מערכת מבוססת CNN לסיווג תמונות פרחים לפי סוג הפרח

* 1. תקציר

המערכת שלנו מציעה שירות של סיווג תמונות על פי האובייקטים המופיעים בהן. בפרט, המערכת שפיתחנו מקבלת אוסף תמונות של פרחים, ומסווגת אותם לפי 5 סוגים שהוגדרו מראש – Daisy, Rose, Dandelion, Sunflower, Tulip. בעיה זו היא שונה מסיווג כללי של תמונות, שכן סיווג תמונות לכאלו שמכילות בני אדם, וכאלו שמכילות מטוסים היא פשוטה יותר, לאור האופי השוני מאוד של הסיווגים. בבעיה שלנו התמונות עשויות להיות דומות מאוד, ועל המערכת להפריד ולסווג בין תמונות אלה. השיטה אותה בחרנו כדי לבצע את משימת הסיווג היא התבססות על מודל של למידת מכונה, המבוסס בעיקר על מודל CNN שאומן כבר על מאגר הנתונים של תחרות ImageNet. לצורך השוואה ומדידת טיבו של המודל המאומן נעזרנו בפרמטרי דיוק כמו precision ו-recall, והשווינו את תוצאות המודל מול מודל מתחרה של צוות מקביל, אשר אומן על מאגר תמונות זהה לשלנו.

* 1. מבוא

תחום סיווג התמונה הוא תחום שבשנים האחרונות נחקר מאוד, ומוצרים חדשים שיוצאים לשוק עושים שימוש נרחב בפתרונות שונים. לדוגמה, מערכות של כלי רכב אוטונומיים דורשים כיום התממשקות עם התשתיות והמאפיינים הקיימים היום בכבישים – הולכי רגל, תמרורים, מעברי חצייה, רכבים אחרים, רמזורים ועוד. כדי התמודד באופן בטיחותי בעולם, על מערכת אוטונומית שכזאת להצליח להבין במדויק ובמהירות את מצב העולם. להצליח להבין האם הרכב ממול נוסע בכיוון הנסיעה או דווקא מגיע ממול, האם ליד מעבר החצייה עומד אדם או דווקא עץ, האם התמרור מגביל את המהירות ל-100 קמ"ש או ל-50 ועוד.

בעיית סיווג הפרחים היא הבעיה שבחרנו לפתור. בניגוד לניסיון לסווג תמונות על פי תוכנן לסיווגים שונים (לדוגמה, אופניים, אדם, משאית, וכו'), סיווג תמונות בתוך אותה קטגוריה, במקרה זה פרחים, מציג אתגרים גדולים יותר לאור העובדה שמחד גיסה ישנו דמיון גדול יותר בין הסיווגים השונים , ומאידך גיסא עשוי להיות שוני רב בתוך אותו סיווג [1] [2]. העובדה האחרונה נכונה במיוחד בעולם הפרחים, שכן פרחים גמישים מטבעם, בניגוד למשל למכונית, עובדה שמאפשרת לאותו האובייקט ממש, לתפוס צורות שונות במרחב, ועל כן להיראות אחרת בתמונות שונות [3].

במערכת זו החלטנו להתבסס על מודל CNN שאומן בעבר. התבססות זאת אפשרה לנו לחסוך זמן פיתוח רב על ידי ניצול ארכיטקטורה קיימת שהציגה ביצועים טובים במערכת אחרות, ובנוסף להשתמש במשקולות קיימות על גבי אותה הארכיטקטורה, אשר אפשרה מצד אחד חסכון של זמן אימון רב, ומצד שני התבססות על ידע רב שנצבר במהלך אימון על dataset גדול [4]. במקרה זה, בחרנו מודל שאומן על ImageNet [5], מאגר נתונים הגדול בשני סדרי גודל מהdataset שזמין לנו. בפרט, בחרנו במודל MobileNet [6]. המודל מציג פשרה טובה בין דיוק בסיווג התמונות, ובין דרישת משאבים נמוכה יחסית למודלים אחרים [6].

