

דו"ח סיכום פרויקט: א'

# Fine-Grained Classification of Objects from Aerial Imagery

Tomer Navon  
Shalev Shaer

תומר נבון  
שלו שער

Netanel Ratner

מנחה: נתנאל רטנר

סמסטר רישום: אביב תשע"ט

תאריך הגשה: אוקטובר, 2019

# תוכן עניינים

1.	תקציר .....	2
2.	מבוא .....	3
2.1	מוטיבציה .....	3
2.2	המטרה – פתרון בעיית הסיווג המורכב .....	3
2.3	הגישה בפרוייקט שלנו .....	4
3.	רקע - למידה עמוקה .....	5
3.1	רקע אלגוריתמי .....	5
3.2	סקירת ספרות .....	8
4.	שלבי הפתרון .....	9
4.1	שלב ראשון – ניסיון Plug & Play .....	9
4.2	שלב שני – ניסיון בשיטות ML קלאסיות .....	10
4.3	שלב שלישי – פתרון בעיית הסיווג למחלקה העיקרית .....	10
4.4	שלב רביעי – פתרון בעיית המאפיינים בצורה היררכית .....	11
5.	תוצאות .....	12
6.	סיכום ומסקנות .....	16
	רשימת מקורות .....	17

# 1. תקציר

בשנים האחרונות השימוש בטכנולוגיות מבוססות Deep Learning (DL) לפתרון בעיות שונות תפס תאוצה משמעותית. אנו עדים כיום לשימוש הולך וגובר של אפליקציות שונות בכלים אלו לפתרון בעיות קיימות. אחד התחומים בהם השימוש בכלים מבוססים DL נפוץ במיוחד, ובו גם יתמקד נושא הפרויקט, הוא תחום הראיה הממוחשבת. למשל, בעזרת שימוש בכלי מתאים, ניתן לאמן רשת נוירונים שתדע לסווג תמונות למחלקות מוגדרות מראש.

במהלך שנת 2019 התבצעה תחרות בתחום הראייה הממוחשבת מטעם משרד הבטחון :

MAFAT Challenge : <https://mafatchallenge.mod.gov.il/>

התחרות כללה סיווג תמונות רחפן לפי קטגוריות ומאפיינים נתונים, תוך שימוש בלמידה עמוקה.

הקטגוריות חולקו למעט מחלקות ראשיות ותתי מחלקות רבות.

בפרויקט זה התמודדנו בהצלחה עם האתגר המוצע בתחרות הנ"ל, תוך שימוש בהרכבת רשתות קונבולוציה על מנת לסווג בצורה היררכית את המאפיינים השונים (עקב תלות במאפיינים) הנדרשים בתחרות.

## 2. מבוא

### 2.1. מוטיבציה

בשנים האחרונות השימוש ברחפנים נעשה פופולרי ויומיומי עקב שיפור הטכנולוגיה והוזלת העלויות. לרחפנים יכולים להיות שימושים רבים, כאשר אחד הנפוצים הוא השימוש הבטחוני. בפרט, זיהוי רכבים לפי מאפיינים נתונים יכול לשמש בפתרון בעיות ביטחוניות כגון: זיהוי ועקיבה של רכב גנוב, זיהוי ועקיבה של רכב מפגע וכו'. לרכבים יכולים להיות מאפיינים רבים, כגון: סוג, צורה, צבע, תוספות על הרכב וכו'. סיווג לפי כמות גדולה יותר של מאפיינים יכולה להביא לידי זיהוי ודאי יותר, אך כרוך באתגר טכנולוגי מורכב.

### 2.2. המטרה – פתרון בעיית הסיווג המורכב

בפרויקט נדרשנו לסווג תמונות של כלי רכב מתצלומי רחפן (בסיס הנתונים נתון) לפי המחלקות והמאפיינים הבאים:

1. מחלקה - כל אובייקט מסווג לאחת מהמחלקות העיקריות הבאות: 'כלי רכב גדולים' או 'כלי רכב קטנים'.
2. תת-סוגים - אובייקטים מסווגים לתתי-מחלקות לפי תפקידם או ייעודם, למשל: מערבל בטון, משאית מנוף, מוביל וכו'. כל אחד מהאובייקטים צריך להיות מוקצה לתת-מחלקה אחת.
3. מאפיינים - אובייקטים מסומנים על פי מאפייניהם. לדוגמא: סולם, הרוס, sunroof וכו'. כל אובייקט עשוי להיות עם מספר תכונות שונות.
4. צבע - אובייקטים מתויגים בצבע הניתן להם (הצבע המוגדר ע"י אדם. למשל: משאית עם קבינה אדומה וארגז לבן תתויג כאדומה). לדוגמא: כחול, אדום, צהוב וכו'. כל אובייקט כולל ערך צבע יחיד.

להלן תיאור מלא של היררכיית התיוג של מערך התחרות:

1. רכב קטן
  - a. תת-סוגים - סדאן, האצ'בק, מיניוואן, ואן, טנדר, ג'יפ, רכב ציבורי.
  - b. מאפיינים - גג שמש, גגון, תא מטען פתוח, תא מטען סגור, הרוס, גלגל חלופי.
  - c. צבעים - צהוב, אדום, כחול, שחור, כסף / אפור, לבן, אחר.
2. רכב גדול
  - a. תת-סוגים - משאית, משאית קלה, מערבל בטון, רכב חקלאי ייעודי, משאית מנוף, מוביל, מכלית, אוטובוס, מיניבוס.
  - b. מאפיינים - תא מטען פתוח, פתחי אוורור, הרוס, תא מטען סגור, סולם, ארגז עם מעטפת רכה, רתום לעגלה.
  - c. צבעים - צהוב, אדום, כחול, שחור, כסף / אפור, לבן, אחר.

