

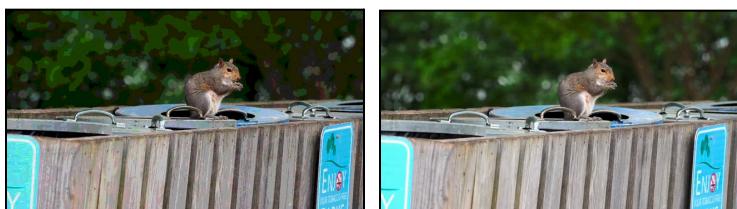
# Image Processing - Exercise 1

Shir Rashkovits, shirashko, 209144013

## מבוא

מטרת התרגיל היא לראות בפועל את השימוש בהיסטוגרמות ככלי חזק לייצוג ועיבוד תמונה, בפרט להבנה בין סכנות בסרטון, וכן לבצע היכרות עם ספריות רלוונטיות לעיבוד תמונה בפייטון. כמו כן, בתרגיל ראיינו דוגמה לקוונטיזציה, שהוא נושא נוסף עליו דיברנו בשיעור, ועל האופן שבו הוא יכול להשפיע על האלגוריתם שלנו במקרים מסוימים, ודורש התיחסות מיוחדת. הטכניקה העיקרית הייתה ייצוג של תמונות, פרימיים מתוך הודהו, על ידי היסטוגרמות (צוברות), וחשבון מרחקים בין היסטוגרמות צוברות של פרימיים עוקבים כדי להבחן בחזק ההבדלים בין פרימיים עוקבים.

ההבדל בין הקטגוריות הוא בכך שסרטונים מקטgorיה 2 מכילים גם פרימיים, תמונות, שעבורו קוונטיזציה לפחות מ 256 נציגות עצומות. לעומת זאת, יש פרימיים שמספר הצבעים בהם מוגבל יותר (יש מספר מצומצם יותר של עצומות שיכולה להופיע) מפרימיים אחרים, לעומת זאת 1 בה יש אחדות לאורך הסרטון. לדוגמה, למטה מצופים שני פרימיים שונים מאותה סצנה מזידאו 4. מימין, תמונה שלא עברה קוונטיזציה ומשמאל פרימם מאותה הסצנה אך עם הגבלת הצבעים, הניכרת ביחס בעיצים ברקע (למשל,



מעט נציגים עבור גווני י록) לאורך התרגיל כשאדרב על פרימיים שעבורו קוונטיזציה, אתיחס לפרימיים שמקילים פחות מ 256 נציגים, אך כמובן שככל פרימם עבר קוונטיזציה כי גם 256 נציגים זה מיפוי של ערכי קלט רבים לערכי פלט מועטים.

## אלגוריתם

- 1) קיבל נתיב לקובץ של סרטון
- 2) נבדק את הסרטון לקבלת ייצוג של הפרימיים המרכיבים אותו בייצוג grayscale
- 3) ניצור היסטוגרמה עם 256 בנים שימוש שמקפת כמה פיקסלים יש מכל עצמה לכל אחד מהפרימיים
- 4) בעזרת החישוב הקודם, נחשב את היסטוגרמה צוברת עם 256 בנים לכל אחד מהפרימיים
- 5) נחשב מרחק וקטורי אלי בין היסטוגרמות הצוברות של פרימיים עוקבים (בעזרת נורמה L1 על הפרשי הווקטורים)
- 6) נחשב ונdziיר את האינדקסים של שני הפרימיים העוקבים שמרחקם מוקסימלי

ההבדל בין האלגוריתם המוצע לבין האלגוריתם שיכל להתאים עבור הקטגוריה הראשונה בלבד, הוא שימוש בהיסטוגרמות צוברות. כאמור, אלגוריתם 1 מدلג על שלב 4 ומשתמש בשלב 5 בהיסטוגרמות שלב 3.