המערכת שפיתחנו הציגה ביצועים טובים יותר מאשר מערכת דומה שפותחה על ידי צוות מקביל. בדיקות השוואה בוצעו הן על ממשק המשתמש של המערכות, והן על איכות ודיוק בזיהוי התמונות השונות. לשם טיבו של הסיווג נעזרנו בפרמטרים שונים אשר חושבו בתום הצגת קבוצת תמונות זהה. תוצאות הניסויי שערכנו מאששות את ההשערה שלנו כי שימוש בTransfer Learning מסייע בהשגת ביצועים עדיפים, במיוחד כאשר נתון dataset קטן.

* 1. רקע ועבודות קודמות
     1. רקע

בבסיס המודל שלנו עומדת רשת קונבולוציה (להלן CNN). לעומת רשת נוירונים קלאסית, CNN מותאמת במיוחד לקלט של תמונה [7]. רשת נוירונים קלאסית מותאמת להתמודד, לדוגמה, בטבלה שמכילה אוסף פיצ'רים, ותכונת מטרה. אוסף הפיצ'רים לרוב יכיל מימדים בלתי תלויים, או שנוכל לבחור כאלו שהתלות בניהם נמוכה [8]. לעומת זאת, כדי לבצע Image Classification אנו נדרשים להתמודד עם תמונות, שבאופן אינהרנטי לכל פיקסל יש משמעות מסויימת רק בתוך ההקשר של שאר הפיקסלים מסביבו. מודל CNN מתחשב בדיוק בנקודה זו [7].

ניתן לחשוב על תמונה כמערך תלת מימדי בגודל WHN– מימד לרוחב התמונה בגודל W, אורך התמונה בגדול H, ומימד נוסף עבור הערוצים השונים – הם RGB, בגדול N. ב-CNN ישנם פילטרים בגודל KKM (כאשר K בדרך כלל שווה ל-3 או 5, ו-M בדרך כלל שווה ל-N, במקרה הנ"ל בשכבת הקלט M יהיה שווה ל-3), אשר מועברים על גבי המערך התלת מימדי, בדומה לשיטת הקונבולוציה מעולם עיבוד האותות. תוצאת המעבר הזה עבור כל פילטר נקראת Feature Map. לאחר מעבר כלל הפילטרים על התמונה המקורית, נקבל מערך תלת מימדי חדש, בגודל W’H’N’, אשר ייצגו את תוצאות הפילטרים השונים, וכאשר N’ שווה לכמות הפילטרים שהפעלנו על התמונה המקורית. כעת ניתן להמשיך לעבד את ה-Feature Map שהתקבל באופן רקורסיבי, בדומה לדרך שבה טיפלנו בקלט המקורי [9]. לאחר מספר שכבות כאלו, ניתן להוסיף שכבת Pooling – זהו פילטר תלת (או דו) מימדי נוסף, אשר מסייע במניעת overfitting של המודל, ובנוסף מצמצם את גודל המימדים על ידי ביצוע של מומצע בין ערכי הפיקסלים עליהם מכילים את הפילטר.

Transfer Learning היא טכניקה מוכרת לרתימת מודל אשר אומן במקור לביצוע משימה אחת, ואימונו מחדש לביצוע משימה שנייה [4]. בבסיס הטכניקה עומדת ההבנה שהשכבות הראשונות ברשתות נוירונים עמוקות קולטות או עוסקות בפיצ'רים פשוטים יחסית, או כלליים. לדוגמה, במערכת לזיהוי פנים, סביר שהשכבות הראשונות במודל יזהו פינות, קווים, ומאפיינים קטנים. שכבות עמוקות יותר יזהו אולי עין שלמה, ורק השכבה העליונה ביותר תדע להרכיב פנים שלמות. אם כן, אם המשימה החדשה שלנו דומה בהיבטים מסוימים למשימה המקורית, נוכל להיעזר באותן שכבות נמוכות מאומנות, וכך לא רק לחסוך זמן אימון יקר, אלא אולי אף לקבל שכבות מאומנות יותר מכפי שהיינו יכולים להשיג עם הdataset שזמין לנו, על ידי שימוש במשקולות של המודל, אשר הושגו על ידי אימון על dataset גדול [4].