מכאן שאנו נדרשים לפתח מודל המסווג את בסיס הנתונים הנתון בבעיה לפי הסיווג המורכב הנזכר לעיל.

כלומר, אם להשתמש בביטויים מתחום הלמידה החישובית, אנו נוקטים בגישה של **למידה מונחית**, שהיא טכניקה בלמידה חישובית על מנת ללמוד פונקציה ("היפותזה") מורכבת על בסיס סט אימון. סט האימון מכיל דגימות מתוך התפלגות משותפת של מרחב האלמנטים ומרחב התיוגים. מטרת הלמידה המונחית היא ללמוד על ההתפלגות המשותפת, ומתוך כך לנבא עבור אלמנט את התיוג שלו. בעיות שבהן עוסקת למידת מונחית כוללות בין היתר רגרסיה ובעיות סיווג.

## **2.3. הגישה בפרויקט שלנו**

בפרויקט שלנו ננסה לפתור את הבעיה על ידי כלים חדשניים בתחום הראייה הממוחשבת. הכלי המרכזי אשר ישמש אותנו הינו אלגוריתם המכונה deep learning ובפרט Deep Convolution Neural Networks. מדובר בתחום בו מיישמים פתרונות של בעיות מממד גבוה על ידי מערכת לומדת מסוג רשת נוירונים עמוקה.

**הגישה ההיררכית** - בפרויקט זה נקטנו בגישת סיווג היררכי של המאפיינים ע"מ לשפר תוצאות ולהתמודד עם הבעיות והאתגרים שבסיווג מורכב מעין זה. בגישה זו השתמשנו בתוצאות הסיווג של המחלקה העיקרית ע"מ לסווג את המאפיינים בצורה תלויה. גישה זו אפשרית עקב התלות בין המאפיינים המסווגים לבין המחלקה העיקרית, כפי שתואר בפירוט המשימה.

### 3. רקע - למידה עמוקה

בפרק זה נסקור בקצרה את הבסיס האלגוריתמי לגישת הלמידה העמוקה, האלגוריתם בו בחרנו ככלי לפתרון הבעיה. ראשית נזכיר את הרקע לאלגוריתם הלמידה. שנית, נסביר את הרקע של רשתות קונבולוציה – (CNN). לבסוף, נתן רקע לאבן הבניין של הסיווג ההיררכי (שיפורט בפתרון שלנו) שהיא רשת הקונבולוציה ResNet.

#### 3.1. רקע אלגוריתמי

למידה עמוקה הינה כאמור תחום התופס תאוצה ובעל פופולריות גוברת. הצלחותיו בתחומים רבים ומגוונים, ביניהם גם בתחום זיהוי אובייקטים בתמונות, עיבוד תמונות וראיה ממוחשבת. **רשת עצבית מלאכותית**, (ANN – Artificial Neural Network) **רשת נוירונים או רשת קשרית** הוא מודל מתמטי חישובי שפותח בהשראת תהליכים מוחיים או קוגניטיביים המתרחשים ברשת עצבית טבעית ומשמש במסגרת למידת מכונה. רשת מסוג זה מכילה בדרך כלל מספר רב של יחידות מידע (קלט ופלט) המקושרות זו לזו. צורת הקישור בין היחידות, המכילה מידע על חוזק הקשר, מדמה את אופן חיבור הנוירונים במוח. השימוש ברשתות עצביות מלאכותיות נפוץ בעיקר במדעים קוגניטיביים ובמערכות תוכנה שונות, בהן מערכות רבות של אינטליגנציה מלאכותית המבצעות משימות מגוונות -זיהוי תווים, זיהוי פנים, זיהוי כתב יד, חיזוי שוק ההון, מערכת זיהוי דיבור, זיהוי תמונה, ניתוח טקסט ועוד.

רשת נוירונים מאופיינת על ידי:

- חיבורים - אופן החיבור בין הנוירונים ברשת
- משקלים - השיטה הקובעת את משקלי החיבורים בין הנוירונים
- פונקציית האקטיבציה, העשויה להיות שונה בכל שכבה (פונקציה לא ליניארית, לרוב  $\text{ReLU}$ ,  $\text{logsig}$  או סיגמואיד)

רשתות נוירונים מורכבות ממספר רב של יחידות עיבוד פשוטות הנקראות נוירונים, אשר מחוברות באופן היררכי ומובנות בשכבות. השכבה הראשונה נועדה לקלוט מידע לרשת, השכבה האמצעית ידועה כשכבה החבויה (במודלים שונים עשויים להיות יותר מאחת כזו), ולבסוף השכבה האחרונה אשר נועדה להחזיר את המידע המעובד כפלט. הצמתים בכל שכבה מחוברים באופן מלא לצמתים בשכבות הסמוכות באמצעות חיבור ישיר בין הנוירונים, כאשר לכל קשר קיים משקל מסוים. המשקל בכל קשר קובע עד כמה רלוונטי המידע שעובר דרכו, והאם על הרשת להשתמש בו על מנת לפתור את הבעיה. כל צומת בשכבת הקלט (השכבה הראשונה) מייצג תכונה שונה מהמבנה, ושכבת הפלטים מייצגת את הפתרון של הבעיה. בשכבה

האמצעית והחיצונית קיימים "ערכי סף" הניתנים לכיול במערכת ממוחשבת, וקובעים את חשיבות הקשרים השונים.

