דו"ח מסכם – פרויקט גמר

Music Notation Software

### מגישים:

חביב עביד 311383780 habib.obeid@gmail.com

רולא עביד 315866905 rolao\_96@hotmail.com



אוניברסיטת חיפה, החוג למדעי המחשב

ספטמבר 2021

###### תוכן עניינים

הגדרת הבעיה ................................................................................... 3

הגדרת הפתרון .................................................................................. 4

מימוש ............................................................................................... 5

הצעות להמשך הפרויקט .................................................................... 13

מה למדנו? ........................................................................................ 14

צוואר-הבקבוק ................................................................................... 15

###### הגדרת הבעיה

לא לכל האנשים יכולת השמיעה היחסית שלהם מפותחת, כלומר הם לא מסוגלים לנגן משמיעת המנגינה או השיר מבלי להתבסס על תווים כתובים.

בנוסף מוזיקאים רבים מתקשים לרשום תווים בזמן שהם מלחינים מוזיקה. קשה ומבלבל להפסיק לנגן כדי לרשום את התווים ואז להמשיך לנגן.

לכן הפרויקט שלנו הוא רעיון לפתור הבעיות שנזכרו למעלה, אז המטרה שלו היא להוציא ולהדפיס תווים מוזיקאליים מתוך וידאו של ידיים מנגנים על הפסנתר.



###### הגדרת הפתרון

בפתרון שלנו בנינו תוכנה כך שמקבלת כקלט וידאו לקלידים של הפסנתר מלמעלה שצולם במצלמה של "איפון 12 מיני" (כמו שמוצג בתמונה). התוכנה מעבדת את הוידאו ומכינה אותו כך שיהיה מתאים לשלבי האיתור.

איתור ראשון יהיה לקלידים עצמם כך שתזהה התוכנה כל קליד לבד, איתור שני הוא לידיים ולאצבעות כך שתזהה התוכנה המיקום שלהם בכל מסגרת בווידאו.

שלב אחרון הוא להחליט אם האצבע לחץ או לא ועל איזה אחד מהקלידים.

ואז נדפיס הפלט כך שיהיה התווים שהוקלדו.

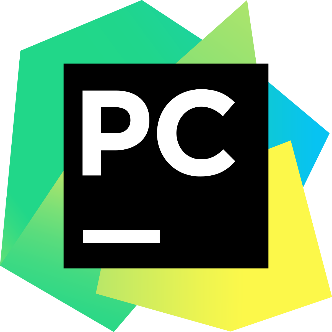


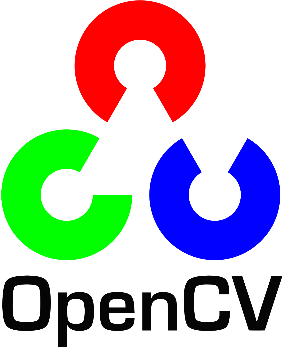


###### מימוש

כתבנו את הקוד בשפת Python גרסה 3.9 תוך שימוש PyCharm גרסה 2021.2.1 .

השתמשנו בספריית OpenCV גרסה 4.5.1 - שנועדה לעזור לפתח יישומים של ראייה ממוחשבת.

השתמשנו גם ב Media Pipe framework שהיא אחת הספריות המשותפות והשימושיות ביותר לעיבוד מדיה בתוך Google.