* + 1. עבודות קודמות

ישנו מחקר רב בתחום עיבוד תמונה בכלל, והתמודדות עם סיווג תמונות בפרט. בבסיס המערכת שפיתחנו עומד מודל CNN, אשר מאז 2012 מציע את הביצועים הטובים ביותר בכל הנוגע לסיווג תמונה [10]. המודל שלנו מבוסס על מודל ה-MobileNet, המציע ארכיטקטורה המכילה מספר שכבות מצומצם לעומת מודלים אחרים, בכדי להתאים את השימוש במודל למערכות mobile ו-embedded [6]. בדומה למערכת שלנו, בוצעו מספר מחקרים [11][12] [13] אשר עוסקים במניעת overfitting, ואחת הטכניקות היעילות ביותר לצמצום הפער שקיים בין הdataset הקטן לבין הרעב של מודלי CNN למידע אשר מוביל ל-overfitting, היא Data Augmentation [11].

בנוסף למחקר שבוצע בתחום, ישנן מספר מערכות אשר פותרות בעיות דומות לזו שהמערכת שלנו מנסה לפתור. המערכת Imagga מבצעת סיווג תמונות למספר רב מאוד של סיווגים, ומציעה API שעל גביו ניתן להעביר תמונות ולקבל חזרה סיווגים. לעומתה, המערכת שלנו מסווגת ל-5 הסיווגים הרלוונטיים לנו, ואינה מציעה API מסודר בשלב זה. בניגוד למערכת שלנו, Imagga מציעה למשתמשים להגדיר סיווגים חדשים בעצמם ולבצע עריכה בסיסית לתמונות שמעלה המשתמש אל המערכת.

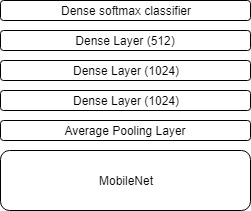
המערכת Cloud Vision מבית Google מציעה אף היא סיווג למספר רב מאוד של סיווגים. בניגוד למערכת שלנו שמציגה ללקוח רק את הסיווג בעל ההסתברות הגבוהה ביותר, Cloud Vision מציגה את ה-Top 6, יחד עם ההסתברות עבור כל סיווג.

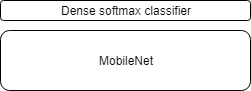
המערכת Clarifi מציעה שירות של סיווג תמונות, בדומה למערכת שלנו, אך גם של ווידאו, בניגוד למערכת שלנו. בנוסף, המערכת מאפשר למשתמש לבחור מבין מספר "מודלים", לדוגמה – מודל "אוכל", אשר יקבע את קבוצת הסיווגים האפשריים עבור התמונה. מערכת זו אינה מציעה API, והעבודה של המשתמש תעשה מול האתר לבד.

* 1. תאור המערכת
     1. מודל CNN
        1. ארכיטקטורה

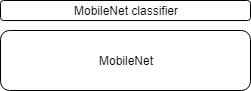
במסגרת עיצוב המערכת, ניסינו להימנע ככל הניתן מהמצאת הגלגל מחדש, ורצינו להסתמך ככל האפשר על מערכות וידע קיימים. לאור זאת, בכדי לבנות את מודל הCNN שלנו, בחרנו להיעזר בטכניקת Transfer Learning [4]. מבין שלל המודלים הניתנים לשימוש בframework של Keras, בחרנו את המודל של MobileNet, שכן מודל זה פותח כאשר הוא מיועד להיות משובץ בתוך מערכות mobile ו-embedded, ולכן מכיל מספר שכבות מועט יחסית, ובנוי משכבות depthwise convolution, אשר מצמצמות את סיבוכיות החישובים בכל שכבה [6].

כדי להמשיך וליהנות מיתרונות עבודה קודמת שנעשתה על ידי אחרים, השתמשנו במשקולות המודל MobileNet אשר נלמדו על ה-dataset של ImageNet. משקולות אלו מהוות עבורנו נקודת התחלה מצוינת, שכן כך אנו מקבלים את השכבות הנמוכות במודל, אשר מזהות מאפיינים, פינות וקווים בסיסיים, כאשר הן כבר אומנו על מעל ל-14 מיליון תמונת, הרבה מעבר ל-dataset המצומצם שיש בידינו, המכיל 4326 תמונת בלבד. מעבר לחסכון הרב בזמן אימון, עובדה זו מסייעת לנו להימנע מ-overfitting.