ברשת ישנן 3 סוגי שכבות:

1. שכבת כניסה (Input Layer) - לכל תא בשכבה זו כניסה אחת. וקטור הכניסות הוא הקלט לרשת. מספר התאים: כמספר המאפיינים (Features). מומלץ לבצע בדיקה בנוגע להשפעתם של המאפיינים השונים על השגיאה, הקשרים וההיררכיה ביניהם. כלל אצבע הוא כי המספר הכולל של המאפיינים לא יעלה על עשירית ממספר הדוגמאות בסדרת הלימוד.

2. שכבות חבויות (Hidden Layers) - לכל תא בשכבה זו מספר כניסות, כמספר תאי הכניסה (Fully Connected). מספר התאים: כמספר התאים בשכבת הכניסה בתוספת תא אחד או שניים. מספר השכבות: מ-0 ועד לאינסוף. בכל השכבות החבויות בדרך כלל אותו מספר תאים. מומלץ לבצע בדיקה בנוגע להשפעתם של מספר התאים ומספר השכבות החבויות על השגיאה. מספר השכבות ומספר התאים מגדירים את גודל הרשת. יש לבחור את הרשת גדולה מספיק, אך לא גדולה מדי. רשת קטנה מדי לא תוכל לקרב בדיוק מספיק את המיפוי הנדרש, ואילו רשת גדולה מדי תמנע לימוד יעיל (זהו ה-Bias Trade Off). רשתות רב-שכבתיות אופייניות הן בעלות שכבה חבויה אחת או שתיים, ומספר הנירונים בשכבות הנסתרות במקרים רבים אינו עולה על 10 (אלא אם כן קיימת חלוקה מודולרית למספר תת-רשתות). בעיות זיהוי תמונה בדרך כלל מאופיינות עם מספר שכבות גבוה מאוד (מאות שכבות). כוונון מספר הנירונים ומבנה הרשת יעשה לרוב אמפירית, תוך שימוש בדוגמאות ולידציה (Validation) או באמצעות אימות-צולב (Cross Validation).

3. שכבת יציאה (Output Layer) - לכל תא בשכבה זו מספר כניסות, כמספר תאי השכבה המוסתרת (Fully Connected). וקטור מוצאי התאים בשכבה זו הוא וקטור היציאה של הרשת. מספר התאים: כמספר ה-Classes.

דרך האימון:

הקלטים לנירון מסוכמים בסכום משוקלל, ואם התוצאה עוברת ערך סף מסוים, הקלט מועבר ליציאה. המשקולות שעל הקשרים בין הנירונים (המשקולות הסינופטיות) הן אלה שאוגרות את הידע של הרשת במודל המפושט, לכן אלגוריתמי הלימוד השונים מבצעים את כוונון המשקולות לערכים שנותנים תוצאות חישוב טובות עבור מערך דוגמאות הלימוד. משימה זאת אינה פשוטה בגלל השכבות החבויות.

באופן יותר ספציפי, בחרנו להשתמש ברשתות קונבולוציה. בלמידת מכונה, רשת קונבולוציה (CNN), או (ConvNet) היא סוג של רשת מלאכותית, המכוונות בדרך כלל לניתוח דימויים חזותיים.

CNNs משתמשים בווריאציה של פרספטרונים, אבן בסיס לרשת עצבית, שתוכננו כך שיצטרכו עיבוד מקדים מינימלי. ברשתות קונבולוציה, בכל שכבה, מתבצע תהליך קונבולוציה בין מסנן לתמונה, ומה שנלמד הוא המסנן שמביא לשגיאה המינימלית.

רשתות קונבולוציוניות קיבלו השראה מתהליכים ביולוגיים, בכך שתבנית הקישוריות בין נוירונים דומה לארגון הקורטקס החזותי של המוח. נוירונים קליפתיים בודדים מגיבים לגירויים רק באזור מוגבל של שדה הראייה המכונה שדה פתוח. שדות פתוחים של נוירונים שונים חופפים באופן חלקי כך שהם מכסים את שדה הראייה כולו.

CNNs זקוקים למעט יחסית עיבוד מראש לעומת אלגוריתמים אחרים. פירוש הדבר שהרשת לומדת את המסננים שבאלגוריתמים המסורתיים היו מתוכננים ידנית. עצמאות זו מידע קודם וממאמץ אנושי בתכנון תכונות היא יתרון גדול.

אחד מאלגוריתמי הלימוד המקובלים ביותר הוא אלגוריתם "חלחול לאחור" (Back-Propagation או BP). זהו אלגוריתם ללימוד מונחה (Supervised Learning). הוא מעביר את קלט הלימוד במורד רשת עם משקולות אקראיות, ומשווה את הפלט המתקבל לפלט הרצוי בכך שהוא מחשב את הטעות עבור כל נוירון פלט.