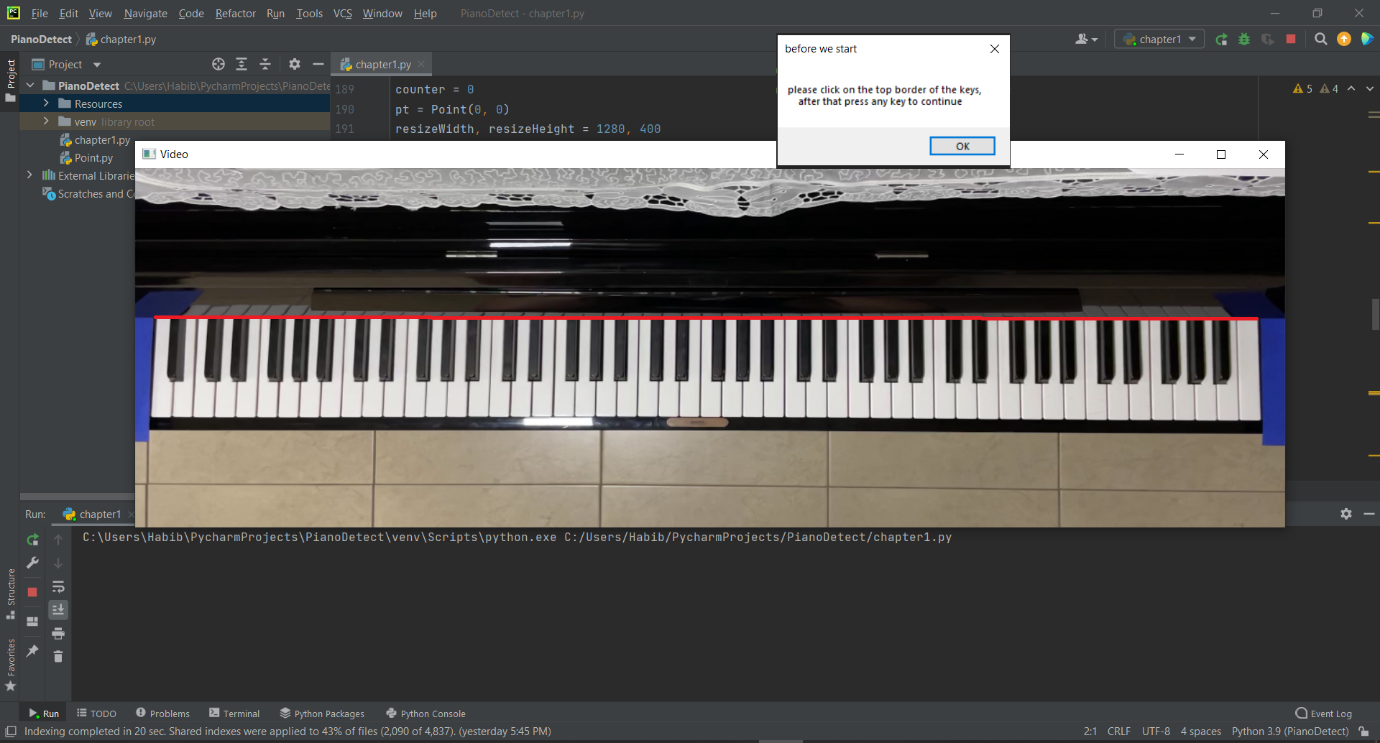


נתאר את השלבים העיקריים במימוש:

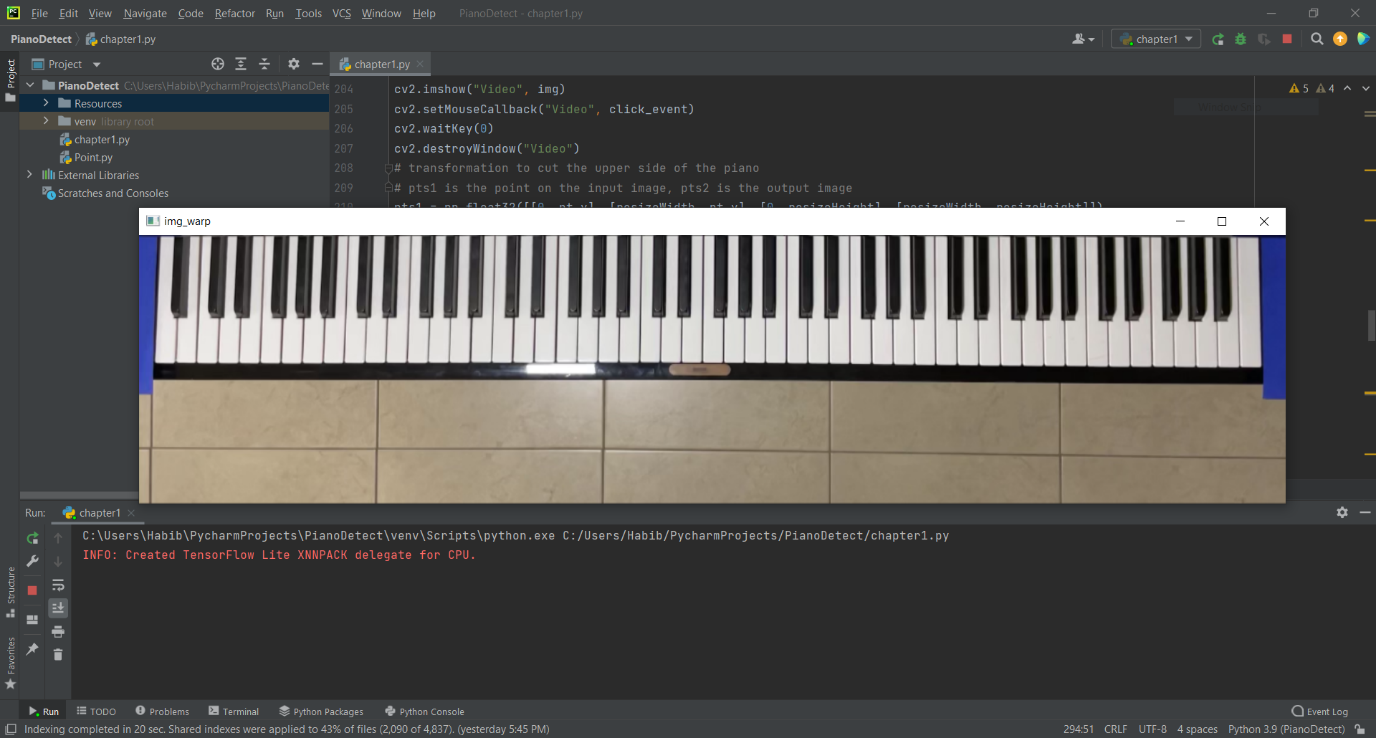
1. מבקשים לבחור וידיו מתוך המחשב



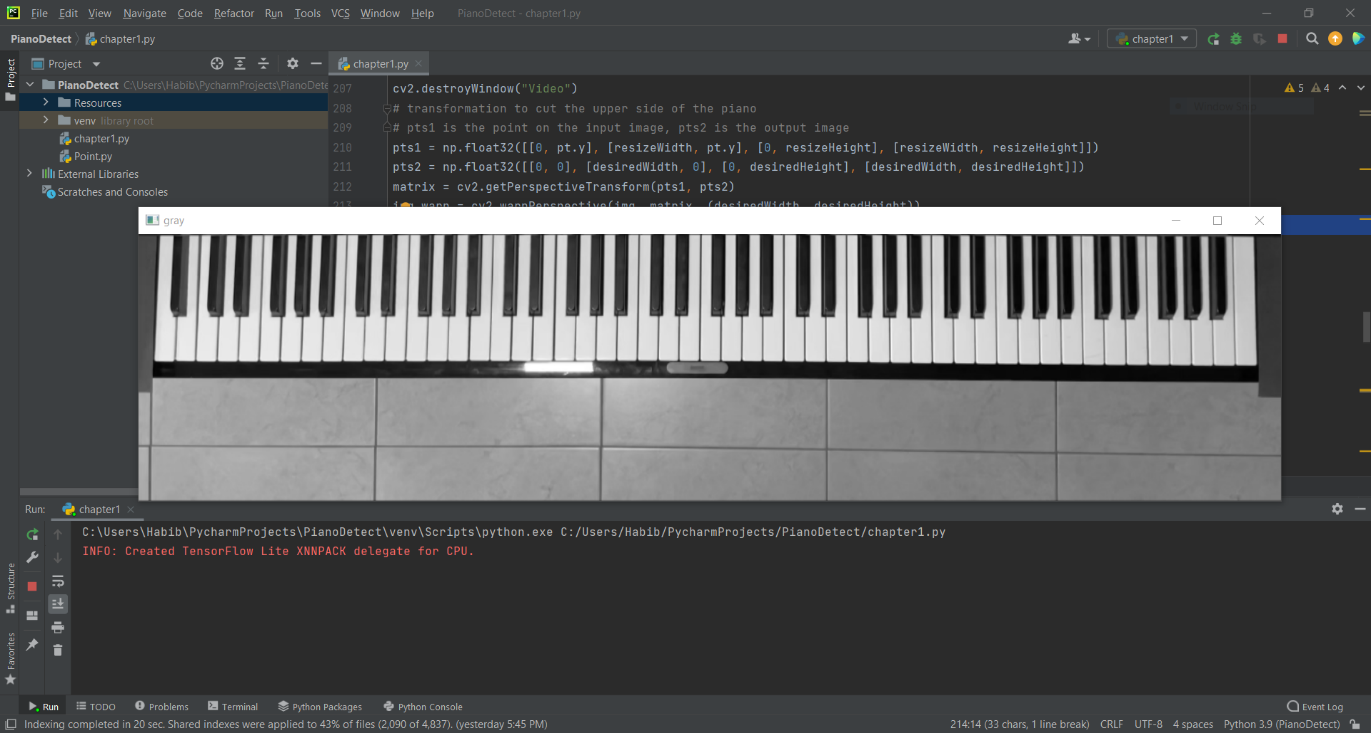
1. מציגים את המסגרת הראשונה מהווידאו (כך שמוצג רק את הפסנתר ללא הידיים) ומבקשים ללחוץ על הקו העלון של קלידי הפסנתר כדי שנחתוך את כל מה שנמצא מעל הקלידים.