כמובן שהיה צורך להחליף את שכבת ה-classifier הקיימת במודל (איור 1) בכזאת המתאימה לבעיה שאנו מנסים לפתור. כזכור, ב-ImageNet ישנם 1000 סיווגים [5], בעוד בבעיה שלנו ישנם 5 בלבד. תחילה הוספנו ישירות מעל שכבות ה-CNN שכבת Dense אחת עם פונקציית אקטיביציה softmax (איור 2), אך לאחר מספר ניסויים שביצענו, הגענו לקונפיגורציה שבה מעל לשכבות הקונבולוציה של MobileNet, ישנה שכבת Average Pooling, מעליה שתיה שכבות Dense בעלות 1024 נוירונים, שכבת Dense נוספת בת 512 נוירונים, ולבסוף שכבת ה-classifier, בת 5 נוירונים, עם פונקציית אקטיבציה softmax (איור 3).



איור 1 - מודל MobileNet



איור 2 - מודל MobileNet עם classifier

איור 3 – ארכיטקטורה סופית

1. אימון המודל

שלב האימון דרש התאמות והתייחסות לעובדה שהמודל משלב שכבות מאומנות ושכבות שאינן מאומנות. בכדי לא לפגוע בשכבות התחתונות, אשר כאמור מזהות מאפיינים בסיסיים אותם תחום הבעיה שלנו חולק עם תחום הבעיה של ImageNet, הקפאנו את משקולות 20 השכבות הראשונות במודל. בחרנו את ה-optimizer שלנו להיות RMSprop, עם learning rate נמוך יחסית של , כדי שלא לשנות יותר מדי משקולות קיימים במודל.

כאמור, הdataset שלנו מצומצם מאוד, ובכדי להגדיל את מספר הדוגמאות שהשכבות העליונות של המודל שלנו רואות, חיפשנו דרכים לבצע Data Augmentation [11]. בין השיטות בהן השתמשנו, נעזרנו בהפיכה אופקית, זום, שינויים ברוחב ובגובה התמונה, וסיבוב התמונה בכיוון השעון.

במסגרת הניסויים השונים שביצענו בשלב בחירת המודל המתאים לבעיה, ניסינו לבחור את מספר הepochים המתאים ביותר, אשר ישלב כמובן בין validation accuracy גבוה, ובין זמן אימון מינימאלי. צמצום הסיבוכיות ורידוד המודל ע"י שימוש ב-MobileNet כמודל עליו התבססנו הובילו, ככל הנראה, גם למשך אימון קצר יותר. במדידות אמפיריות שערכנו, עולה כי epoch בודד במודל על גבי מודל MobileNet לקח כשמינית מן הזמן שלקח לנו לבצע במודל עמוק יותר, לדוגמה Inception V3 של GoogLeNet[14]. כמו כן, הצלחנו להשיג validation accuracy של 87% לאחר 8 epoch-ים בלבד. ניסיון להגדיל את מספר ה-epoch-ים לא הראה שיפור משמעותי בביצועי המודל, והחל ממספר epoch-ים כלשהו, באופן צפוי התחלנו לראות ירידה ב-validation accuracy.

1. ממשק משתמש

הבעיה אותה המערכת מנסה לפתור היא בעיה קלה להבנה עבור בני אדם, כזאת שהקלט והפלט הצפויים עבורה פשוטים וברורים. בפיתוח ממשק המשתמש, בחרנו לשמור ככל הניתן על פשטות זאת, כדי לאפשר למספר רב ככל האפשר של משתמשים להיעזר ולהשתמש במערכת, עם דרישות מינמאליות ככל הניתן של ידע קודם בשימוש במחשבים.

ממשק המשתמש מחולק גיאוגרפית לשני חלקים, בהתאם לשתי אפשרויות השימוש במערכת. החלק הראשון, המופיע גיאוגרפית בחלקו העליון של ממשק המשתמש, עוסק בטעינת נתונים, ודרישה מהמודל לפסק סיווג עבור הנתונים. בנוסף לציון נתיב לתיקייה שבה נמצאות התמונות אותן ברצון המשתמש לסווג, נדרש המשתמש לספק גם נתיב לקובץ h5 המכיל את המודל השלם (כלומר, ארכיטקטורה ומשקולות).

