בס"ד

**מגיש:** מיכאל לוגסי **ת.ז:** 305536575

# דו"ח פרויקט גמר

בתחילה בחנתי את הנתונים הנמצאים בתיקיות, ראיתי מה אופי התמונות מה גודלן, כמובן שביצעתי את הדיוגים בהתאם להוראות והבנתי תוך ראשית את הקושי בזיהוי תמונות מורכבות ואת זה שדאטה ככל הנראה יגיע עם שגיאות – כי אני הרגשתי שבחלק מהמקרים לא הייתי מספיק מדוייק.  
החשש שלי על סמך הדאטה שלי היה בעיקר, שאין מספיק נתונים לאימון ולכן אני עלול ליפול לoverfitting, שהתמונות הן מאוד גדולות וה resize עלול לפגוע באיכות שלהן, וכמובן שאין לי מושג מהו ה LR המתאים למודל ולכן עליי לגשש ולהבין את הכיוון.  
  
בפרויקט בחרתי לעבוד עם **YOLO** (גירסת ה large שלו) כי הוא מציע שילוב מצוין של דיוק ומהירות, מה שהופך אותו לאידיאלי למשימות זיהוי אובייקטים בזמן אמת.  
  
התחלתי בהרצות ראשוניות (שאינן מתועדות) על מנת לראות אם הרשת שבניתי נבנתה כמו שצריך ושאני בכיוון הנכון ולאחר מכן התחלתי לבחון את הרשת על הנתונים בהתאם להוראות.  
  
בכל האימונים שהרצתי המודל עבד עם אופטימיזר של Adam עם מנגנון של Early Stopping ועם ערך דיפולטיבי של המומנטום.  
בחקר המודל בחרתי לשחק עם מספר פרמטרים שלפי הבנתי הם הכי משפיעים על איכות המודל ויוכלו להוביל לשיפורים במודל ולהניב תוצאות טובות יותר:

* Learning Rate
* Pretrained/Not Pretrained
* With/Without Augmentation
* Image Size

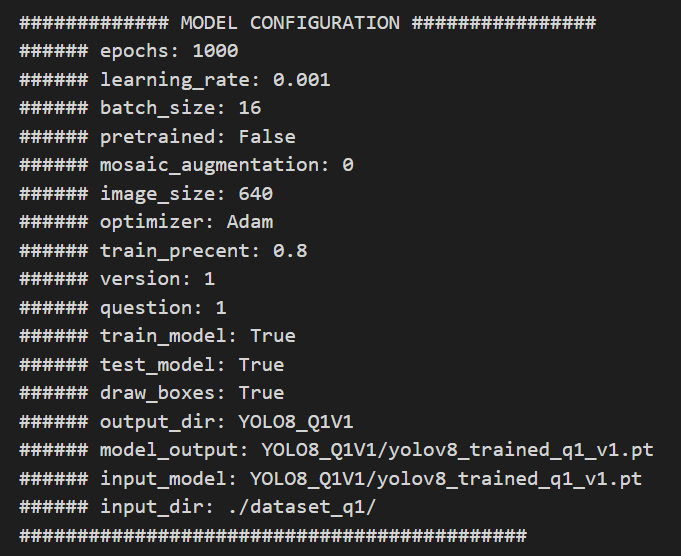
אני מודע לכך שיתכן שהגבלה למספר הפרמטרים אלו תוביל אותי לתקרת זכוכית מסויימת – אך מאחר ויש הרבה מאוד אפשרויות לשחק איתן ומאחר ולא רציתי להתפזר בחרתי מספר פרמטרים שהאמנתי שיוכלו להוביל אותי לתוצאות טובות ואיתן הלכתי.  
בנוסף – לאור אילוצים של זיכרון של ה GPU נאלצתי לפעמים לשנות את ה batch\_size – לא ציינתי זאת בתהליך מאחר ויחסית לזה השפעה שולית ביחס לפרמטרים האחרים שאיתם רציתי להתעסק.  
  
לטובת קבלת אומדן עבור איכות המודל תייגתי חלק מהתמונות של הטסט - כ 30 תמונות (הרבה – קרה לי בטעות) מהדאטה של שאלה 1 וכ 10 עבור הדאטה של שאלה 2.

בנוסף לנתונים המתקבלים מהרשת המשמשים לטובת הבנת איכות המודל הוספתי עוד כמה אומדנים שעזרו לי גם כן להבין בפועל איך המודל עובד עבוד התמונות של הטסט שתייגתי.  
את המודל בחנתי כמובן על סמך תוצאת ה IoU עבור כל תיבה, כאשר למעשה ביצעתי התאמה בין התיבה שנחזתה לבין התיבות שהיו מתוייגות בתמונה בצורה גרידית, כלומר נעשתה התאמה בין תיבה מתוייגת לתיבה שהמודל חזה עם ה IoU הגבוה ביותר.

אלו האומדנים שהוספתי:

* **Average Matches IoU –** ממוצע של ה IoU של כל אחת מהתיבות אך רק עבור תיבות שנמצא להן התאמה בין הצפי לחוזי.
* **Average Matches&Unmatched IoU -** ממוצע של ה IoU של כל אחת מהתיבות כולל התיבות שלא נמצאה להן התאמה בין הצפי לחוזי.
* **Average Unmatches Predictions –** מספר התיבות היתרות שהמודל חזה בממוצע.
* **Average Unmatches Actual –** מספר התיבות החסרות שהמודל לא חזה ולא נמצאה להן התאמה.