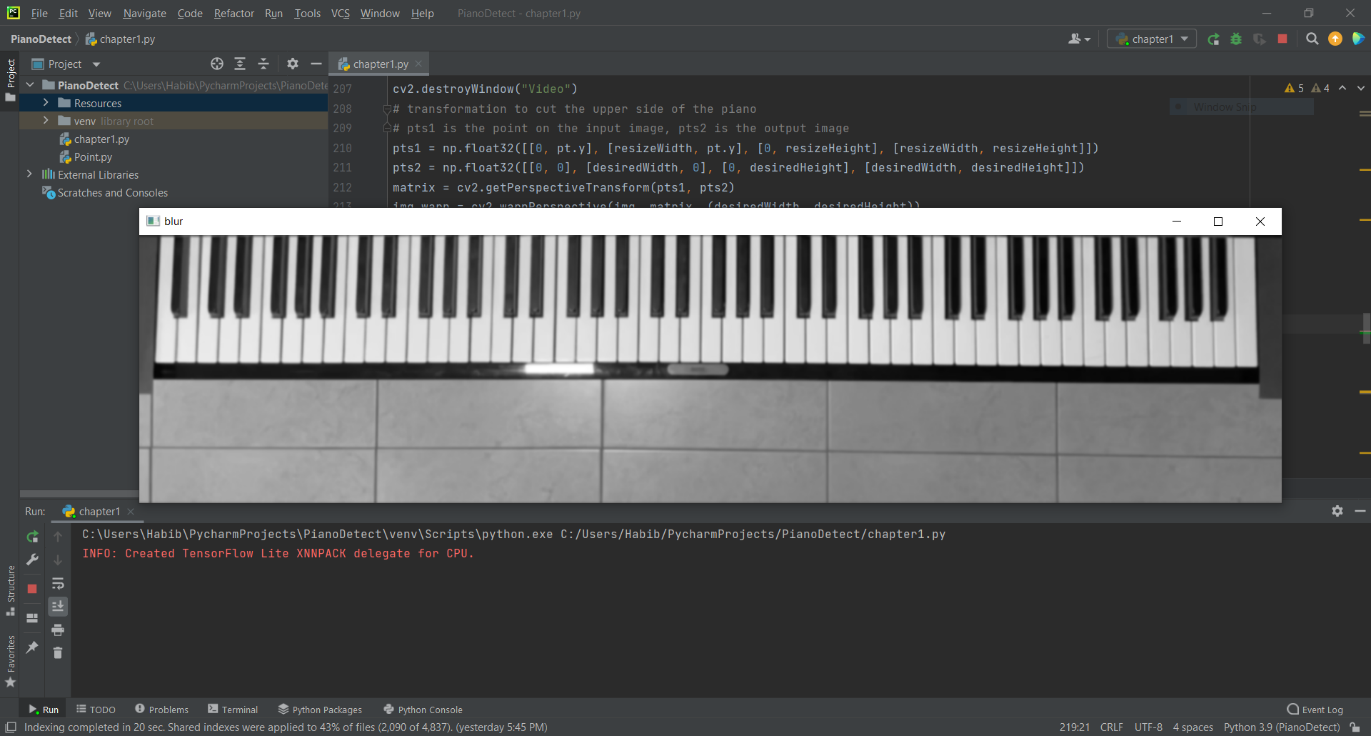
1. עושים warp לתמונה ע"י שימוש ב cv2.warpPerspective.