חלקה השני של המערכת, המופיע גיאוגרפית בתחתית ממשק המשתמש, עוסק בהצגת תוצאות הסיווג ועיבוד נתונים. בטבלה, אשר תופסת את מרבית שטח הממשק, יוצגו שמות הקבצים אשר נטענו מתוך הנתיב אשר אותו סיפק המשתמש, ולצידם יוצגו הסיווגים אשר התקבלו על ידי המודל. מצד שמאל של חלקו התחתון של הממשק, הוספנו אפשרות לסנן את התוצאות המוצגות בטבלה, ע"י בחירת הסיווג. כלומר, אם בחרנו Daisy, יופיעו רק תמונות אשר קיבלו את הסיווג Daisy על ידי המודל. ניתן לבחור כל שילוב שרירותי של סיווגים.

1. בדיקת וניסויים

את הבדיקות והניסויים אשר ביצענו למערכת ניתן לחלק ל2 סוגים עיקריים: ראשית רצינו לבדוק את טיבו של האלגוריתם שכתבנו ולהשוות את איכות הסיווג ביחס למערכת אחרת כאשר התנאים הם זהים. בנוסף, רצינו לבדוק את איכותו של ממשק המשתמש: האם השימוש הסטנדרטי במערכת הינו נוח, טבעי ומאפשר לבצע את הפעולות בצורה פשוטה יחסית וכמו כן גם להבין את הפלט של המערכת באופן ברור ואינטואיטיבי.

כאמור, רצינו לבדוק את איכותו של האלגוריתם. לשם כך ביצענו ניסוי אשר בדק את יעילות ודיוק סיווג המערכת שלנו על קבוצת תמונות מסוימת לעומת מערכת אחרת שביצעה אותה מטרה.   
ראשית, לשם הניסוי, שתי המערכות ביצעו אימונים על אותו Train set (קובץ אשר הכיל אלפי סוגים של תמונות שונות מחמישה סוגים שונים של פרחים) אשר נקבע מבעוד מועד. זאת ועוד, 2 המערכות הופעלו על חומרה זהה (אותו מחשב). לאחר אימונים שלקחו מספר שעות, יצרנו Test set (קובץ בדיקה בעל 50 תמונות – 10 מכל קבוצה – שונות) ועליו הרצנו את המערכת שלנו במטרה לקבל סיווג בעבור כל תמונה.

לאחר קבלת התוצאות מ2 המערכות השונות, נעזרנו בפרמטרים כדי לאמוד את רמת הדיוק של כל מערכת:  
Precision :דיוק התוצאות שהתקבלו לתוצאות האמת (כמות תוצאות הנכונות שחזרו מתוך כלל התוצאות שחזרו)  
Recall **:** מתייחס ל"רגישות" התוצאות (כמות התוצאות הנכונות שחזרו לעומת כמות התוצאות הנכונות שהיו אמורות להתקבל).  
חישוב פרמטרים אלו מתחיל בסיווג התוצאות ל4 תתי סוגים של תוצאות:  
חיובי נכון (True positive): תמונה אשר סווגה מסוג מסוים ואכן כזו.

חיובי לא נכון (True negative): תמונה אשר סווגה מסוג מסוים והיא למעשה לא כזו.

שלילי נכון (False positive): תמונה אשר סווגה כאחת מסוג אחר ואכן היא כזו.

שלילי לא נכון (False negative): תמונה אשר סווגה כאחת מסוג אחר והיא מהסוג הנכון.  
ולאחר הרצת המערכת ביצענו את חישוב הפרמטרים ע"פ הנוסחאות הבאות:

להלן מוצגים הסיווגים השונים שהתקבלו בעבור המערכות השונות. בסימונים מטה, המערכת

שפיתחנו מסומנת ב-System A, והמערכת של הצוות המקביל מסומנת ב-System B:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| System A | | | | | |
|  | Daisy | Dandelion | Tulip | Rose | Sunflower |
| True Positive | 10 | 8 | 9 | 9 | 10 |
| False Positive | 0 | 0 | 1 | 1 | 2 |
| True negative | 40 | 40 | 39 | 39 | 38 |
| False negative | 0 | 2 | 1 | 1 | 0 |
| Precision | 1 | 1 | 0.9 | 0.9 | 0.8334 |
| Recall | 1 | 0.8 | 0.9 | 0.9 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| System B | | | | | |
|  | Daisy | Dandelion | Tulip | Rose | Sunflower |
| True Positive | 10 | 9 | 10 | 0 | 6 |
| False Positive | 1 | 2 | 10 | 0 | 0 |
| True negative | 39 | 38 | 30 | 40 | 40 |
| False negative | 0 | 1 | 0 | 10 | 4 |
| Precision | 0.8334 | 0.75 | 0.5 | 0 | 1 |
| Recall | 1 | 0.9 | 1 | 0 | 0.6 |