הרעיון העיקרי של BP הוא שהטעויות מחלחלות חזרה במעלה הרשת כך שמתבצעים כוונוני משקולות בעזרת אלגוריתם Gradient descent, שמנסה להקטין את הטעויות בפלט. אם תיקוני המשקולות נעשים בשיעור מספיק קטן, גדלים הסיכויים להתכנסות של הרשת לטעות מינימלית. אומנם תיתכן התכנסות למינימום מקומי בלבד, אך למעשה השיטה נותנת תוצאות טובות כיוון שניתן להפעילה שוב עם משקולות התחלתיות שונות. אף על פי שהלימוד בשיטה זו יכול להיות איטי, השימוש ברשת אחרי הלימוד הוא מהיר מאוד.

כיום יש מימושים רבים ושונים של רשתות קונבולוציה, כאשר אנו בחרנו להשתמש ב- ResNet לפתרון בעיית הסיווג. **רשת עצבית שיווי (ResNet)** היא רשת עצבית מלאכותית (ANN) מסוג שבונה על המבנים הידועים מן תאים פירמידליים בקליפת המוח. רשתות עצביות שיוויות עושות זאת על ידי שימוש בחיבורי דילוג, או באמצעות קיצורי דרך לקפיצות מעל כמה שכבות.

אחת המוטיבציות לדלג על שכבות היא להימנע מבעיית היעלמות הגרדיאנטים (הכפלה בכל שכבה בערך שבין 0 ל-1 מקטין את הגרדיאנט ומקשה על הלימוד), על ידי שימוש חוזר בפעילויות משכבה קודמת עד שהשכבה הסמוכה לומדת את משקולותיה.



דילוג מפשט את היעילות של הרשת, תוך שימוש בפחות שכבות בשלבי האימונים הראשוניים. הדילוג מזרז את הלמידה על ידי הקטנת ההשפעה של מדרגות נעלמות, מכיוון שיש פחות שכבות להפיץ. לאחר מכן הרשת משחזרת בהדרגה את השכבות המדלגות כאשר היא לומדת את מרחב התכונות.

## 3.2. סקירת ספרות

כהכנה לפרויקט זה עסקנו בעיקר בסקירה של עולם הלמידה העמוקה. מכיוון שבעיית הסיווג היא הבעיה שעימה פרצה הלמידה העמוקה לתודעה, רוב המאמרים שסקרנו היו מאמרים "קלאסיים" בתחום. בנוסף, קראנו גם מאמרים בתחום של פתרון בעיות נפוצות בבעיות סיווג ומאמר על סיווג היררכי.

- הרשת שעימה פרצה הלמידה העמוקה לתודעה, הייתה רשת הסיווג AlexNet [1]. ע"י הבנה של עקרון הפעולה של רשת זו למדנו על רשתות סיווג בכלליות.
- הרשת שאיתה בחרנו לעבוד הייתה ResNet [2], שנבחרה עקב היעילות הרבה שלה ואחוזי הדיוק הגבוהים בבעיות סיווג.
- במהלך העבודה נתקלנו בבעיית חוסר האיזון במידע הנתון, ולכן נעזרנו במאמר [3]. בעזרת מאמר זה הגענו להחלטה על אוגמנטציות סלקטיביות ומשקול שונה של פונקציית המחיר כפי שיתואר בהמשך.
- את ההשראה לסיווג היררכי קיבלנו ממאמר [4]. במאמר זה מתואר סיווג היררכי של טקסטים בעזרת רשתות קונבולוציה.

## 4. שלבי הפתרון

### 4.1. שלב ראשון – ניסיון Plug & Play

בשלב הראשון ניסינו לפתור את בעיית הסיווג למחלקה העיקרית – רכב גדול \ רכב קטן. אימנו רשת מסוג ResNet18 עם המידע הנתון וקיבלנו תוצאות מעולות של 96%. לאחר בחינה של התוצאות, הבנו שהרשת לא מסווגת בצורה טובה (על אף התוצאות ה"מרשימות"), שכן הרשת סיווגה את כלל התמונות כ"רכב קטן". תוצאה זו נבעה מהבעיה המרכזית איתה התמודדנו בפרויקט – בעיית ה **Imbalanced Data**. בעיית ה **Imbalanced Data** היא בעיה שכיחה בלמידת מכונה (בעיית הסיווג), כאשר יש יחס לא פרופורציונאלי של דגימות בכל מחלקה. מכיוון שבלמידה עמוקה אנו מנסים למזער פונקציית מחיר, מידע לא מאוזן בצורה קיצונית יכולה לגרום תוצאת חיזוי דטרמיניסטית (למשל, אם 95% מהדגימות הן ממחלקה '1', אם הרשת תחזה תמיד '1' היא תדייק ב-95% מהמקרים, מה שיגרור כנראה מחיר נמוך כנדרש).

דוגמא לחוסר האיזון במידע הנתון לנו בבעיה :

מחלקה עיקרית		
גודל	כמות	אחוז מכלל המידע
רכב קטן	8651	96.40%
רכב גדול	317	3.60%
צבע		
גודל	כמות	אחוז מכלל המידע
שחור	939	11%
כחול	576	6.80%
ירוק	55	<1%
צהוב	57	<1%
אדום	252	3%
אפור/כסף	3004	36%
לבן	3157	37%
אחר	407	5%

מכאן שהבנו שפתרון הבעיה כרוך בעבודה נוספת מעבר ל **Plug & Play** .

