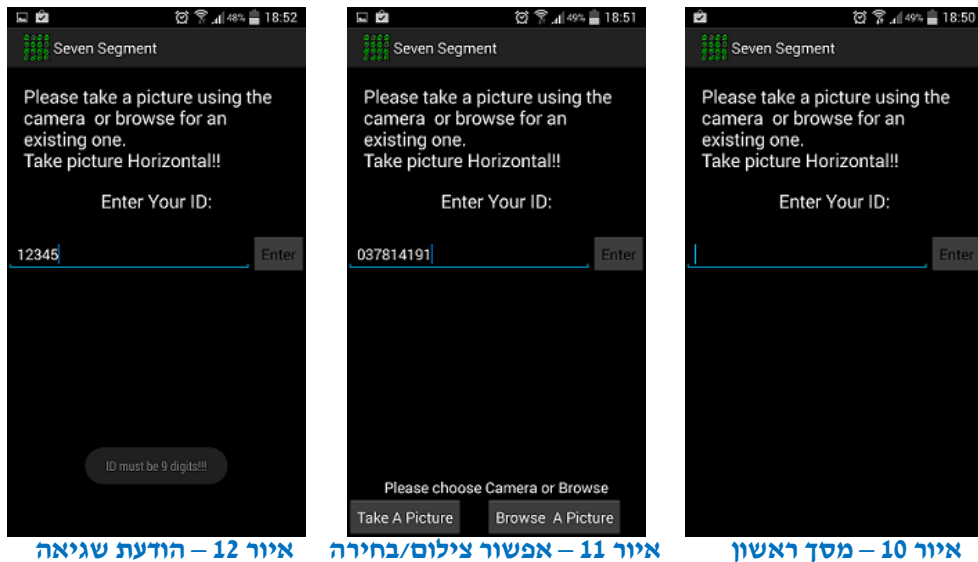
בפרויקט נעשה שימוש באפליקצית אנדרויד כדי לשלוח את התמונה של המבחן ולקבל תשובה. הדרך הטובה ביותר לבצע זאת כדי להתגבר על הבדלים בסוגי הטלפונים החכמים היא לשלוח את התמונה לשרת ששם יתבצע העיבוד תמונה ללא קשר למאפייני הטלפון. את האפליקציה ניתן לחלק ל-3 מכלולים עיקריים כמתואר באיור 9:



איור 9 - תרשים זרימה של האפליקציה

4.3.3.2. לקיחת תמונה

זה המסך הראשון של האפליקציה שנקרא גם פעילות. יש הוראות שימוש והוא מחכה לקלט מהמשתמש של תעודת זהות כמתואר באיור 10, רק אם הזין 9 ספרות תיפתח לו האפשרות ללקיחת תמונה או בחירת תמונה מהזיכרון של הטלפון החכם כמתואר באיור 11. לאחר בחירת תמונה, תפעיל את המסך השני. במקרה ולא הזין 9 ספרות המשתמש יקבל הודעת שגיאה כמתואר באיור 12.



איור 12 – הודעת שגיאה

איור 11 – אפשרות צילום/בחירה

איור 10 – מסך ראשון

4.3.3.3. הורדת קנה מידה לתמונה

זהו המסך השני הוא מציג את התמונה של המבחן ושדה של תעודת הזהות שהוזנה כמתואר באיור 13. האלגוריתם בודק את גודל התמונה אם גדול מ-0.8 מגה פיקסל שזה 1024x768 הוא מוריד קנה מידה לרזולוציה הנ"ל. ישנה אפשרות לחזור למסך קודם זה יגרור מחיקת התמונה שצולמה. אם המשתמש מאשר את התמונה הוא יעבור למסך הבא.