## פרטי מימוש

- הסרטון מתפרק מהתוך נתיב במחשב, ולאחריה הסרטון למערך הפרויימים בפורמט grayscale.
- נעזרת בספריה mediapy.
- קבועתי פרמטר של מספר הבינים בהיסטוגרמות שווה למספר העצומות 256, כך שכל בין מכיל את כל הפיקסלים עם עוצמה 0 כך שמספרם שלם בין אפס ל-255. בחירה זאת נבעה מכך שהסרטונים בתרגיל (וקרייתם עם mediapy של `read_video`) מיצגים עם 8 bpp (גם אלו שעברו קומונטיזציה עם עצומות בטוחה של 0-255). מכיוון שהאלגוריתם מנוסה למצוא הבדלים בין פרויימים על בסיס סוג וכמויות העצומות שמצוות בתמונות המרכיבות את הסרטון, זאת הייתה הבחירה הגדאית ביותר, מכיוון שמספר קטן יותר או הכנסה של פיקסלים באופן אחר אל הבינים היו עלולים להוביל לאיבוד מידע (ميزוג של כמויות עצומות שונות לביניהם היה מוסף בין ייחד) וחינוי למציאת הבדל בין פרויימים, ומספר גדול יותר של בינים היה מוסיף לבנים ריקים ולכן היה חסר משמעות או שהוא מפריד פיקסלים שדוחקם יחד כדי שנוכל לזהות כמהות כל עצומה בתמונה.
- נעזרת בספריה `umpy` לצורך חישוב ההיסטוגרמות ולאחר מכן לחישוב ההיסטוגרמות הצוברות.
- המרחק הוקטוריאלי בין ההיסטוגרמות הצוברות של פרויימים עוקבים חושב לפי הנוסחה שראינו בשיעור. לצורך חישוב עיל של המרחק בין ההיסטוגרמות הצוברות (שלמעשה מיוצג על ידי וקטור), נעזרת בספריה `skimage` עבור פונקציונליות של חישור, ערך מוחלט וסכימה.
- לבסוף חישוב הפרויימים העוקבים שהתקבל בעברם מערך מקסימלי געשה על ידי מציאת `argmax` במערך המרחקים בין הפרויימים.
- כמו כן נעזרת בספריה `matplotlib` לצורך יצירת גרפים להציג התוצאות.

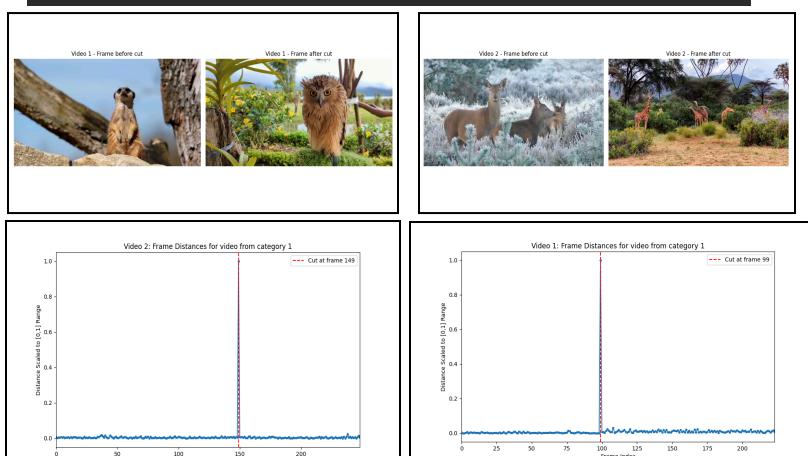
הकושי העיקרי שנטקלתי בו בתרגיל הוא חוסר ניסיון רחוב בספריות הנדרשות כגון `numpy`, `mediapy` וכו'. לכן חלק ניכר מהתרגיל היה הינה מקדימה של `bootcamp` וקריאה על הספריות הללו, בניסיון להבין איך אפשר להשיג בעזרתן באופן יעיל את הפונקציונליות שרציתנו למשמש.