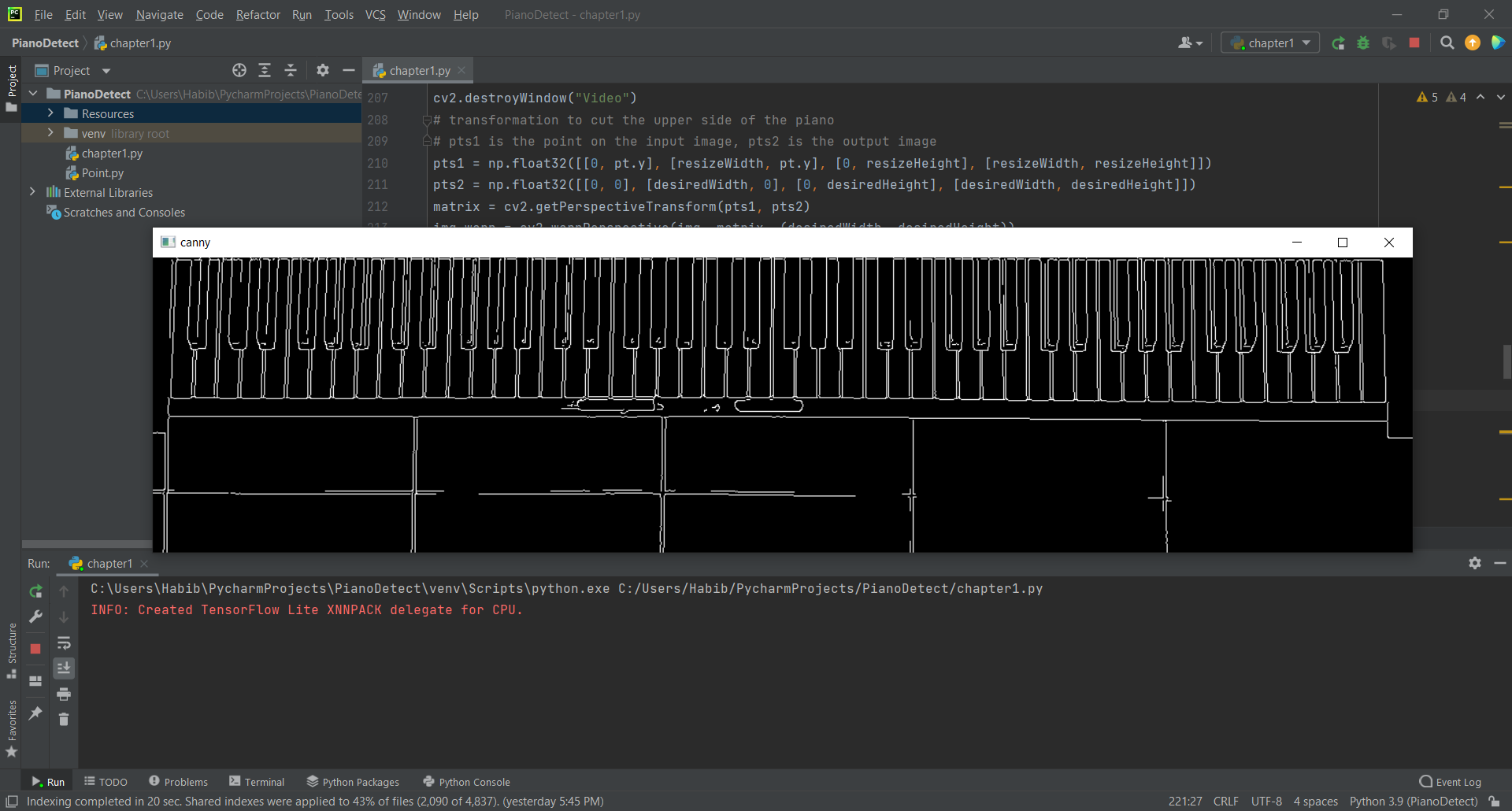
1. הופכים את התמונה לתמונה אפורה על יד שימוש ב cv2.cvtColor .



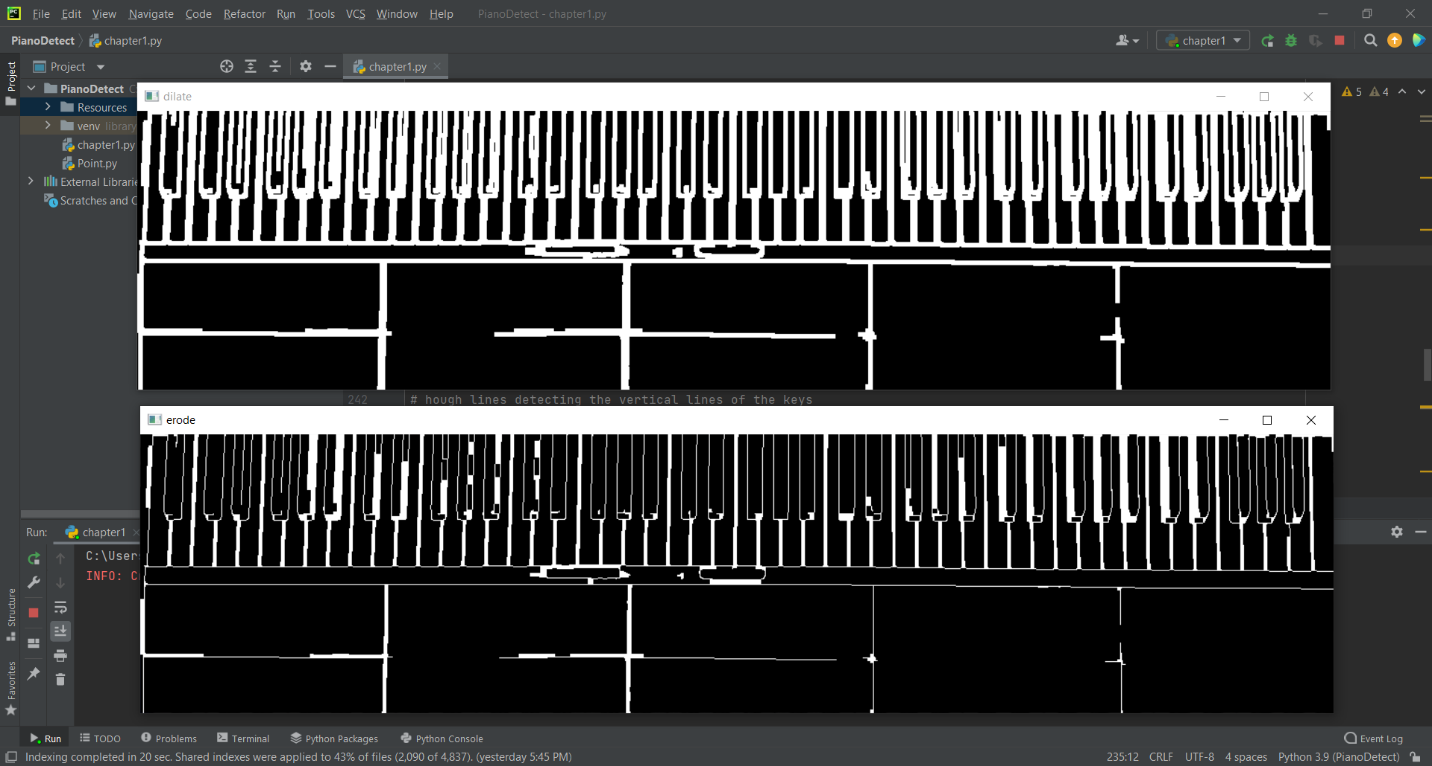
1. מטשטשים אותה בעזרת cv2.GaussianBlur .