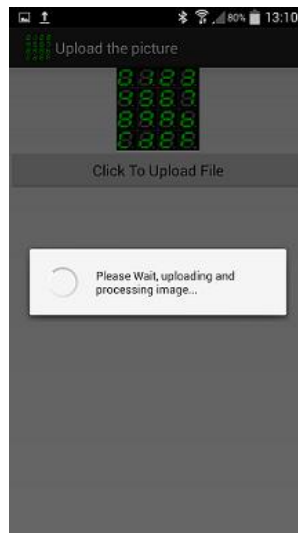
איור 13 – מסך שני

4.3.3.4. העלאת תמונה לשרת והמתנה לתשובה

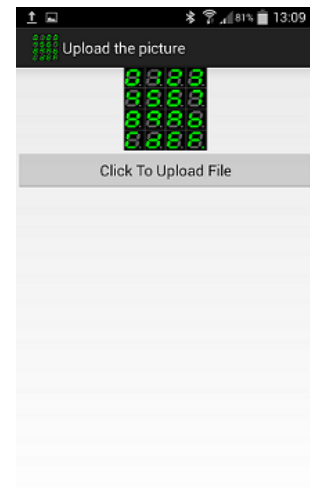
זהו המסך השלישי כמתואר באיור 14 הוא אחראי על העלאת התמונה לשרת בעל כתובת קבועה. כאשר המשתמש לוחץ על כפתור העלאה נפתח קישוריות לאינטרנט דרך רשת סלולרית / רשת אלחוטית כמתואר באיור 15 ומחכה לתשובה חזרה מהשרת. השרת מחזיר תשובה של Success/Failed. אם השרת החזיר Success הוא קורא את התשובה מהשרת ומציג אותה במסך ורושם אם הצליחה או נכשלה ביחס לקובץ תבנית כמתואר באיור 16.



איור 16 – תשובה מהשרת



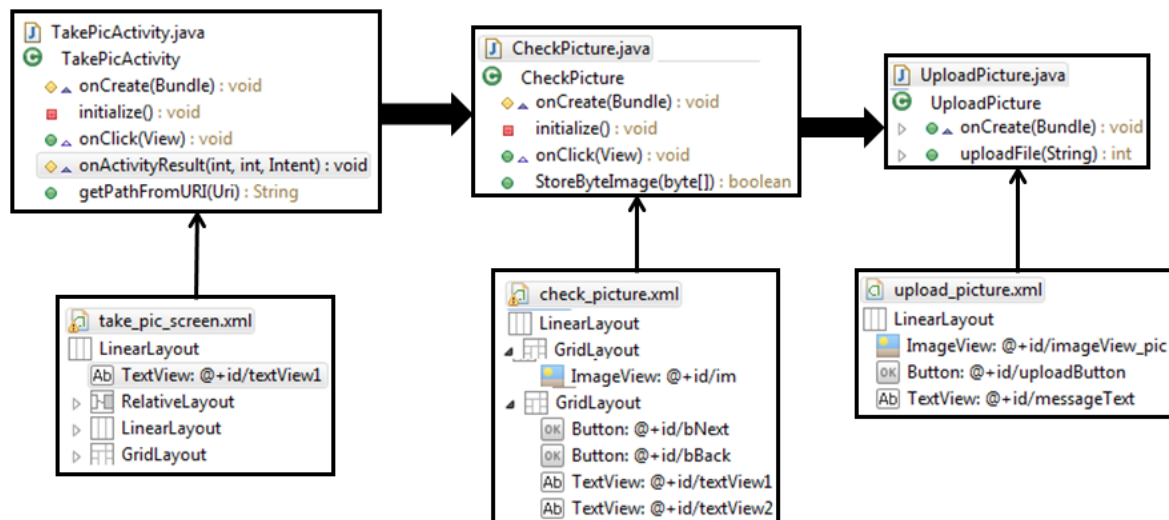
איור 15 – העלאת תמונה



איור 14 – מסך שלישי

4.3.3.5 Class Diagram - דיאגרמת מחלקה

דיאגרמת מחלקה כמתואר באיור 17 מציג את מבנה המערכת של האפליקציה ע"י הצגת מחלקותיה והקשרים ביניהם.

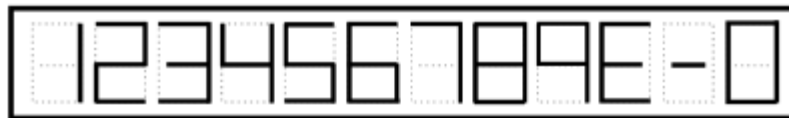


איור 17 – דיאגרמת מחלקות

4.3.4 טופס תשובות

4.3.4.1 תכנון טופס תשובות

כדי להתחיל את בניית האלגוריתם של עיבוד תמונה היה צורך לבנות דף תשובות עם מידע על כל פרט בדף ברמת הפיקסל. ישנם כמה הגדרות עיקריות שעליו התבסס עיצוב הדף, כל מבנה של ספרה יהיה באורך 70 פיקסל ורוחב 45 פיקסל. טופס התשובות כמתואר באיור 13 מורכב מהזנת 5 ספרות אחרונות של ת"ז ו 5 תשובות אפשריות כשכל תשובה כוללת 7 מספרים. ישנם בסה"כ 12 אפשרויות לכתיבה כמתואר באיור 18.