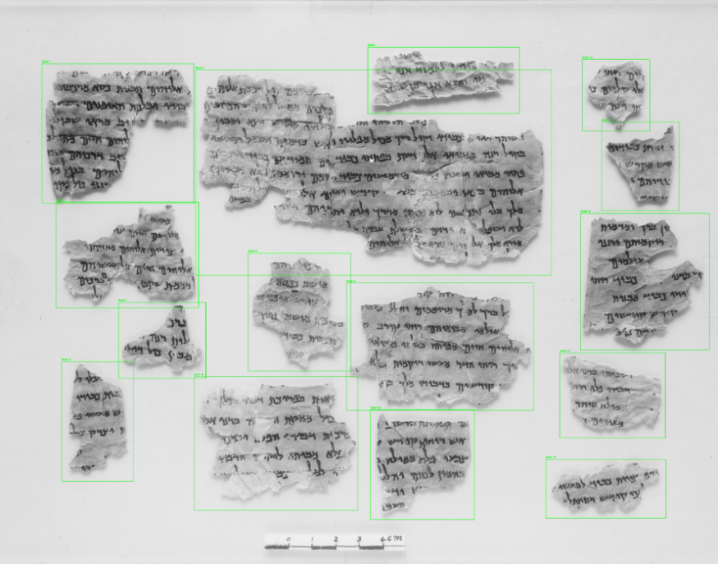
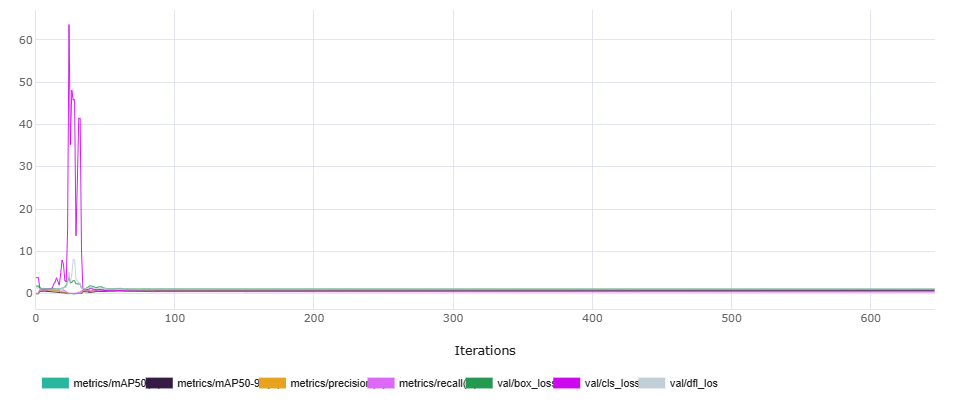
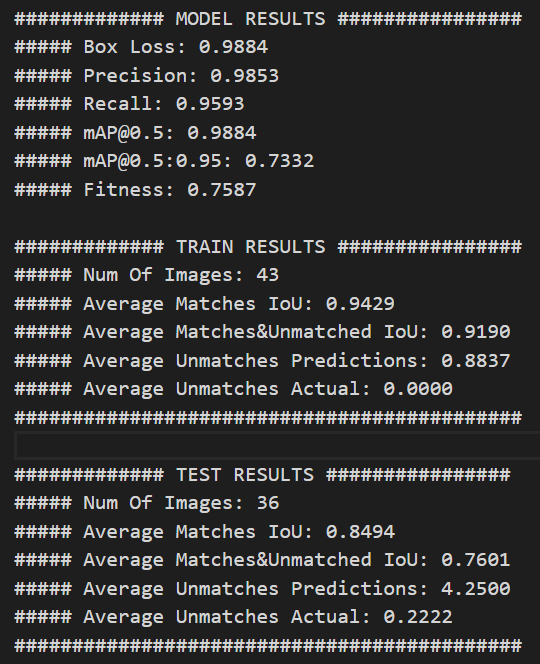
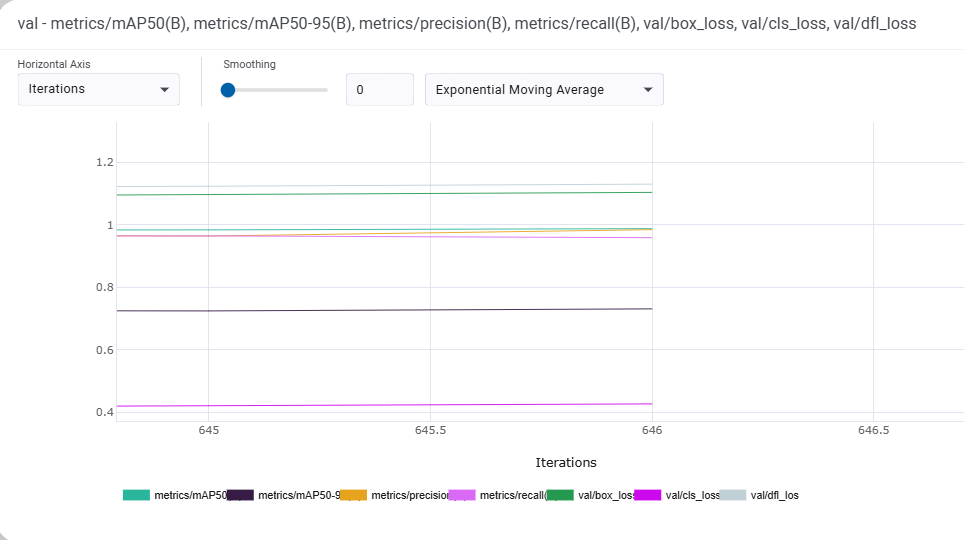
בנוסף יצרתי תמונות עם התיבות שנחזו ע"י המודל על מנת לבחון בעיניים את איכות המודל.  
  
**\*\*\* בסוף הדו"ח מצורך הסבר כיצד יש להריץ את הקבצים על מנת לאמן ולבחון את המודל \*\*\***  
  
**בעמודים הבאים אתאר את תהליך האימון בכל שאלה ובכל גירסת דאטה.**מאחר ורציתי להגיע לתוצאות טובות הרצתי הרבה רשתות עם פרמטרים שונים ומשולבים במקבלים – כאן אציג רק חלק מהתהליך וכמובן אסביר את הדרך שלי והתהליך של הבדיקה.