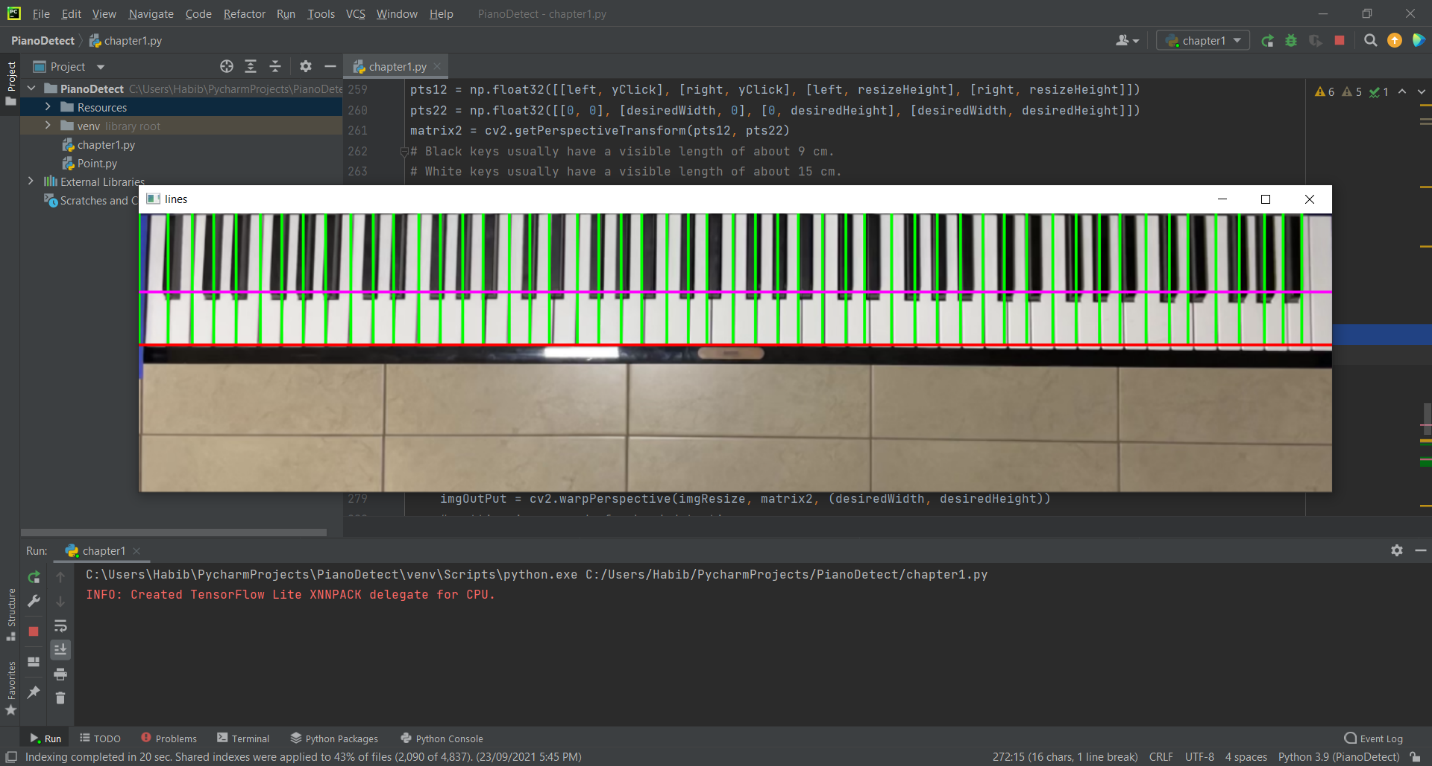
1. עושים איתור לספות ע"י שימוש ב cv2.Canny.



