

מציגים – שון אסולין, אורי טביבי  
מנחות – ד"ר אירינה רבייב, ד"ר מרינה ליטבק

## תקציר

זיהוי מגדר לפי כתב יד נחשבת לאחת הבעיות הנחקרות ביותר כיום ומעוררת עניין רב בשל מגוון התחומים הרחב שנעשה בהם שימוש, כמו רפואה, פסיכולוגיה, ניתוח מסמכים היסטוריים וחקירות קרימינולוגיות. פרויקט זה עוסק בזיהוי המגדר של כותב באופן אוטומטי באמצעות רשתות נוירונים לפי תמונה סרוקה של כתב יד. המחקר מתבסס על מאגר כתבי יד HebG בשפה העברית, וכן על מאגר QUWI המכיל כתבי יד בשפות האנגלית והערבית.

## טכנולוגיות



Python



Keras



TensorFlow



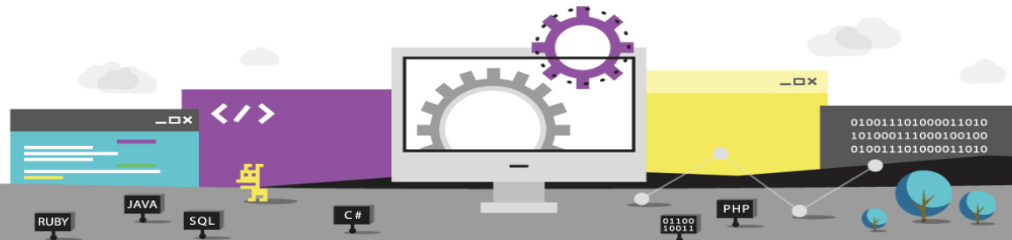
JavaScript



Flask

## מטרות ויעדים

- מחקר ואימון מודלים המותאמים לכתבי יד בשפה העברית.
- יצירת מערכת אמינה בעלת אחוזי דיוק גבוהים.
- תהליך עיבוד זיהוי מהיר ויעיל.
- פיתוח מערכת (אתר) לביצוע זיהוי המגדר באופן אוטומטי.





מציגים – שון אסולין, אורי טביבי  
מנחות – ד"ר אירינה רביב, ד"ר מרינה ליטבק

## שיטת העבודה

## חלוקת התמונה ל - Patches

## טופס מקור



## עיבוד וחילוץ אזור הכתב

כחל - רכב על פולמוס נגד שטן פולחן, אף מר קדרי פולחן  
באור אשט אולחן? חולחן מחדש, פולחן החולחן פולחן  
פולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן  
פולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן  
פולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן  
פולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן

לחולחן נגד שטן פולחן  
חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן  
חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן  
חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן  
חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן, חולחן

## הזנת נתונים למודל

CNN

## זיהוי לפי Patch

Female

Female

Male

זיהוי סופי

Female

## תוצאות ומסקנות

- לאחר ניתוח תוצאות הסקר בקרב הקהל הרחב, הדיוק הממוצע עומד על 67.4%.
- באמצעות שימוש ברשתות נוירונים הגענו לאחוז דיוק העומד על 85%.
- ניתן לראות שיפור משמעותי באחוזי הדיוק באמצעות שימוש ברשתות נוירונים, דבר המשאיר מקום להמשך מחקר ושימוש בכלים אוטומטים נוספים.

## ניסויים

- במסגרת המחקר ביצענו ניסויים על מאגר כתבי יד בשפה העברית - HebG.
- על מנת להרחיב את המחקר ולהעריך את טיב המודלים, ביצענו ניסויים נוספים על מאגר QUWI המכיל כתבי יד בשפות האנגלית והערבית.
- ערכנו סקר שהופץ לקהל הרחב על מנת להשוות בין זיהוי זיהוי אנושי לבין רשתות נוירונים.