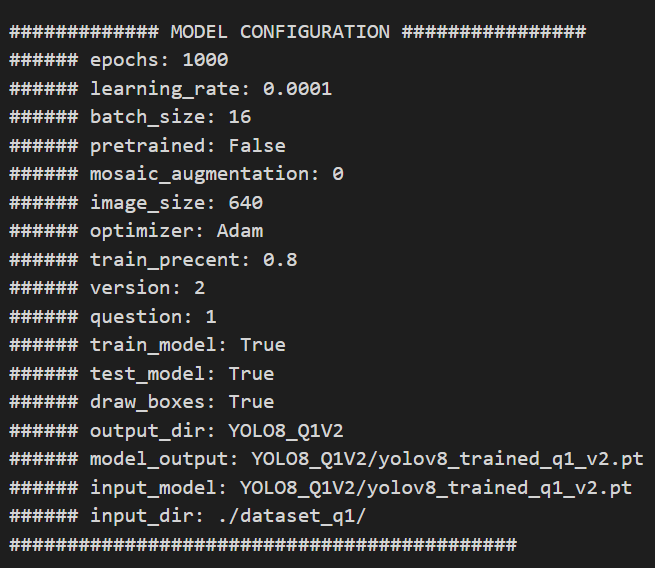
***שאלה 1:***

***מודל ראשוני - ערכים שרירותיים***

התחלתי עם מודל עם ערכים דיפולטיבים בשביל הרגשה ראשנית של הרשת, לראות תוצאות ראשוניות שלה ולראות כיצד היא משתפרת בהמשך.

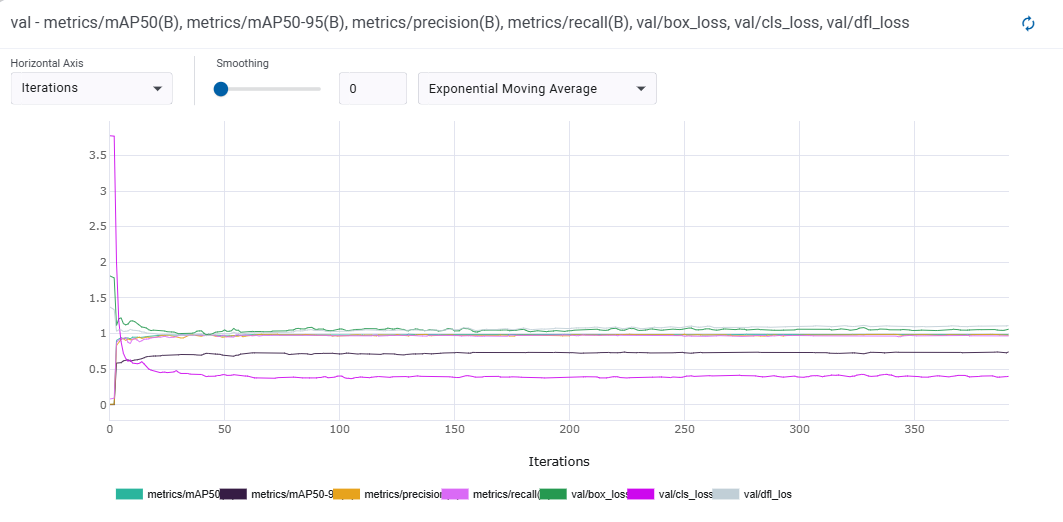
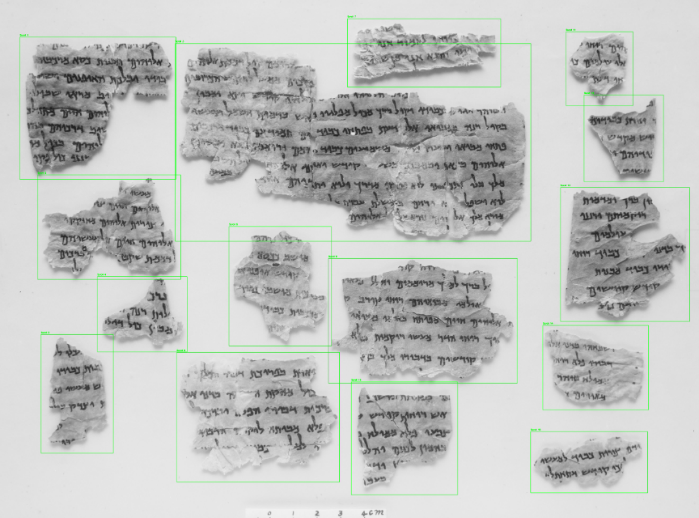
* Learning Rate – 0.001 – מספר שרירותי שבחרתי
* Not Pretrained
* Without Augmentation
* Image Size – התמונות עברו resize ל 640X640

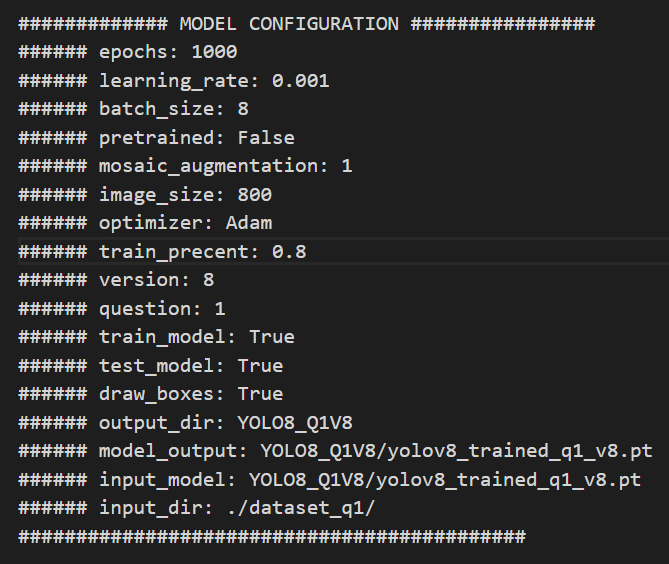
************המודל רץ כ 600 איטרציות והניב תוצאות יחסית טובות להתחלה. המודל היה יחסית יציב לאורך האימון והתקדם עד שנתקע עד שנעצר.  
הציונים יחסית טובים, אין אוברפיטינג, אמנם יחסית יש הרבה חיזוי לא נכון של מסגרות אך על סמך תוצאות החוזי בעיניים ניתן לראות המודל עובד מצויין ושהציונים לא בדיוק משקפים בעיה אמיתית במודל.  
  