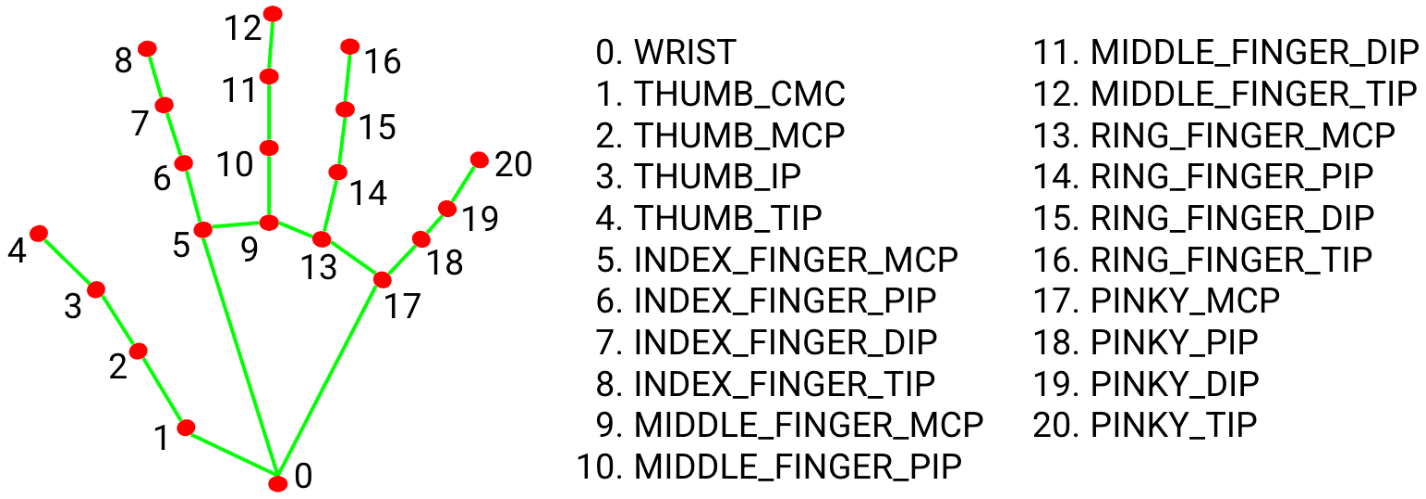
1. מפעילים Hough Line detection על תמונת ה Canny ע"י שימוש ב cv2.houghLines, כדי למצוא את הקווים האנכיים, וממינים אותם ומוצאים את הקו הראשון ואת הקו האחרון שהם צדדי הקלידים של הפסנתר.
2. נחתוך את שני הצדדים המיותרים ע"י cv2.warpPerspective ואז שומרים את המידות החדשות שקיבלנו כדי לעשות warp לכל frame בווידאו.
3. כדי שה Hough Line detection יהיה יותר מדויק, נצטרך להשתמש בשתי הפונקציות cv2.dilate ו cv2.erode כך שהראשונה מגדילה את רוחב הקווים בתמונה, כדי שכל שני קווים קרובים יותר מדי אחד לשני יתלכדו ואז הופכים לקו אחד, אחר כך הפונקציה השנייה עושה את ההפך כלומר מקטינה את רוחב הקווים.



1. מאתרים את הקווים האופקיים ע"י שימוש ב cv2.houghLines, מחפשים ושומרים את הקו התחתון של הקלידים. (הקו האדום בתמונה)
2. מאתרים את הקווים האנכיים גם ע"י שימוש ב cv2.houghLines.
3. כדי שנבחר את הקווים המפרידים בין כל שני קלידים מתוך כל הקווים האנכיים שמצאנו, בנינו פונקציית delete\_extra\_lines, מניחים שהפער בין שני קלידים הוא 3 פיקסלים, קיימים 52 קלידים, הרוחב של החלון שמציגים בו הפסנתר הוא 1280, אז : (1280-(53\*3))/52 =~21, כלומר רוחב כל קליד כמעט 20, לכן מסננים את הקווים ומשאירים את אלה שהמרחק ביניהם הוא 20. (הקווים הירוקים בתמונה)
4. מאתרים את הקו האופקי התחתון של הקלידים השחורים ע"י חישובים ליחס בין אורך הקלידים הלבנים והקלידים השחורים. (הקו הוורוד בתמונה)



1. עכשיו מאתרים את הידיים ע"י שימוש בפתרון קיים מmedia Pipe, הפתרון מגדיר את היד כ21 נקודות (landmarks) ומאתר מיקום כל נקודה, לכן עכשיו עוברים על כל frame בווידאו ושומרים ברשימה עבור כל frame את הקואורדינאטות x ו- y לכל נקודה בכל יד נמצאת בווידאו.