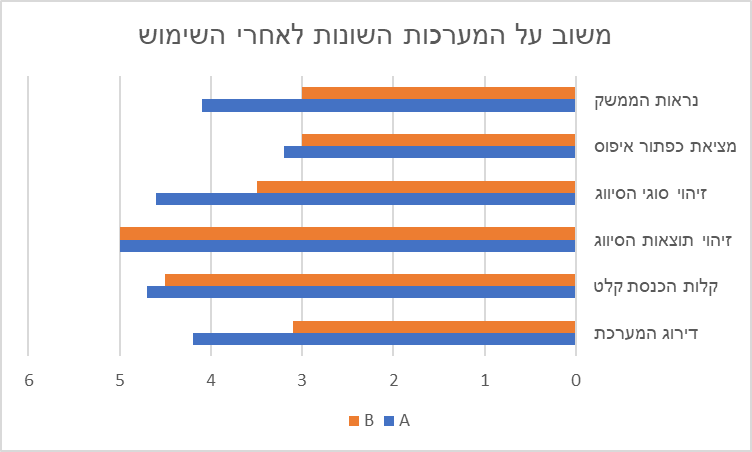
לאחר ריכוז התוצאות הצגנו את הנתונים השונים בגרף של precision כתלות בrecall לצורך בדיקת טיב המערכת. (כאשר דגמנו בעבור כל סוג פרח את הנתונים הנ"ל)

מנתוני הגרף ניתן לראות כי המערכת שלנו, System A, הינה יציבה ומחזירה תוצאות מדויקות יותר לעומת המערכת השנייה. היציבות באה לידי ביטוי בכך שבעבור כל סוג של פרח שאובחן במערכת הראשונה ניתן לקבל דיוק יחסית זהה בעוד שהמערכת השנייה מציגה ביצועים לא יציבים ומשתנים בעבור כל סוג פרח. בכך איששנו את השערת הניסוי ואכן המערכת שלנו יציבה יותר ומדויקת יותר מהשנייה.

לטובת בדיקת איכות הממשק, יצרנו והעברנו שאלון מקדים באוכלוסיית הנבדקים כדי לאפיין את שגרת שימושם. בשאלון התייחסנו לשימוש במכשיר הסלולארי: מהו השימוש השכיח ביותר, תדירות השימוש במכשיר והכרת המכשיר. בנוסף בדקנו את שגרת שימושו בעולם הפרחים: מהי תדירות שבה הנשאל מטייל ומהי מידת ההיכרות שלו עם סוגי הפרחים. לאחר מכן ביצעו הנסיינים סדרת פעולות אשר קיימת ב2 המערכות :

* 1. בחרו תיקייה ובה תמונות הפרחים אותם רצו לסווג.
  2. בחרו את המודל הרצוי לסיווג (קובץ מסוג h5).
  3. הריצו את מערכת הסיווג.
  4. לאחר החזרת התוצאות:
     1. בחרו סוג פרח מסוים, כמה תמונות הן מהסוג הנבחר?
     2. כמה סוגי פרחים ישנם?
  5. בצעו reset למערכת.

לאחר ביצוע סדרת הפעולות הנ"ל בדקנו בעזרת שאלון נוסף את שביעות רצונם של המשתמשים השונים. להלן התוצאות:



1. סיכום

בשגרת יומנו , עולם הצומח הינו חלק אינטגרלי שמקיף אותנו בהמון תחומים. עם התקדמות הטכנולוגיה ובאופן האינטנסיבי בו השימוש במצלמה הפך להיות כמעט מיידי – עיבוד תמונות בזמן אמת הופך להיות צורך גדול מיום ליום. במקביל, בעולם הצומח, זיהוי תמונות של פרחים בפרט הופך להיות בעל חשיבות גדולה. החל מצורך בסיסי אשר קיים כבר שנים ברצון לזיהוי סוגי פרחים )וקיבל מענה ממגדירי פרחים כאלו ואחרים( ועד לצורך שגדל לאחרונה בשוק החקלאות כאשר נכתבים פיצ'רים ופטנטים שונים לאחרונה ויש צורך לזהות האם תמונות מסוימות של פרחים כבר קיימות במאגר או לא [1]. העבודה הידנית הינה מאוד אינטנסיבית ואיטית ועל כן נדרש מענה אוטומטי ומהיר.