## 4.2. שלב שני – ניסיון בשיטות ML קלאסיות

לאחר שהבנו את בעיית חוסר האיזון, ניסינו לפתור את בעיית הסיווג למחלקה העיקרית בעזרת שיטות קלאסיות:

- סיווג בעזרת ניתוח מאפיין השטח של האובייקט המסווג. בפועל גילינו שמאפיין זה לא מפריד את המידע בצורה מושלמת, ולכן הגענו לרמת ביצועים של כ 70% (לא מספק).
- סיווג בעזרת ניתוח מאפיין המימד הארוך של האובייקט המסווג. בפועל גילינו שמאפיין זה לא מפריד את המידע בצורה מושלמת, ולכן הגענו לרמת ביצועים של כ 75% (לא מספק).

## 4.3. שלב שלישי – פתרון בעיית הסיווג למחלקה העיקרית

בשלב זה חזרנו להשתמש בלמידה עמוקה תוך התמודדות עם בעיית חוסר האיזון בעזרת הצעדים הבאים:

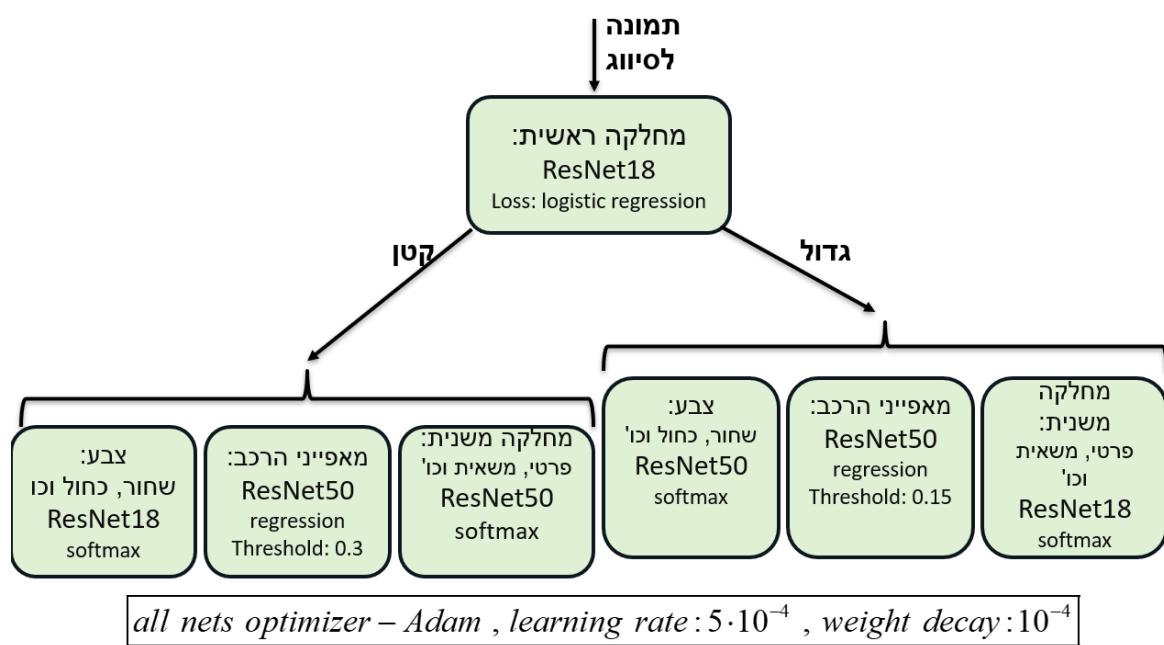
- **אוגמנטציה סלקטיבית.** מכיוון שנקודת העבודה בבעיה זו לא תלויה בזווית של האובייקט, ביצענו אוגמנטציה סיבובית לאובייקטים מסוג "רכב גדול", וכך הגדלנו את כמות הדגימות מהמחלקה. פעולה זו לא הביאה לשוויון מלא בין כמות הדגימות, ולכן השתמשנו בנוסף במשקול פונקציית המחיר.
- **משקול שונה של פונקציית המחיר.** כפי שתיארנו ברקע התאורטי, פונקציית המחיר נותנת "קנס" על שגיאה בסיווג. מכיוון שהיו לנו דוגמאות מעטות מהמחלקה של "רכב גדול", הגדרנו קנס גדול יותר (בצורה הפוכה לשכיחות שלו במידע) לשגיאה במחלקה "רכב גדול" מאשר לשגיאה בסיווג של "רכב קטן".
- **ניקוי המידע.** בבסיס הנתונים הנתון היו כ 3% של תמונות שנלקחו מגובה גבוה משמעותית משאר התמונות (מה שיצר נקודת עבודה שונה). ניקוי תמונות אלה מבסיס הנתונים הביא לשיפור משמעותי בתוצאות.

לאחר פעולות אלה אימנו רשת ResNet18 והגענו לכ 98% דיוק עבור סט הבחן.

## 4.4. שלב רביעי – פתרון בעיית המאפיינים בצורה היררכית

בשלב סיווג המאפיינים, ניסינו תחילה לסווג את המאפיינים ע"י רשת אחת שאינה תלויה בתוצאות הסיווג למחלקה העיקרית. רשת זו הייתה אמורה לסווג לפי כל המאפיינים (גם אם הרכב המקורי הוא קטן – נסווג גם את המאפיינים ששייכים רק לרכב גדול, ולהפך). פתרון זה הביא לתוצאות מניחות את הדעת, אך לא לרמה של State Of The Art.