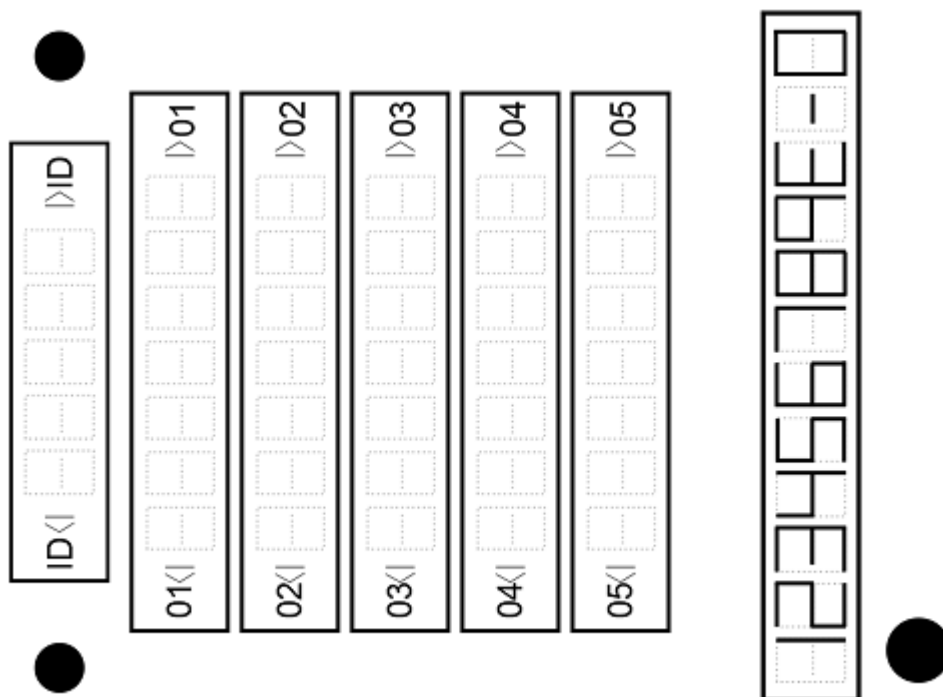


איור 18 – אפשרויות הכתיבה

4.3.4.2. הוראות צילום ומילוי טופס בחינה.

- יש למלא את הספרות כקווים ישרים ולא להשאיר רווחים בין קו לקו.
- יש לצלם את הדף לרוחב כאשר הקצה העליון בצד שמאל כמתואר באיור 19.
- יש לצלם כאשר המצלמה אופקית.
- יש לצלם את כל הדף אך לצלם קצת מעבר לקצה הדף ושהמצלמה מקבילה לדף.

4.3.4.3. טופס הבחינה

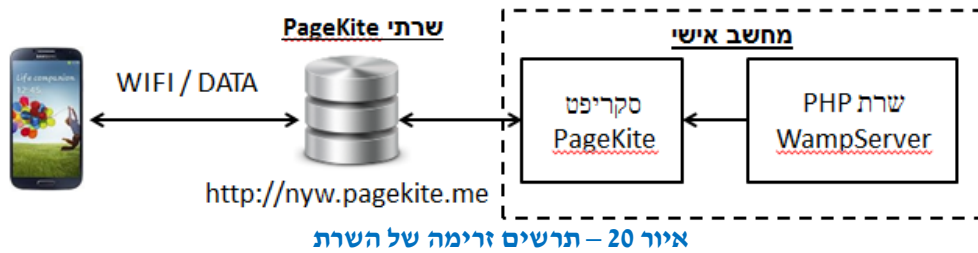


איור 19 – טופס הבחינה

4.3.5. שרת PHP

4.3.5.1. תכנון שרת PHP

בעבודה זו הותקן שרת PHP על המחשב האישי שמחובר לאינטרנט כמתואר באיור 20. היה צורך בתכנות סקריפט PHP שמקבל תמונה מהטלפון החכם ומפעיל סקריפט מטלב לעיבוד תמונה. בסיום ישלח את המידע חזרה לטלפון הנייד ויסגור את הקישור. בעקבות התקנת השרת על המחשב האישי השרת עדיין לא היה ציבורי כדי שהטלפון החכם יכול לתקשר איתו. לשם כך הותקן סקריפט פיטון של חברת PageKite אשר נותן לך כתובת קבועה <http://nyw.pagekite.me> וכך הופך את השרת לציבורי.