דוגמא לתמונה שהוא חזה:

***מודל שני – בחינת שיפור ה LR***

אחרי שהיה לי מודל ראשוני רציתי לבחון את השפעת ה LR על אימון המודל ותוצאותיו. האינטואציה אמרה לי ש LR נמוך יותר אולי יגרום לרשת ללמוד באופן מדוייק יותר (גם מאחר שהמודל הקודם הופסק כמעט באמצע הדרך).

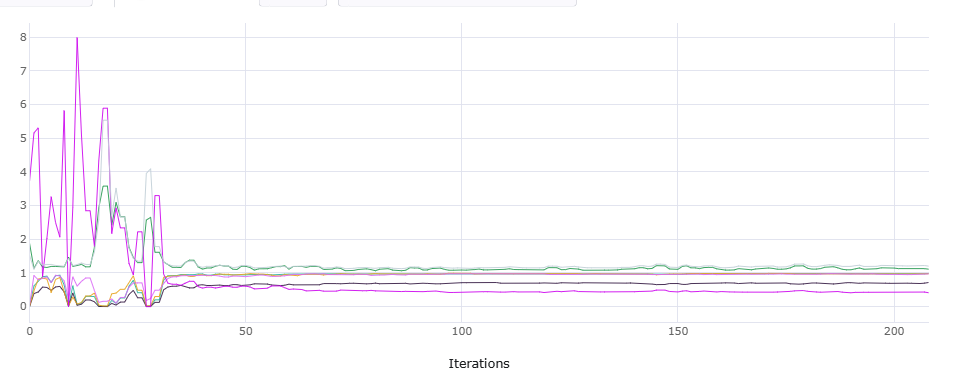
* Learning Rate – 0.0001 – הקטתי מפעם שעברה
* Not Pretrained – ללא שינוי
* Without Augmentation - ללא שינוי
* Image Size – התמונות עברו resize ל 640X640 – ללא שינוי

***A screenshot of a computer

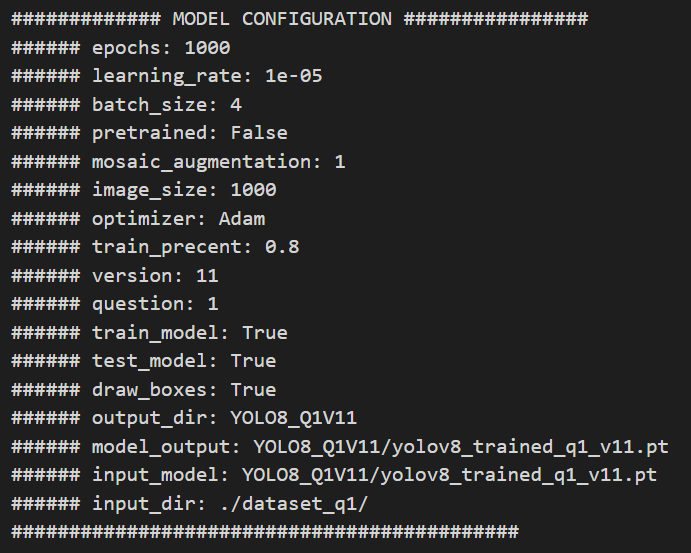
AI-generated content may be incorrect.***המודל רץ כ 400 איטרציות, כנראה הגיע למינימום מקומי ולא הצליח לצאת ממנו מהר יותר ביחס למודל הקודם.  
 המודל הניב תוצאות משופרות מאוד (3% ב IOU הכללי) וגם מקטין כמעט בחצי (!) את החוזי העודף – מה שכמובן עוזר להעלאת אחוז הדיוק.  
שוב ניכר שאין אובר פיטיג ושוב על סמך בחינה של התוצאה בעיניים שוב ניכר שהמודל עדיין עובד מאוד יפה.  
דוגמא לתמונה שהוא חזה:***מודל שלישי – בחינת שיפור האוגמנטציה + הגדלת התמונה***

לאחר שזיהיתי שהקטנת התמונה משפיעה לרעה על המודל (לא ציינתי כאן אך יהיה ניתן לראות בסיכום של החקירה) בחרתי לחזור ל LR הראשוני והפעם להוסיף אוגמנטציה והגדלת התמונה מתוך תקווה לראות אולי הגדלת הרזולוציה והוספת תמונות לאימון ישפרו את המודל.

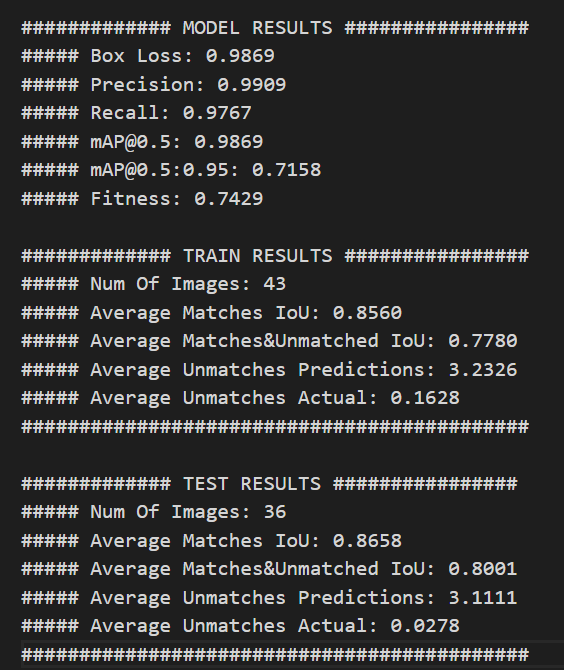
* Learning Rate – 0.001 – חזרתי לראשוני
* Not Pretrained – ללא שינוי
* With Augmentation - רוצה לבחון את השפעת האוגמנטציה
* Image Size – התמונות עברו resize ל 800X800