לאור פתרון בעיית המחלקה העיקרית באחוזי דיוק גבוהים, ובהשראת מאמר הסיווג ההיררכי למסמכי טקסט [4], נקטנו בגישה הסיווג ההיררכי המתוארת להלן :



כעת בעזרת הארכיטקטורה ההיררכית התלויה, הבאנו לנטרול גרירת חוסר האיזון לבעיות המשניות, שכן עבור כל תת מחלקה (רכב גדול \ רכב קטן) המידע מאוזן ביחס למאפיינים. בנוסף, הסיווג ההיררכי התמודד עם הבעיה שנוצרה מהגדרה שונה של מאפיינים עבור רכב גדול ורכב קטן (הגדרה של משאית כחולה שונה בתכלית מהגדרה של רכב פרטי כחול).

כל מודל המרכיב את הארכיטקטורה הוא רשת העומדת בפני עצמה, המאומנת רק עם אובייקטים המתאימים לה (הרשתות הימניות מאומנות עם רכבים גדולים, והשמאליות עם רכבים קטנים).

ארכיטקטורה היררכית זו מתאפשרת בעיקר בעקבות אחוז דיוק גבוה בסיווג המחלקה העיקרית, שכן הדיוק הסופי מוכפל בדיוק זה.

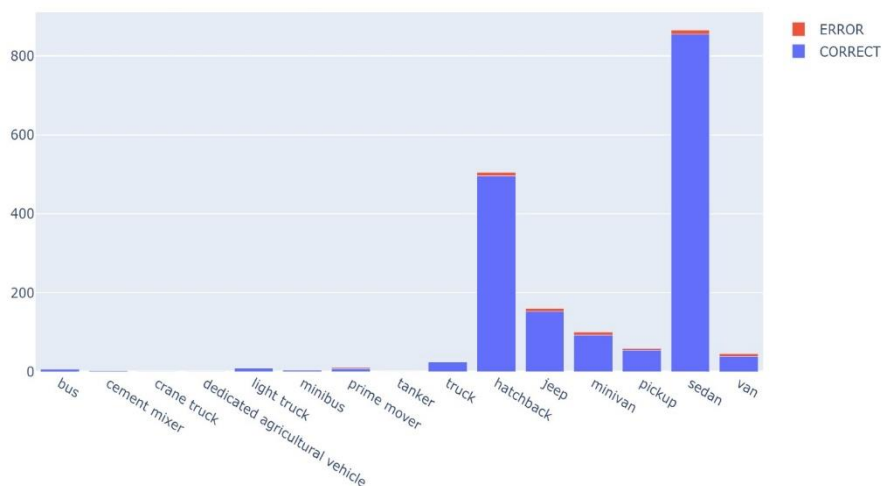
## 5. תוצאות

סט האימון שלנו הורכב מ 1795 אובייקטים (20% מכלל בסיס הנתונים) שנבחרו בצורה רנדומלית, ומייצגת את התפלגות המידע בכל מחלקה. כמו שתיארנו בבית המידע הלא מאוזן, חשוב לבחון את התוצאות לא רק בצורה כללית, אלא גם בצורה שתמחיש שהסיווג הצליח גם עבור המחלקה עם מעט דגימות (כדי לראות שהמודל לא מתייג תמיד כמחלקה עם הייצוג הגדול). במקרה שלנו, הוספנו תוצאות עבור המחלקה עם הייצוג הנמוך ע"מ להמחיש שגם עבורן קיבלנו תוצאות טובות.

כמו שכבר הזכרנו, עבור המחלקה העיקרית הגענו לכ – 98% דיוק, כאשר הדיוק עבור רכבים גדולים (המחלקה עם כמות דגימות נמוכה יותר) הוא 100%.

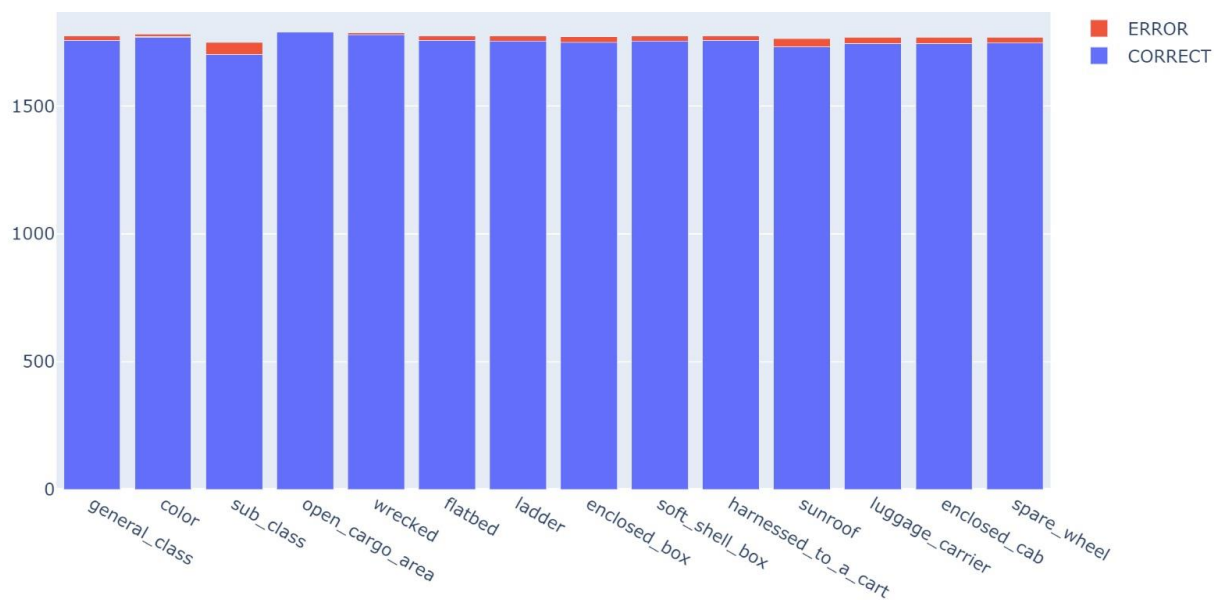
### עבור סיווג המחלקה המשנית :