4.3.6. עיבוד תמונה

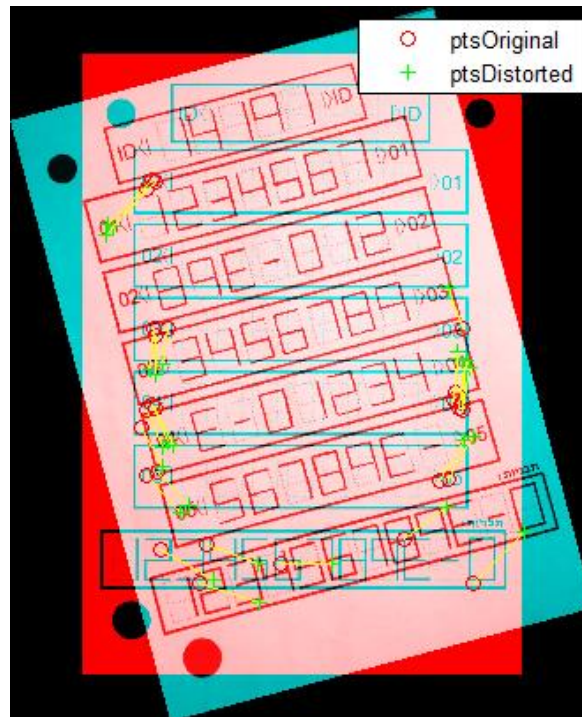
4.3.6.1. סיבוב ותיקון קנה מידה

כאשר השרת מקבל תמונה לא בהכרח שהיא בסיבוב הנכון או בקנה מידה הנכון. לכן בעזרת סקריפט מטלב אנו מקבלים את התמונה מהטלפון הנייד ומשווה אותו לטופס המבחן המקורי. נעשיית השוואה בין נקודות דומות בתמונות ואז נעשה חישוב של כמה סיבוב וכמה תיקון קנה מידה צריך לעשות. בסיום נשמרת התמונה אחרי התיקון.

נעשית עבודה עם אלגוריתם SIFT – Scale Invariant Feature Transform כדי לתקן את התמונה. נקודות SIFT מחולצות מתמונה, אובייקט מאותר בתמונה ע"י השוואה עם התמונה המקורית ע"י משוואת מרחק אוקלידית כמתוארת להלן:

$$\text{if } p=(p_1, p_2), q=(q_1, q_2) \rightarrow d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2}$$

לאחר קבלת הבדלים במרחקים בין 2 התמונות כמתואר באיור 21 נקבל זוויות לתיקון הסיבוב ויחס לתיקון קנה מידה של התמונה.



איור 21 – אלגוריתם SIFT לתיקון סיבוב ויחס קנה מידה

$$\theta = \arctan \frac{y}{x}, \quad r = \sqrt{x^2 + y^2} \quad \text{: המשוואות לזווית והזזה}$$