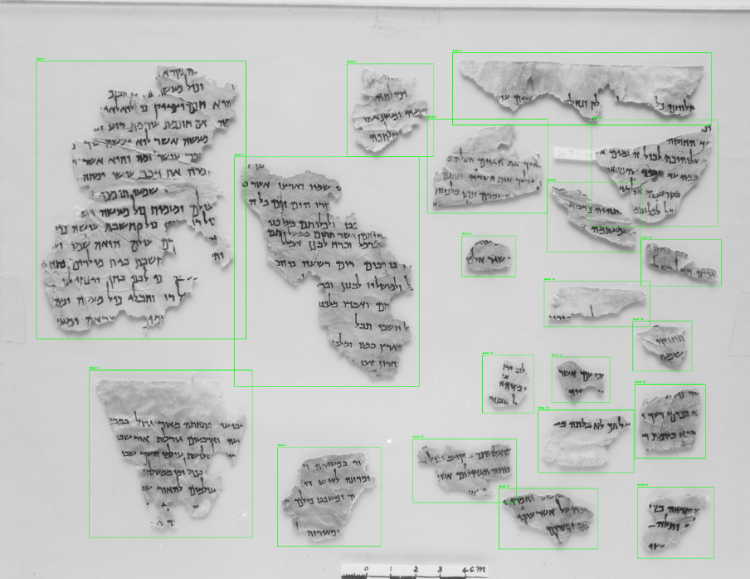
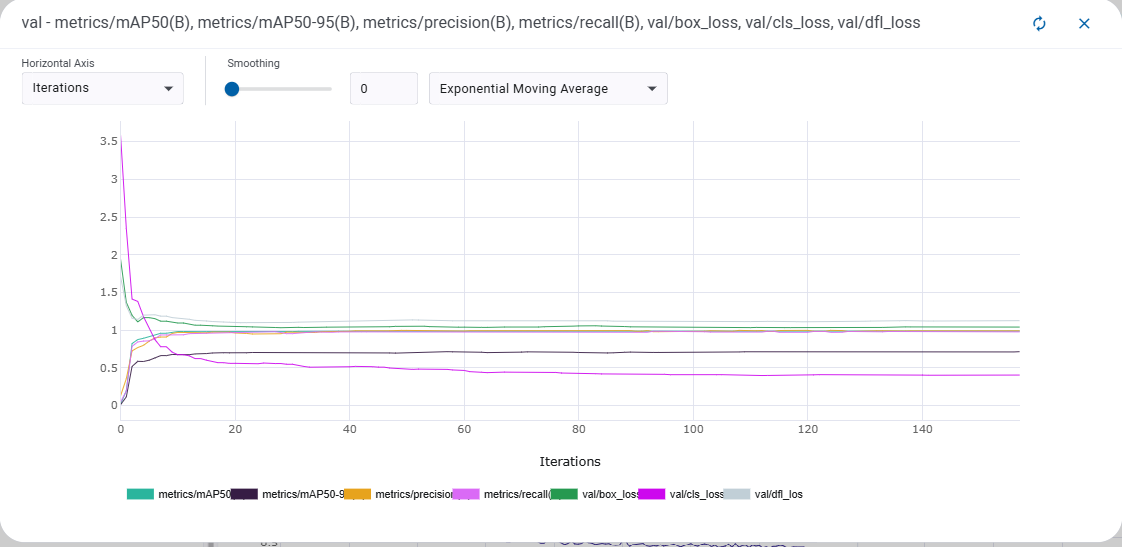
***A screenshot of a white board with many pieces of paper

AI-generated content may be incorrect.A screenshot of a computer screen

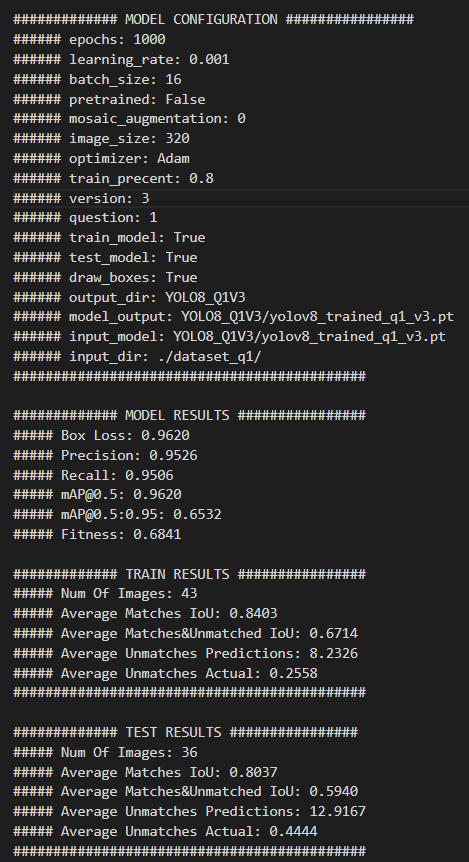
AI-generated content may be incorrect.***המודל רץ כ 200 איטרציות והניב תוצאות פחות טובות מהמודל הקודם, אמנם יש שיפור בדיוק ביחס למודל הראשון (אחוז +-) אך יצר יותר חיזויים עודפים – מה שמעיד שבמה שחזה נכון עשה זאת טוב יותר אך עדיין יצר יותר חיזויים עודפים.  
אך שוב ביחס לבדיקת התוצאה בעיניים ניכר שהמודל עובד מעולה.  
דוגמא לתמונה שהוא חזה:***מודל רביעי – המודל עם התוצאות הטובות ביותר***

מאחר ואיני רוצה להאריך ולהלאות אתמצת את המשך החקר שנעשה והמסקנות שהובילו אותי למודל הבא.  
בחנתי כל מיני היבטים של שינויים בפרמטרים שבחרתי לשחק איתם. המגמה שהבנתי היא שככל התמונה גדולה יותר החוזי מדוייק יותר, אוגמטציה עוזרת למודל שה LR נמוך טוב לאימון.  
לכן במודל זה יישמתי את הדברים.