Accuracy of the bus network on the 7 test images: 100 %
Accuracy of the cement mixer network on the 3 test images: 100 %
Accuracy of the crane truck network on the 1 test images: 100 %
Accuracy of the dedicated agricultural vehicle network on the 1 test images: 100 %
Accuracy of the light truck network on the 9 test images: 100 %
Accuracy of the minibus network on the 4 test images: 100 %
Accuracy of the prime mover network on the 10 test images: 80 %
Accuracy of the tanker network on the 0 test images: NaN
Accuracy of the truck network on the 25 test images: 100 %
Accuracy of the hatchback network on the 505 test images: 98 %
Accuracy of the jeep network on the 160 test images: 96 %
Accuracy of the minivan network on the 101 test images: 92 %
Accuracy of the pickup network on the 58 test images: 95 %
Accuracy of the sedan network on the 865 test images: 99 %
Accuracy of the van network on the 46 test images: 85 %

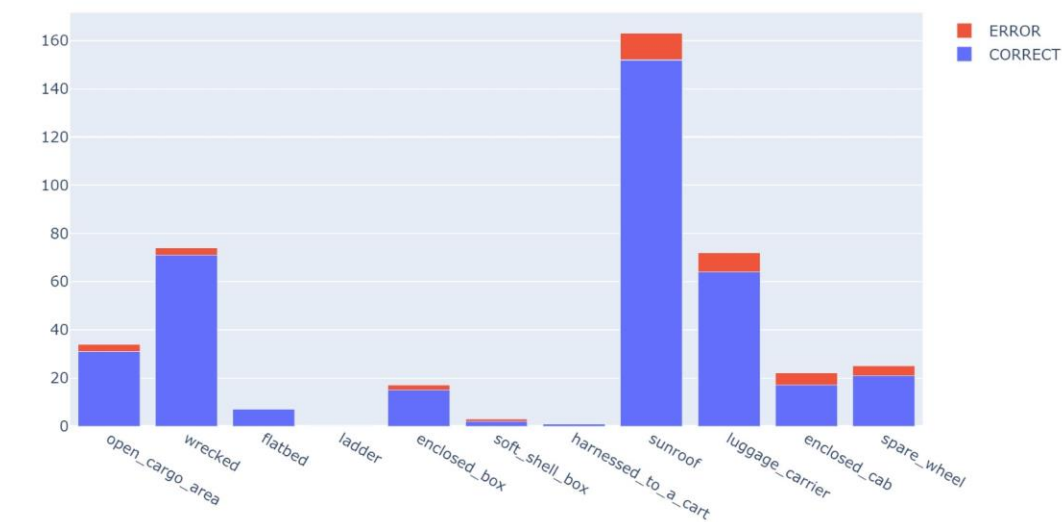


## עבור סיווג המאפיינים :

### עבור כלל סט המבחן -



### עבור דגימות חיוביות בלבד (ייצוג נמוך) –

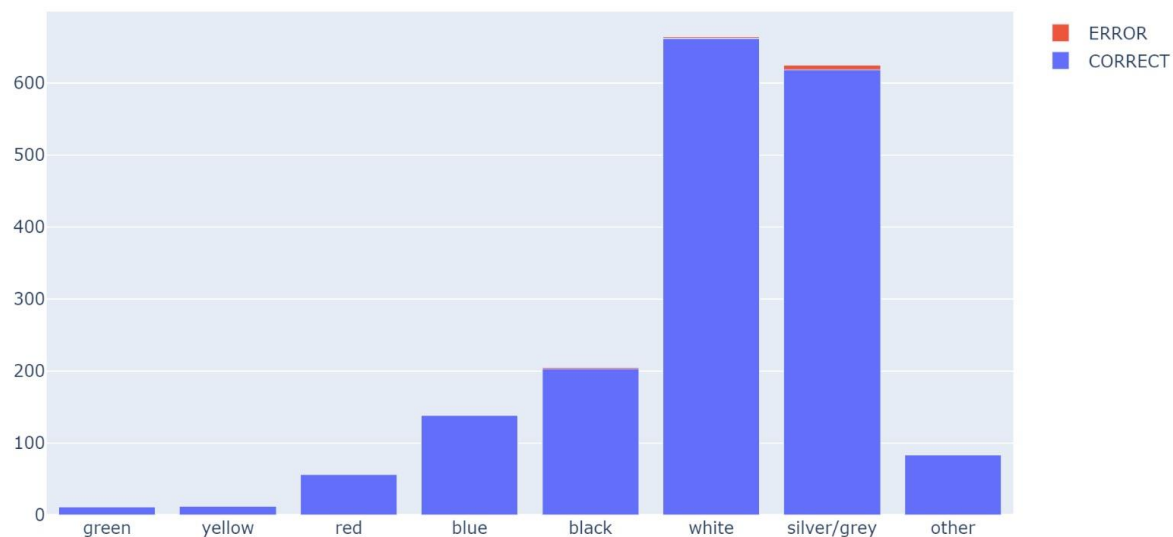


```

Accuracy of the open_cargo_area network on the 1795 test images: 99 %
Accuracy of the open_cargo_area network on the 34 positive test images: 91 %
-----
Accuracy of the wrecked network on the 1795 test images: 99 %
Accuracy of the wrecked network on the 74 positive test images: 96 %
-----
Accuracy of the flatbed network on the 1795 test images: 98 %
Accuracy of the flatbed network on the 7 positive test images: 100 %
-----
Accuracy of the ladder network on the 1795 test images: 98 %
Accuracy of the ladder network on the 0 positive test images: NaN
-----
Accuracy of the enclosed_box network on the 1795 test images: 98 %
Accuracy of the enclosed_box network on the 17 positive test images: 88 %
-----
Accuracy of the soft_shell_box network on the 1795 test images: 98 %
Accuracy of the soft_shell_box network on the 3 positive test images: 67 % (2 of 3!)
Accuracy of the harnessed_to_a_cart network on the 1795 test images: 98 %
Accuracy of the harnessed_to_a_cart network on the 1 positive test images: 100 %
-----
Accuracy of the sunroof network on the 1795 test images: 98 %
Accuracy of the sunroof network on the 163 positive test images: 93 %
-----
Accuracy of the luggage_carrier network on the 1795 test images: 98 %
Accuracy of the luggage_carrier network on the 72 positive test images: 89 %
-----
Accuracy of the enclosed_cab network on the 1795 test images: 98 %
Accuracy of the enclosed_cab network on the 22 positive test images: 77 %