כמו כן היו לי מחשבות שונות למימושים אפשריים כגון שימוש בייצוג מקומי ולא רק בייצוג גלובלי של כל פרויים (כפי שדיברנו בשיעורי היסטוגרמה בשיעור), או שימוש ב-`threshold` מסוים במקום `max arg`. לגבי האופציה הראשונה, קיים טרייד אוף מרכיב יחסית וניצרכים סרטונים רבים כדי לקבוע היפר פרמטר של מספר החלקים ואופן החלוקה כדי ליצר "יצוגים מקומיים" וכן לבסוף החלטתי שנדרש פתרון שפותחות דורש הסתכילות של מידת מכונה ושלצורך התרגיל האלגוריתם שסיפקתי יהיה לתוצאות הרצויות. לגבי הסוף, מכיוון שאנו יכולים להסתמך על כך שיש לבדוק סצנה אחד, הבנתי שהבחירה המתאימה יותר במקרה זה היאבחירה של ההבדל המקסימלי (לעומת מצב שבו למשל לא מובטח שהיא מעבר סצנה, ואז יתכן שכדי להציב רף, כי אחרת יתכן שנזהה מעבר סצנה כשלעצמה אין מעבר כלל).

Video 1 (Category 1): Cut frame index: 99, Next frame index: 100  
 Video 2 (Category 1): Cut frame index: 149, Next frame index: 150

## תוצאות קטgorיה 1

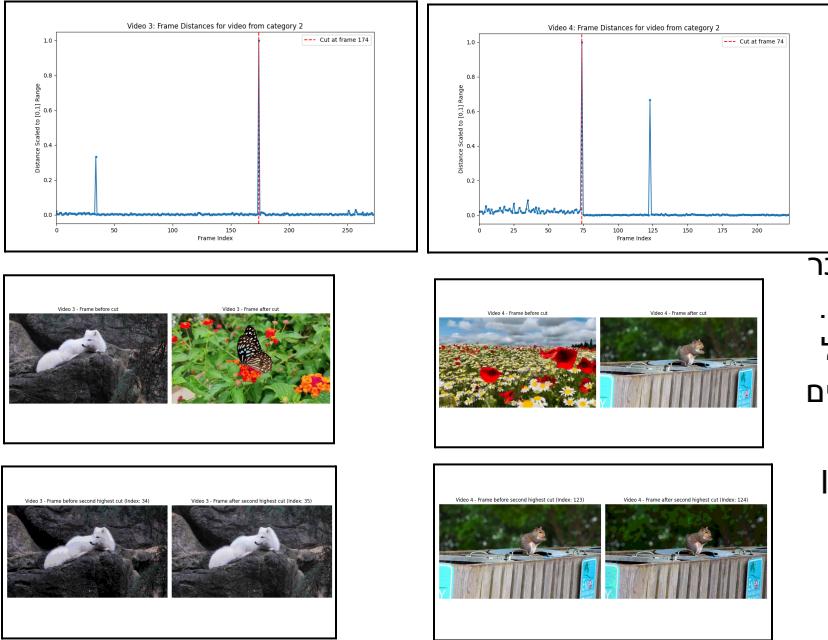
מצורפות התוצאות שמחזיר האלגוריתם, כמו כן הפרויימים המתאים להמחשה. האלגוריתם מצליח לזהות את הפרויימים בהם חיל המעבר בין הסצנות ללא הסחתה עבור קטgorיה זו. זה ניכר בגרפים משמאל, בהם ניתן לראות



שמתקנים הבדלים מינוריים במרחקים עברו פריים מאותה הסצנה, ביחס להבדל הבולט המתקיים מעבר בין הסצנות השונות. כזכור, מעבר הסצנה מתקיים על בסיס המרחקים הללו.

## תוצאות קטgorיה 2

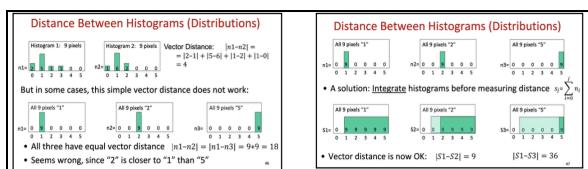
Video 3 (Category 2): Cut frame index: 174, Next frame index: 175  
 Video 4 (Category 2): Cut frame index: 74, Next frame index: 75



גם בקטgorיה 2 ניתן לראות שהאלגוריתם הצליח להבחן מעבר בין הסצנות, כאשר הפלט שהאלגוריתם מחזיר מצורף למללה. אולם, בקטgorיה זו ניתן לראות שקיימות עלייה נוספת נוספת בגרפים. משמאלו מוצגים גם הפריים המתאים לשתי העליות החודדות, כאשר מלמעלה הפריים עברו העלייה החדה יותר שם האלגוריתם הכריע שה cut scene מתרחש. ניתן לראות לפי הפריים עם המרחק השני הגדל ביותר, שהפיק השני הוא הסחה של הפריים (או מפריים) שעבר קוונטייזציה. אולם, ניתן לראות שהאלגוריתם הצליח להתגבר על ההסחה ולהבחן נוכנה מעבר בין הסצנות גם בקטgorיה זאת (הקו האדום בגרף המרחקים מסמן את הפריים הנכון).