* Learning Rate – 0.00001 – מאוד קטן
* Not Pretrained
* With Augmentation
* Image Size – התמונות עברו resize ל 1024X1024 –

******המודל רץ כ 150 איטרציות והניב שוב תוצאות יחסית זהות אך עם שיפור קל ביחס למודל השני שהרצתי (שהיה הטוב ביותר עד כה). המודל שוב היה יחסית יציב לאורך האימון והתקדם עד שנתקע עד שנעצר. באופן מפתיע המודל מניב תוצאות טובות יותר עבור הטסט מאשר האימון עצמו מה שמעיד אולי על מודל שיוכל לעבוד גם עבור תמונות חדשות במידה טובה.  
וכמובן – על סמך בדיקה בעיניים – המודל שוב עובד מעולה!  
דוגמא לתמונה שהוא חזה:  
***סיכום חקר ואימון מודל שאלה ראשונה***

כבר באימונים הראשונים היה ניכר הרשת יודעת לעבוד עם הדאטה הנתון ושגם עבור לכאורה דאטה מועט מצליחה להניב תוצאות בהחלט יפות ומרשימות.  
בעיקר התרשמתי מהתוצאות בעיניים – מרגיש לי שאומנם בין מודל למודל הצלחתי לשפר פה אחוז שם אחוז אבל בפועל מאחר והתיוג עצמו לא מאה אחוז מדוייק סביב המגילה הסטייה של ה IOU לא כ"כ משמעותית (כי עשיתי ממוצע באומדן שלי – אילו הייתי עושה ממוצע ל IOU שהוא מעל אחוז סביר מסויים כמו למשל 0.75 כנראה שהמודל היה מגיע לתוצאות כמעט מושלמות (על סמך הממוצע שקיבלתי כמובן).  
אני בחרתי במודל האחרון כמודל הטוב ביותר עבורי – מאחר שהוא היה בין הטובים ביותר וגם הראה שהוא עובד כמעט אותו דבר גם לאימון וגם לטסט – אז אולי הוא יותר גלובלי ויכול להתאים לעוד מקרים.  
דבר אחד שממש הפריע לי ולא הבנתי איך בדיוק לתפתור – היה ניכר שהאימון מגיע למיצוי יחסית מהר – בשום אימון לא הגעתי לאימון מלא של 1000 איפוקים – יתכן שזה בגלל שהדאטה קטן ולכן המודל נעצר בשביל למנוע אוברפיטינג – אך שוב לא הצלחתי לפתור את הבעיה שאולי היתה יכולה לשפר לי את איכות המודל.  
מצ"ב עוד כמה קונפיגורציות ומודלים שהרצתי כחלק משלבי החקירה

A screenshot of a computer

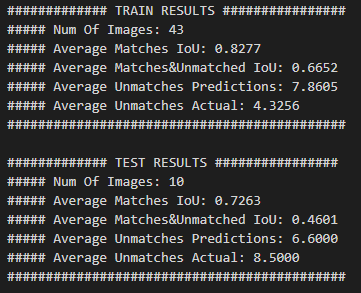
AI-generated content may be incorrect.

***שאלה 2:***

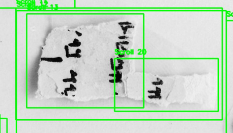
***מודל ראשוני - ערכים שרירותיים***

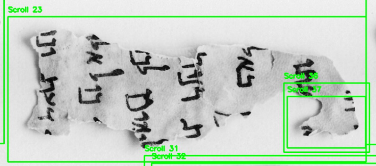
שוב התחלתי עם מודל עם ערכים דיפולטיבים בשביל הרגשה ראשנית של הרשת, לראות תוצאות ראשוניות שלה ולראות כיצד היא משתפרת בהמשך. בחרתי באותם ערכים כמו בשאלה הראשונה.

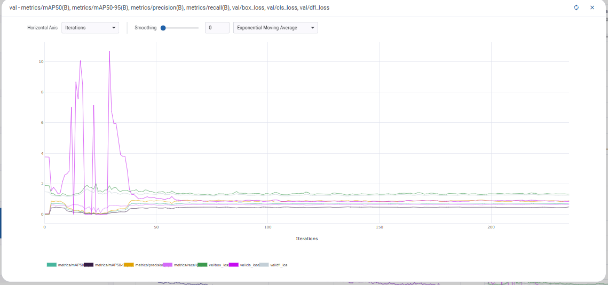
* Learning Rate – 0.001 – מספר שרירותי שבחרתי
* Not Pretrained
* Without Augmentation
* Image Size – התמונות עברו resize ל 640X640

המודל רץ כ 250 איטרציות והניב תוצאות לא טובות – לא עבור האימון ולא עבור הטסט. מאחר והמודל נעצר יחסית בתחילת האימון נראה שהאימון נעצר מאחר והמודל לא הצליח להשתפר ולא בשביל למנוע אוברפיטינג.  
נכיר גם שכמות הפספוסים (כ 9 חוסרים וכ 7 עודפים) של המודל גדולים – מה שכמובן משפיע מאוד על הציון – ניתן לראות שללא הפיספוסים המודל יחסית עם ציון סביר של 72% - סה"כ ציון סביר בהתחשב במורכבות הבעיה.   
בהסתכלות על התוצאות בעיניים – נראה שסה"כ המודל מייצר חיזוי סביר – ניכר שיש פיספוסים אך גם הצלחות.  
דוגמא לתמונה שהוא חזה – כולל הצלחה ושגיאות:

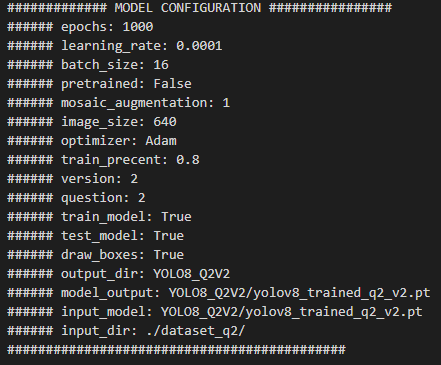
A screenshot of a white board with green squares

AI-generated content may be incorrect.זיהוי יפה:

זיהוי שגוי:

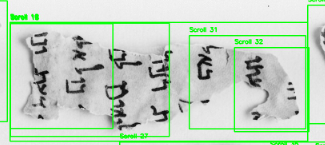
A white piece of paper with green rectangles

AI-generated content may be incorrect.פספוס:

***מודל שני - בדיקת השפעת ה LR ואוגמנטציה***

מאחר והיה ניכר בשאלה הראשונה שהקטנת ה LR והוספת אוגמנטציה הובילה לשיפור במודל – בחרתי הפעם ללכת בכיוון הזה הזה גם בפעם.