-----
Accuracy of the spare_wheel network on the 1795 test images: 98 %
Accuracy of the spare_wheel network on the 25 positive test images: 84 %

```

## עבור סיווג הצבע:



Accuracy of the green network on the 11 test images: 100 %  
 11 11  
 Accuracy of the yellow network on the 12 test images: 100 %  
 12 12  
 Accuracy of the red network on the 56 test images: 100 %  
 56 56  
 Accuracy of the blue network on the 139 test images: 99 %  
 138 139  
 Accuracy of the black network on the 205 test images: 99 %  
 203 205  
 Accuracy of the white network on the 664 test images: 99 %  
 662 664  
 Accuracy of the silver/grey network on the 625 test images: 98 %  
 618 625  
 Accuracy of the other network on the 83 test images: 100 %  
 83 83



## 6. סיכום ומסקנות

בפרויקט זה התמודדנו עם אתגר ראייה ממוחשבת של מפא"ת, שדרש סיווג מורכב של תמונות רכבים ממבט הרחפן. מטרת הפרויקט הייתה ליצור מסווג תמונות של רכבים למחלקה עיקרית, מחלקה משנית, מאפיינים שונים וצבע.

בסיס הנתונים כלל דגימות לא מאוזנות, ולכן עיקר האתגר היה להתגבר על בעיה זו. דרכי ההתמודדות שלנו עם בעיה זו כללו בעיקר אוגמנטציות סלקטיביות ומשקול שונה של פונקציית המחיר.

בנוסף, עיקר החידוש (בעינינו) בפרויקט זה הוא שימוש בארכיטקטורה היררכית תלויה ע"מ להתמודד עם המורכבות שבסיווג, שכן שימוש בארכיטקטורה זו הביאה לביצועים מעולים ללא שימוש ב Boosting.

המסקנות שלנו מפרויקט זה:

- ניתוח התפלגות המידע ביחס למחלקות הסיווג הוא קריטי עבור בעיות סיווג.
- במקרה של מידע לא מאוזן, אין משמעות לאחוז דיוק עבור מדגם מייצג. צריך לייצר מדגם מאוזן ולבדוק עבורו.
- שימוש בסיווג היררכי תלוי עבור הבעיה המורכבת שקיבלנו, הוביל לתוצאות טובות מאד. **המסקנה העיקרית מפרויקט זה היא שכדאי לחקור את ההשפעה של סיווג היררכי תלוי של תמונות, גם אם מורכבות הבעיה לא נובעת מעודף מאפיינים לסיווג**
- למשל, אם צריך לסווג תמונות של רכבים לסוגים הבאים: פרטי, ג'יפ, משאית ואוטובוס, עם מידע לא מאוזן, נציע לסווג ל- גדול \ קטן ואז לסווג לפי תוצאת הסיווג הראשון.

## רשימת מקורות

- [1] Krizhevsky Alex, Ilya Sutskever and Geoffrey E. Hinton, "ImageNet classification with deep convolutional neural networks", Advances in neural information processing systems, 2012.
- [2] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian sun, "Deep Residual Learning for Image Recognition", 2015.
- [3] Vicente Garcia, Ana I. Marques and J. Salvador Sanchez "Improving Risk Predictions by Preprocessing Imbalanced Credit Data", 2012.
- [4] Kamran Kowsari , Donald E. Brown, Mojtaba Heidarysafa, Kiana Jafari Meimandi , Matthew S. Gerber and Laura E. Barnes, "HDLTex: Hierarchical Deep Learning for Text Classification", 2017.