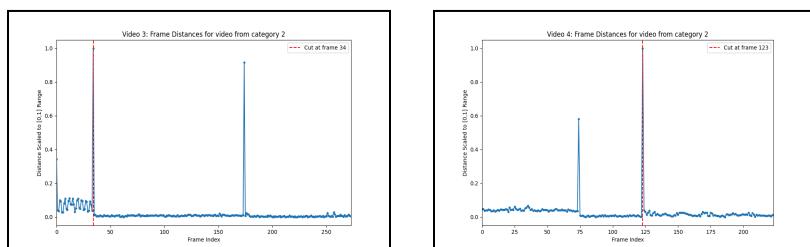
ההבדל בין שתי הקטgorיות הוא שבקטgorיה 2 יש מעבר באחת מתחר שתיהן הסצנות מפריים לא קוונטייזיה לפריים שעברו קוונטייזציה (או הסדר הפוך). עברו הקטgorיה הראשונה, פחות מסוכן לעבד עם היסטוגרמה רגילה, מאחר שפריים השיכים לאוותה סצנה מכלים כמוות דומה מאד של פיקסלים מכל עצמה. אולם, כפי שראינו בשיעור, מרחק בין היסטוגרמות כואمدنן לדמיון תמונות אינם רגיש מספיק לאחר שמתיחס להבדל בין עצמות קרובות באותו האופן המתיחס לעצמות רוחקות (חישוב המרחק הוקטוריאלי מושפע רק מדמיון של ביןיהם בהתאם מדוקת בין היסטוגרמות, ככלומר התרומה של בין זיהוי בהיסטוגרמה 1 היא רק בהבדל שלו מבין זיהוי בהיסטוגרמה 2). כך בקטgorיה השנייה, יש מעבר מסיבי מפריים שכל 256 העצמות יכולות להתבטא, לפריים שעבר קוונטייזציה. ככלומר, יש עצמות שאין לנו נציגים, פיקסלים שלפני קוונטייזציה היו מיוצגים על ידי עצמות שלא ניתן לבטא יותר, ובקוונטייזציה הם הוזזו לעצמות אחרות, קרובות יחסית (כי בקוונטייזציה רוצים לאבד כמה שפחות מידע וליצג באופן כמה שיותר נאמן את התמונה). בגלל אי הריגשות המתווארת של היסטוגרמות, במקרה של הקטgorיה השנייה סביר שההיסטוריה רגילות יגיעו לביצועים לא טובים, בזיהוי של מעבר מסצנה שעברה קוונטייזיה לסצנה ללא קוונטייזציה כמעבר הסצנה (בגלל אי הריגשות במרחב בין עצמות קרובות לעצמות רוחקות). לעומת זאת, כשהנשתמש בהיסטוגרמות צוברות, נצפה לתוצאות הרבה יותר דומות עבור פריים מאותה סצנה, גם אם מספר נציגי העצמות לא זהה, בغالל שגם אם הפיקסלים הוזזו מעט באחד הפריים, כצוברים אותם אז במבט כללי, בלבד מעט ביניים, נקבל ייצוג יחסית דומה, כי מספר הפיקסלים בסביבה מסוימת של עצמה דומה בשני הפריים, מה שיביא לכך שההו הרבה פחות הפרשים בין זיהוי בהיסטוגרמה 1 לבין זיהוי בהיסטוגרמה 2. זאת לעומת מעבר בין פריים מותאם של

סצנות שונות, ביןיהם מספר הפיקסלים בסביבה מסוימת של עצמה לרוב יישאר שונה גם כמספרotal על ההיסטוגרמות הצبورות, בגלל שבאופן טיפוסי בסצנות שונות יהיה הרכב שונה של כמותם של צבעים (או גווני אפור), וכן יהיו עליות באזוריים שונים בהיסטוגרמות הצبورות.