* Learning Rate – 0.0001 – הקטנתי במעט
* Not Pretrained
* With Augmentation
* Image Size – התמונות עברו resize ל 640X640

A screenshot of a computer

AI-generated content may be incorrect.המודל רץ פחות איטרציות כ 160 סה"כ ולהפעתי הניב תוצאות פחות טובות.  
בדיוק המודל פחות או יותר זהה לקודם (קצת פחות טוב אמנם) – אך לא שיפר ואף הרע את הטעויות של המודל, מספר החוזי העודף גדל כמעט בפי 2, אך שיפר ממש במעט את המסגרות שהמודל פיספס.  
שוב הסתכלות בעיניים על התוצאות – הן נראות סבירות – אך שוב ניכר שהמודל לא מדוייק.  
בחרתי לראות את הפער ביחס לאותה תמונה ונראה שהמודל לא כ"כ השתפר ביחס לקודם.  
החוזי השגוי נהיה גדול יותר

**A screenshot of a white board with torn paper

AI-generated content may be incorrect.**A piece of paper with writing on it

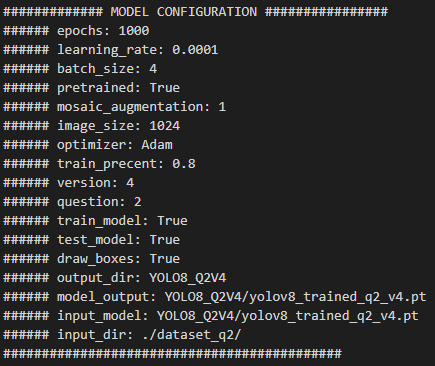
AI-generated content may be incorrect.הפעם פיספס את ההצלחה הקודמת:

A screenshot of a computer screen

AI-generated content may be incorrect.ושוב פיספס את המגילה הזו:

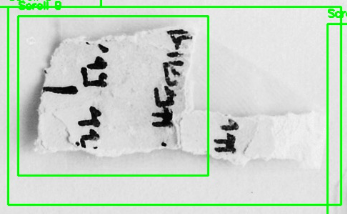
**A screen shot of a graph

AI-generated content may be incorrect.**

***מודל שלישי - שימוש במודל מאומן והגדלת התמונה עם אוגמנטציה***

בחרתי לנסות שוב את הגדלת התמונה, מתוך אמונה שככול שיהיה יותר פרטים למודל הוא אולי יהיה טוב יותר. וגם בחרתי להשתמש במודל מאומן – רציתי לבחון אולי מודל מאומן יוכל אולי להקפיץ את האיכות במודל הזה.

* Learning Rate – 0.0001
* Pretrained
* With Augmentation
* Image Size – התמונות עברו resize ל 1024X1024

A screenshot of a computer

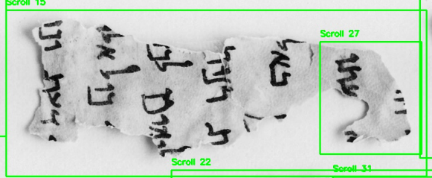
AI-generated content may be incorrect.המודל רץ שוב כ 150 איטרציות – הציג פחות או יותר תוצאות כמו המודלים הקודמים – אך עדיין פחות טובים מהמודל הראשוני מבחינת הפספוסים והחיזוי העודף של המודל.  
בבחינת התוצאות בעיניים – שוב הכל נראה סביר אך עדיין רואים את הטעויות.  
בחרתי שוב לבחון ביחס לאותה תמונה כמו קודם – נראה שביחס למודל הראשון יש דברים שהוא השתפר אך כאלה שהתגרע.  
ההצלחה של המודל הראשון כעת הצלחה חלקית:

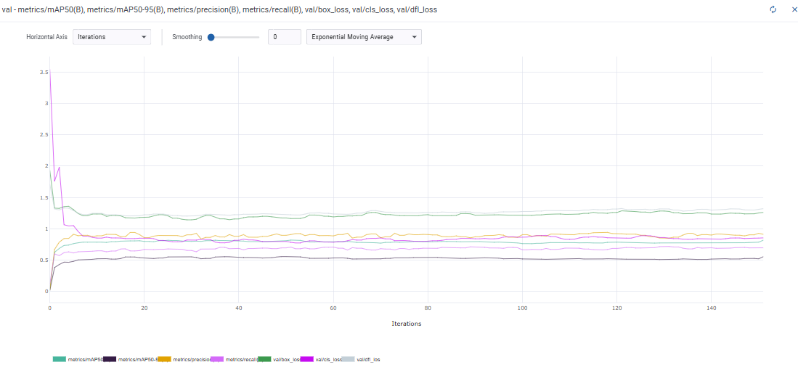
A screenshot of a computer screen

AI-generated content may be incorrect.

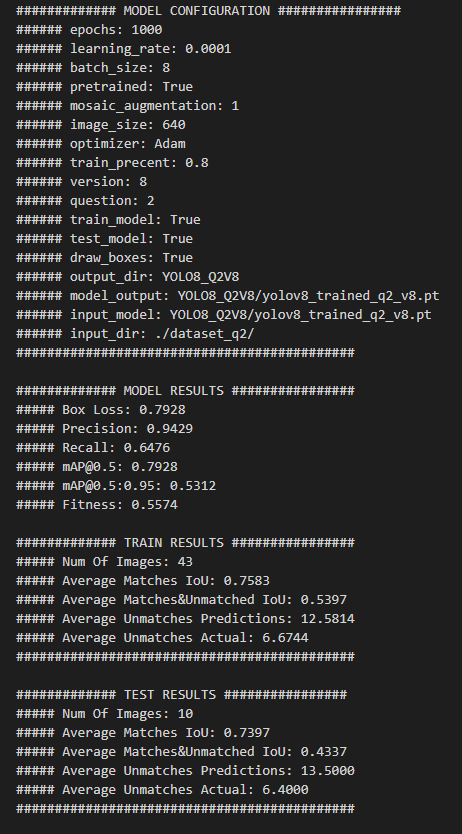
A close-up of a piece of paper

AI-generated content may be incorrect.הפעם הפספוס הפך להצלחה:

השגיאה נותרה שגיאה אך הפעם פחותה:

****

***סיכום חקר ואימון מודל שאלה שנייה***

**A screenshot of a computer program

AI-generated content may be incorrect.**כבר באימונים הראשונים היה ניכר הרשת מתקשה לעבוד עם הדאטה הנתון ושבניגוד לתוצאות של השאלה הראשונה התוצאות נהיו גרועות בצורה מאוד משמעותית וישנם המון פספוסים.  
הפעם גם בהתרשמות מהעיינים היה ניכר שהמודל לא כ"כ מצליח להתמודד עם האתגר – אם כי הוא לא השתבש אלא פשוט לא הבין בדיוק איך לפתור אותה – נראה שהוא כן שומר על זיהוי המגילות כמו שהיה בשאלה הקודמת – אך את תתי המגילות מאוד קשה לו להבחין.  
ציינתי בדו"ח רק שלוש מודלים שבחנתי אך ניסיתי הרבה פרומטציות עם המודל אך שום דבר לא הוביל לשיפור משמעותי – בחרתי את היותר מעניינים בעייני – מצרף פה כמה פרינטים מהריצות האחרות.  
אני בחרתי במודל הראשון כמודל הטוב ביותר – מאחר שהוא באמת הכי פחות יצר שגיאות.  
חשוב לי לציין, מאחר ואני תייגתי בעצמי את התמונות אני יכול להגיד שאני ממש מבין את המודל 😊  
היה לי מאוד קשה בעצמי לזהות אם באמת יש חיבור בין שתי מגילות או שזה סתם כפל או משהו אחר שמבלבל אותי, אני אפילו חושש שהתיוג שלי שיתכן מאוד שהיה לא מוצלח ומבלבל והוא זה שגרם למודל לא לעבוד טוב – כך שאני קצת סלחן עם המודל והתוצאות שלו – אני באמת חושב שהבעיה כאן היא בדאטה עצמו ולאו דווקא במודל עצמו, או הדאטה לא מתוייג בצורה נכונה או שאין מספיק תמונות לאימון בשביל לפתור את הבעיה הזו.  
עוד בעיה – שאני חושב שזו בעיה – שכמו בשאלה הראשונה עדיין נותרה – במודלים נעצרים בשלבים מאוד מוקדמים ואין אימון ארוך וממושך שאולי היה יכול באמת לשפר את המודל.  
  
מצ"ב עוד כמה קונפיגורציות ומודלים שהרצתי כחלק משלבי החקירה

**A computer screen shot of a program

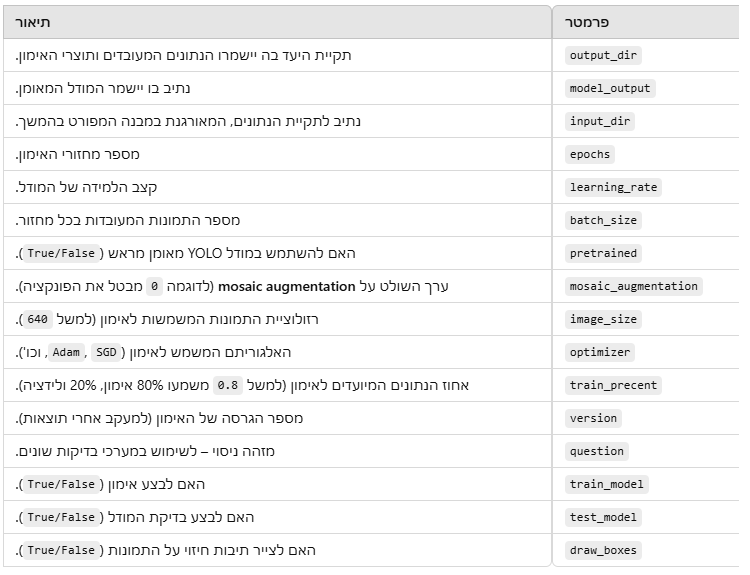
AI-generated content may be incorrect.**

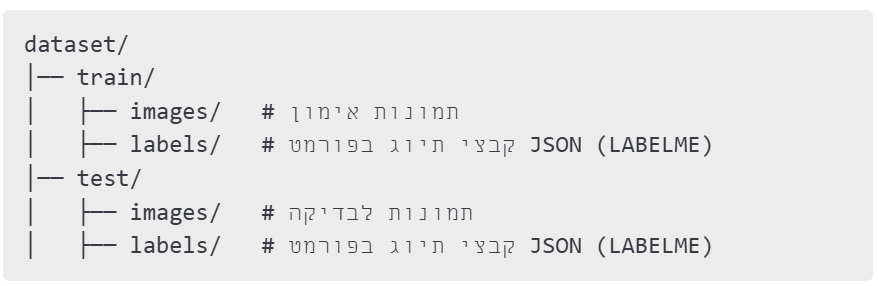
**מדריך לשימוש בפרויקט YOLOv8**

**סקירה כללית** פרויקט זה מכיל סקריפטים לאימון והערכת מודל YOLOv8 לזיהוי אובייקטים. כל אחד מהסקריפטים יכול לרוץ באופן עצמאי, אך יש להגדיר את הקונפיגורציה המתאימה לפני כן.

**קונפיגורציה** הגדרות הפרויקט מאוחסנות במילון config, אשר קובע את ההתנהגות של האימון, ההערכה והכנת הדאטה.

**פרמטרים עיקריים בקונפיגורציה** לפני הרצת הסקריפטים, יש לוודא שהפרמטרים הבאים מוגדרים כראוי:

****

**מבנה תיקיית הנתונים** הנתונים צריכים להיות מאורגנים במבנה הבא:  
  


**הרצת חיזוי באמצעות הפונקציות process\_detailed\_bounding\_boxes ו־predict\_process\_bounding\_boxes** :

על מנת להפעיל פונקציות אלו, יש לוודא שבתיקייה של הסקריפט קיים מודל מאומן בשם:

* **best\_model\_q1.pt**
* **best\_model\_q2.pt**

קישור למודלים שלי:  
https://drive.google.com/drive/folders/1ufNu-\_b6C4wl7\_5X6vfHshuGrprwAkrj?usp=sharing