על כן , הבעיה אשר אנו באים לפתור הינה סיווג הפרחים. הסיווג התבצע תוך התבססות על מודלים של למידת מכונה, אשר במרכזו מודל CNN שאומן כבר על מאגר תמונות של ImageNet – בכך נחסכו עבורנו שעות פיתוח ארכיטקטורה ואימון רבות. לאחר בניית מודל מאומן אשר נבדק מול dataset של תמונות שונות (בשמו mobileNet), פיתחנו מערכת אשר בהינתן קובץ מאומן ותיקיית תמונות מאפשר למשתמש לקבל סיווג ברור בעבור כל תמונה אשר שלח. זאת ועוד , ביצענו בדיקה אל מול מערכת אחרת לצורך השוואת איכות הסיווג והתגלה כי הדיוק של המערכת שלנו כמדויק יותר (לצורך השוואות התוצאות נעזרנו בפרמטרים precision ו – recall).

אנו מוצאים מספר נקודות אשר ניתנות לפיתוח נוסף או שיפור: ראשית ניתן לשפר את ביצועי המודל המאומן שלנו. לדוגמא, בעזרת שימוש בmobileNet #2, או בהגדלת ה-Training Set עליו מאומן המודל. כמובן שניתן לחקור מודלים מאומנים רבים אחרים, אשר מגבלת הזמן לא אפשרה לנו לחקור את ביצועיהם. שנית, נוכל לשפר את שימוש המערכת עצמה בכך שנשלבה במסגרות שונות כדוגמת פלאפונים סלולריים וכיוצא בזה, ולאפשר פונקציונאליות לפעולות זריזות וישירות לגבי חיזוי. בנוסף, ניתן לפתח את השימוש גם בממשקים הקיימים: אפשור ברמת הצגת תוצאות הסיווג – אחוז הדיוק שעלה במודל, הצגת התמונה עצמה ועוד.

נדמה שככל שהשימוש במכשירים ניידים הופך שכיח אצל מרבית האוכלוסייה בעולם, שימוש בעיבודי תמונה בכלל ובסיווג תמונות בפרט תופס מקום גדול יותר ויותר במרוצת השנים בעקבות המהירות והדיוק שניתן להגיע בעת הסיווג.

1. ביבליוגרפיה

|  |
| --- |
| [1] Guru, D. S., YH Sharath Kumar, and S. Manjunath. "Textural features in flower classification." Mathematical and Computer Modelling 54.3-4 (2011): 1030-1036.‏ |
|  |
|  |  |
| [2] Gehler, Peter, and Sebastian Nowozin. "On feature combination for multiclass object classification." 2009 IEEE 12th International Conference on Computer Vision. IEEE, 2009.‏ |
|  |  |
| [3] Nilsback, Maria-Elena, and Andrew Zisserman. "Automated flower classification over a large number of classes." 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing. IEEE, 2008.‏ |
|  |
|  |  |
| [4] Shin, Hoo-Chang, et al. "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning." IEEE transactions on medical imaging 35.5 (2016): 1285-1298.‏ |
|  |
| [5] Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009.‏ |
|  |

[6] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).‏

[7] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.‏

[8] Dreiseitl, Stephan, and Lucila Ohno-Machado. "Logistic regression and artificial neural network classification models: a methodology review." *Journal of biomedical informatics* 35.5-6 (2002): 352-359.‏

[9] LeCun, Yann, et al. "Object recognition with gradient-based learning." *Shape, contour and grouping in computer vision*. Springer, Berlin, Heidelberg, 1999. 319-345.‏

[10] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).‏

[11] Perez, Luis, and Jason Wang. "The effectiveness of data augmentation in image classification using deep learning." *arXiv preprint arXiv:1712.04621* (2017).

[12] Wang, Baiyang, and Diego Klabjan. "Regularization for unsupervised deep neural nets." *Thirty-First AAAI Conference on Artificial Intelligence*. 2017.

[13] Xiang, Sitao, and Hao Li. "On the effects of batch and weight normalization in generative adversarial networks." *arXiv preprint arXiv:1704.03971* (2017).

[14] Szegedy, Christian, et al. "Going deeper with convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.