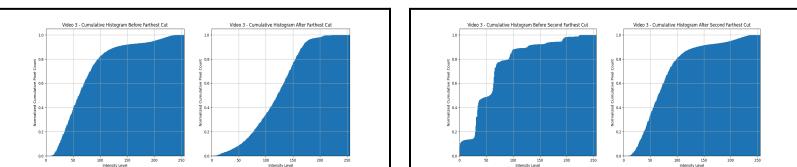
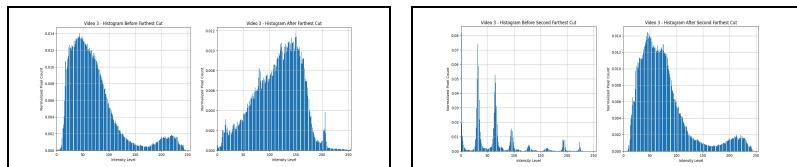


משמאל למפה מצורפים גרפים המראים את המרחקים שהתקבלו עם אלגוריתם 1 (בין ההיסטוגרמות, ללא היסטוגרמות צבורות) עבור הסרטונים בקטgorיה 2. ניתן לראות שאלגוריתם זה

אכן מושך על ידי המתואר לעיל ו שימוש בו היה מוביל אותנו לתוצאה שגוייה.



כפי שתיארתי, לאחר שאגוריתם 2 משתמש בהיסטוגרמה צבורת, על אף שיש הבדל גדול יחסית בין פריים שעבר קוונטיזציה לצה שלא, ההיסטוגרמות של פריים מאותה סצנה עדין דומות יותר, ולכן יש שנחסר וננסcum את הפרשי בערך מוחלט של הבנים, נקבל שבסך הכל בין רוב הבנים יהיה הפרש בערך מוחלט פחות משמעותית מאשר בעבר הסצנה). כעת אציג גרפים שמטרתם לנסוט להמחיש את הסיבה שהאלגוריתם הראשון נכשל בעוד שהאלגוריתם השני מצליח. מלמטה הצד שמאל ניתן לראות את ההבדלים בין ההיסטוגרמות להיסטוגרמות הצבורות עבור פריים שביניהם חל מעבר הסצנה, ומימין ניתן לראות את ההבדלים עבור פריים שביניהם חל מעבר מפריים עם קוונטיזציה לפרקאים. מימין, ניתן לראות בהistogramma הרגילה שמאחר ויש מספר נציגים מוגבל, יש שניים רב בין ההיסטוגרמות, אך יחד עם זאת ה"פיקים" קורים באזוריים דומים. لكن ההיסטוגרמות הצבורות המתקבלות הן עם מגמה יחסית דומה, של עלייה משמעותית בחצי הראשון, יש בעיקר פיקסלים כהים. ממשאל לעומת זאת, ניתן לראות שההיסטוגרמות הרגילה שמאחר ויש מספר נציגים מוגבל, יש את מכיוון שבפרקאים השיר לסצנה הראשונה יש בעיקר פיקסלים כהים ובפרקאים השיר לסצנה השנייה יש ריכוז גבוה יותר של פיקסלים בתחום האמצע (לא בהירים ולא כהים במיוחד).



## מסקנות

لتמונה דומות היסטוגרמות דומות, ועל ידי שימוש בתובנה זאת ניתן לאבחן את נקודת שינוי הסצנה בידאו. אולם, טענה זאת לא מדוקנית, ונכון יותר לומר שלתמונה דומות יש היסטוגרמות צבורות דומות. בתרגום ראייתי

בפועל כיצד ההיסטוגרמות מייצגות תמונה, ומה החישובות של ההיסטוגרמה צבורת ושל התיחסות מיוחדת למצב של ייצוג שונה של תמונה בשילוב אלגוריתם רובוטי שמנע מטעויות. כמו כן, ביצועי היכרות עם ספריות שימושיות לעיבוד תמונה.