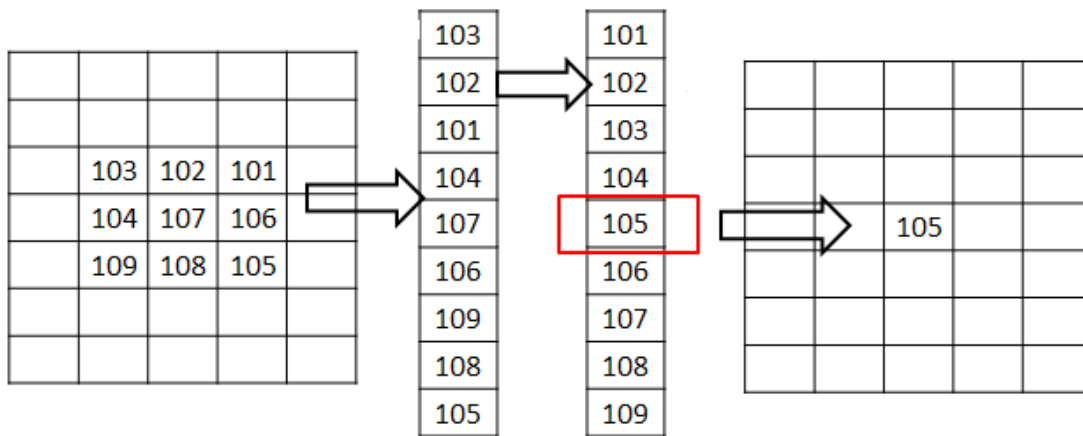
4.3.6.2. חיתוך וזיהוי

לאחר תיקון סיבוב וקנה מידה, נחתכים כל התשובות, ונעשיית השוואה מול מסד נתונים. נציג את האלגוריתם מציאת ספרה המתואר באיור 23.

- המספרים בטופס התשובות במקומות קובעים לכן נחתוך ספרה ספרה.
- הפיכת תמונה לשחור לבן זה השלב הראשון כדי להתאים לתמונת המקור. יש לנו תמונה שמיוצת במטריצה דו-מימדית 1024×768 בכל פיקסל יהיה ערך מ-0-255.
- כל ערך מנומל ע"י 255 כך שנקבל בסוף מטריצה דו מימדית ובכל פיקסל יהיה ערך 0-1.

$$X_{i,j} = \frac{X_{i,j}}{255} : \text{כל פיקסל עובר את המשוואה הבאה}$$

- העברת הספרה דרך Median Filter כדי לטפל ברעשי "Salt and pepper" כמתואר באיור 22.



איור 22 – דרך פעולה של Median פילטר 3X3

משוואה של Median פילטר 3X3 ניתן להרחיב את הפילטר בהתאם לגודל המטריצה אך לשלם בזמן הרצה.

$$x_{i,j} = \text{median} \{ x_{i-1,j-1}, x_{i,j-1}, x_{i+1,j-1}, x_{i-1,j}, x_{i,j}, x_{i+1,j}, x_{i-1,j+1}, x_{i,j+1}, x_{i+1,j+1} \}$$

- בסוף נעשית השוואה של הספרה מול מסד הנתונים ישנם 12 תבניות לשם השוואה. כל השוואה בשיטת Normalized 2-D Cross-Correlation תחזיר תשובה בין 0-1 וניקח את התשובה המקסימלית בתור התשובה שלנו.



איור 23 – אלגוריתם מציאת ספרה

נבדקו מספר רב של טפסים לשם בדיקת תקינות של Normalized 2-D Cross-Correlation נציג בטבלה 1 מדגם של כמה מהניסויים.

ניסוי מספר	צבע	גודל	מספר תשובות	מספר הצלחות	אחוז הצלחה
1	שחור	0.5 Pixels	40	36	90.0%
2	שחור	0.5 Pixels	40	36	90.0%
3	כחול	0.5 Pixels	40	37	92.5%
4	כחול	0.5 Pixels	40	38	95.0%
5	שחור	0.7 Pixels	40	38	95.0%
6	שחור	0.7 Pixels	40	38	95.0%
7	כחול	0.7 Pixels	40	35	87.5%
8	כחול	0.7 Pixels	40	34	85.0%

טבלה 1- תוצאות זיהוי ספרות

במהלך בדיקת הטפסים נחקר מה ההסתברות לזיהוי ומהו הספרה השגויה שנקבל במקרה וטעות זיהוי. נוכל לראות מהתוצאת שמרוכזות טבלה 2 שרוב השגיאות נובעות מזיהוי של ספרות 8 ו-6 בתור ספרה 5, 6 ו-8.

מספר	זיהוי 1	זיהוי 2	הסתברות זיהוי	מספר	זיהוי 1	זיהוי 2	הסתברות זיהוי
1	7		גבוהה	7	3	1	גבוהה
2	8		גבוהה	8	5	6	נמוכה
3	8		גבוהה	9	4		בינונית
4	9		גבוהה	E	6	8	גבוהה
5	6	8	בינונית	-			גבוהה
6	5	8	נמוכה	0	8		גבוהה