1. מחפשים על הframe הראשון כך שכל היד נמצאת מעל הקלידים " start\_f" (בודקים שכל קצוות האצבעות מעל הקו האדום שהסברנו בסעיף 10), בנוסף מחפשים על הframe הראשון כך שכל היד נמצאת מחוץ לפסנתר " end\_f" (בודקים שכל קצוות האצבעות מתחת הקו האדום). אחרי שנמצא אותם נמחק את כל ה frames שהם לפני ה start\_f ואחרי ה end\_f.
2. מבין כל ה21 נקודות של היד שומרים בכל frame, לכל אחד מהאצבעות את הקואורדינטות של שתי נקודות: נקודת ההתחלה של האצבע ונקודת הקצו שלו. (לדוגמא: לאצבע מורה אנו שומרים את הנקודות 5 ו 8 -מוצג בתמונת סעיף 14).
3. בכל frame מהווידאו אנו מחשבים את אורך כל אחד מהצבעות (מחשבים המרחק בין שתי הנקודות ששמרנו לכל אצבע) .
4. כדי שנחליט אם יש לחיצה על קליד או לא בחרנו לכל אצבע סף מסוים, אם אורך האצבע יורד מתחת לסף הזה אז יש לחיצה על קליד. לכן עבור כל frame אנו עוברים על האורכים של האצבעות ובודקים באיזה frame קיימת לחיצה.
5. לכל לחיצה על קליד מסוים הנחנו שזמן הלחיצה הוא לאורך 25 frames, לכן אחרי שמצאנו שיש לחיצה באיזשהו frame, נשמור אותו ונתעלם מה 25 frames שאחריו.
6. עכשיו יש לנו כל הנתונים כדי להדפיס את התווים שהוקלדו (המיקום של כל אצבע בכל frame, וה-frames שאתרנו בהם לחיצה על קליד) כל מה שנשאר לעשות הוא להדפיס את שם התו.
7. נציג את הסרטון וצובעים את הקלידים בזמן הנגינה.

###### הצעות להמשך הפרויקט

הדרך שבה היינו ממשיכים את הפרויקט היא:

1. במקום להשתמש במצלמה אחת שממוקמת מלמעלה אפשר להוסיף עוד שתי מצלמות מהצדדים לאיתור ירידת האצבעות, דבר זה יעזור כך שתהיה התוצאות עוד יותר מדויקות.
2. נחשב כמה זמן נלחץ כל קליד.
3. זיהוי אוטומטי של קלידי הפסנתר, כלומר בלי לבקש מהמשתמש ללחוץ על תמונה.
4. לעשות איתור ללחיצות בזמן אמת, כלומר לקבל דאטה ישירות   
   ממצלמה לייב, ולהציג על המסך את הוידאו עם צביעת הקלידים.



###### מה למדנו?

התחום כולו של הראייה ממוחשבת היה חדש עבורנו, ולכן נאלצנו לחפור לעומק (צפיה בסרטונים, קריאת מאמרים...) לפני שנוכל ליישם משהו.

קודם כל שנינו לא השתמשנו אף פעם בPython לפני זה, אז למדנו שפה חדשה לגמרי תוך העבודה על הפרויקט.

גם למדנו איך להשתמש בספריית OpenCV, צפינו בהרבה סרטונים וקורסים בחינם כדי שנוכל להשתמש בפונקציות שנמצאים בה.

בנוסף למדנו איך להיעזר ב MediaPipe framework כדי שנוכל לעשות איתור לידיים ואיך לאסוף מידע ולהשתמש בו.



###### צוואר-הבקבוק

החלק הכי קשה בפרויקט שבו נתקענו המון זמן הוא איתור הקלידים. היה לנו הרבה מכשולים שבגללם לקח לנו הרבה זמן כדי למצוא את הדרך היעילה ואידיאלית כך שיהיה האיתור הכי מדויק שאפשר.

הפסנתר היה ישן ולכן היו כמה קלידים פגומים כלומר המרחק בין שני קלידים לא היה זהה ואחיד אצל כולם.

האור המוחזר בגלל הצבע השחור של הפסנתר היה בעיתי מאוד.