טבלה 2- טבלת זיהוי הסתברות והטעויות הנפוצות לכל ספרה

5. סיכום ודיון

5.1. סיכום

בספר זה מפורטים מוצג מודל מקדם מתאם מנורמל ושימושיו בעולם העיבוד תמונה. במהלך הספר הצגתי את הפירוט המתמטי למודל וכן את ההבדלים שנעשו מהמודל הקודם מקדם מתאם ומהם חסרונותיו ביחס למקדם מתאם מנורמל. במהלך העבודה הצגתי 3 שימושים בולטים למודל ואיך הם באים לידי ביטוי בעולם העיבוד תמונה. 2 השימושים הראשונים הינם שימושים בולטים בתחום העיבוד תמונה, אחד השימושים הוא זיהוי אובייקטים בתוך תמונה שימוש זה נרחב בעולם העיבוד שבבי במכונות SMT. שימוש שני נפוץ הוא רכישה של מטרה ועקיבה לאחר המטרה בכ ל השלבים, שימוש זה נרחב בעולם הצבאי. השימוש השלישי הוא הפרויקט המחקרי שביצעתי במסגרת תואר בהנדסת חשמל, כאשר השימוש שלו הוא זיהוי ספרות ועיבוד ומתן ציון לטופס מבחן. שלבי העבודה הנדרשים להפעלת אפליקציית אנדרויד שתצלם / תבחר טופס תשובות ותשלח לשרת חיצוני את התמונה לעיבוד תמונה. בשרת יבוצע סיבוב ותיקון קנה מידה לתמונה בעזרת אלגוריתם SIFT. בעזרת Normalized 2-D Cross-Correlation ניסינו למצוא את הספרות שבטופס הבחינה. נמצא כי השימוש Normalized 2-D Cross-Correlation נותן לנו הצלחה גבוהה של 85%-100%. המערכת מצליחה לאתר בממוצע מעל 37 מתוך 40 מספרים בטופס הבחינה.

5.2. מסקנות

הפרויקט מציג יכולות מאוד גבוהות ומדויקות בזיהוי אובייקטים ומהם השימושים הנפוצים כיום בשוק. במהלך העבודה נחשפתי לעולם העיבוד התמונה ואיך בעזרת מודלים מתמטיים וסטטיסטיים ניתן להגיע לתוצאות גבוהות ויעילות בשונה מהיכולת של בן אדם. הפרויקט בהנדסת חשמל מראה כי ישנה אפשרות לזהות את התשובות הנכונות בטפסי המבחן ובעזרת פיתוח נוסף ניתן להגיע לפרויקט שיציג תשובה של טופס התשובות באופן מלא. מעבר לכך הפרויקט מראה שיש אפשרות של פיתוח טכנולוגי שניתן להשתמש בו ללימודי הנדסת חשמל ואלקטרוניקה לממש יכולות טלפון חכם - שרת - מטלב על מנת לבצע פרויקטים מתקדמים עתידיים.

- [1] J. P. Lewis, "Fast Template Matching", *Vision Interface*, p. 120-123, 1995.
- [2] Y.R.Rao, N.Prathapani and E.Nagabhooshanam, "Application of Normalized Cross Correlation to Image Registration. IJRET, 2014
- [3] J.N.Sarvaiya, Dr.S.Pathnail, S.Bombaywala, "Image Registration By Template Matching Using Normalized Cross-Correlation, International Conference on Advances in Computing, Control, and Telecommunication Technologies, 2009
- [4] EE368/CS232 Digital Image Processing <http://web.stanford.edu/class/ee368/index.html>
- [5] Android Developers <http://developer.android.com/index.html>
- [6] Android Point <http://sunil-android.blogspot.co.il/2013/09/file-upload-on-php-server-in-android.html>
- [7] MathWorks <http://www.mathworks.com/>
- [8] Eclipse <https://www.eclipse.org/>
- [9] Android (operating system) [http://en.wikipedia.org/wiki/Android_\(operating_system\)](http://en.wikipedia.org/wiki/Android_(operating_system))
- [10] WolframMathWorld <http://mathworld.wolfram.com/Cross-CorrelationTheorem.html>
- [11] Scale Invariant Feature Transform <http://www2.it.lut.fi/kurssit/07-08/CT20A6100/seminars/2009-2010/SIFT.pdf>
- [12] Wikipedia Cross Correlation <http://en.wikipedia.org/wiki/Cross-correlation>
- [13] Wikipedia Pearson correlation coefficient http://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient
- [14] Image Processing Wikipedia <http://www.disi.unige.it/person/RovettaS/rad/image-processing-wikipedia-book.pdf>
- [15] Bahram Javidi Image Recognition and Classification, Algorithms, Systems, and Application, 2001