

Image Processing - Exercise 1

Ofek Avidan, ofek.avidan, 318879574

מבוא

התרגיל נועד ליצור תוכנית Python המסוגלת לזהות אוטומטית שינוי סצנה בסרטון. המונח "שינוי סצנה" (או חיתוך סצנה) מתייחס לנקודה בסרטון שבה מתרחש מעבר מסביבה או הקשר אחד לאחר, המעיד על שינוי בתוכן החזותי. המטרה העיקרית היא לפתח אלגוריתם שיכול לזהות ביעילות (בזכות ספריות חשובות ב-Python וחיסכון בלולאות) את המעברים הללו בוידאו. הטכניקה המרכזית המופעלת בתוכנית זו היא ניתוח היסטוגרמה, ובפרט חיסור בין היסטוגרמות של פריימים עוקבים (הרצאה 1, שקופית 46). היסטוגרמות מספקות ייצוג סטטיסטי של התפלגות עוצמות הפיקסלים בתמונה. על ידי הרחבת הרעיון הזה לפריימים של וידאו, נוכל לבחון את ההתפתחות של התפלגות עוצמת הפיקסלים לאורך זמן. הרעיון הוא ששינוי בסצנות תואם לרוב בשינוי ניכר בין היסטוגרמות של שני פריימים עוקבים כפי שלמדנו בהרצאה. בחירת טכניקת הניתוח תלויה בקטגוריית הסרטון. אם הסרטון מסווג כסוג 2, נעשה שימוש באלגוריתם חזק יותר. אלגוריתם זה כולל חישוב היסטוגרמות מצטברות ולאחר מכן ניתוח ההבדלים בין היסטוגרמות מצטברות עוקבות. מצד שני, לסרטונים מסוג 1, מספיק אלגוריתם חלש יותר. במקרה זה, היסטוגרמות רגילות של פריימים עוקבים מופחתות זו מזו.

סרטוני וידאו בקטגוריה 1 מורכבים משתי סצנות עם חתך סצנה בודד, המציגים שינויים מינימליים בצבע (וכתוצאה מכך גם בגווי אפור) בתוך הסצנות. האלגוריתם של קטגוריה 1, בהיותו מתון יותר, מתאים למעברים הדרגתיים ללא שינויי צבע גדולים. לעומת זאת, סרטוני קטגוריה 2, הכוללים חתך סצנה ושינויים ניכרים בצבע (שוב, או ברמת אפור) בתוך הסצנות, דורשים אלגוריתם חזק יותר. עיבוד זה מטפל הן בחיתוכים פתאומיים והן בשינויים משמעותיים בצבע או ברמת אפור, המשקפים את המאפיינים הייחודיים של סרטוני קטגוריה 2.

אלגוריתם

האלגוריתם שבו השתמשתי כדי לזהות שינוי סצנה הוא:

1. עיבוד מוקדם של וידאו (קריאת הנתיב לסרטון הקלט והמרה לרצף של פריימים בגווי אפור באמצעות מטריצת המרת RGB לגווי אפור מוגדרת מראש, עליה למדנו בהרצאה)
2. חישוב היסטוגרמה (עבור כל פריימים בסרטון, המייצגות את התפלגות עוצמות הפיקסלים)
3. בחירת אלגוריתם (בהתבסס על הקטגוריה של הסרטון, נקבע אם להשתמש באלגוריתם חלש יותר (חיסור היסטוגרמות רגילות) או באלגוריתם חזק יותר (חיסור היסטוגרמות מצטברות)).
4. חיסור היסטוגרמות (אם קטגוריה 1, נחשב את ההבדלים בין היסטוגרמות רגילות של פריימים עוקבים. אם קטגוריה 2, נחשב את ההבדלים בין היסטוגרמות מצטברות עוקבות)

5. חישוב נורמה אוקלידית (נחשב את הנורמה עבור כל שתי שורות עוקבות של המטריצה שמייצגת את הפרשי ההיסטוגרמה)
6. זיהוי נורמה מקסימלית (נזהה את הפריים עם הנורמה האוקלידית המקסימלית, כלומר, את הפריימים העוקבים ה"שונים ביותר" מבחינת היסטוגרמה, שמציינים את שינוי הסצנה)
7. יצירת פלט (נחזיר את הסצנה שזוהתה כ-tuple, המייצגת את אינדקס הפריימים האחרון של הסצנה הראשונה ואת אינדקס הפריימים הראשון של הסצנה השנייה)

השינויים העיקריים בין אלגוריתמי זיהוי שינוי הסצנה עבור שתי קטגוריות הווידאו כוללים את בחירת האלגוריתם ואת השיטה של הפרשי ההיסטוגרמות. עבור סרטוני קטגוריה 1, נעשה שימוש באלגוריתם "חלש" יותר, המפחית היסטוגרמות רגילות כדי ללכוד וריאציות עדינות בתוך סצנות. לעומת זאת, סרטוני קטגוריה 2 משתמשים באלגוריתם "חזק" יותר, המנתח הבדלים בין היסטוגרמות מצטברות עוקבות כדי לתת מענה לשינויים בולטים יותר, כולל חיתוכים בסצנה ושינויים משמעותיים בצבע (או ברמת אפור). התאמות אלו נותנות מענה למאפיינים הייחודיים של כל קטגוריה, ומבטיחות את יכולת ההסתגלות והדיוק של האלגוריתם בזיהוי שינויים בסצנה על פני תרחישים מגוונים של תוכן וידאו.

פרטי מימוש

ביישום אלגוריתם זיהוי חיתוך הסצנה, התהליך מתחיל בקריאת סרטון הקלט באמצעות ספריית ה-'mediapy' והמרת הפריימים מ-RGB לגווני אפור באמצעות מטריצה מוגדרת מראש (עליה למדנו בהרצאה). ספריית numpy משמשת לאחר מכן לביצוע פעולות מטריצה, חישוב היסטוגרמות וחישוב סכומים מצטברים. בהתאם לקטגוריית הווידאו, נבחר אלגוריתם חלש יותר או חזק יותר. עבור סרטוני קטגוריה 1, היישום כרוך בהפחתת היסטוגרמות רגילות של פריימים עוקבים, בעוד עבור סרטוני קטגוריה 2, מנותחים הבדלים בין היסטוגרמות מצטברות עוקבות. הנורמה האוקלידית מחושבת באמצעות פונקציונליות האלגברה לינארית של numpy, מה שמקל על כימות השינויים הכוללים בהתפלגות עוצמת הפיקסלים. הפונקציה 'argmax' מזהה את הפריים עם הנורמה המקסימלית, ומציינת את שינוי הסצנה.

יישום האלגוריתם מסתמך בעיקר על פונקציונליות קיימת מספריות 'numpy' ו-'Numpy'. 'mediapy' חיוני עבור פעולות מערך, חישובי היסטוגרמה וחישובי אלגברה לינארית, בעוד ש-'mediapy' מטפל במשימות קריאת וידאו ומניפולציה. ספריות אלו נבחרו בשל יעילותן ותמיכה מקיפה בפעולות מספריות ועיבוד מולטימדיה ב-Python.

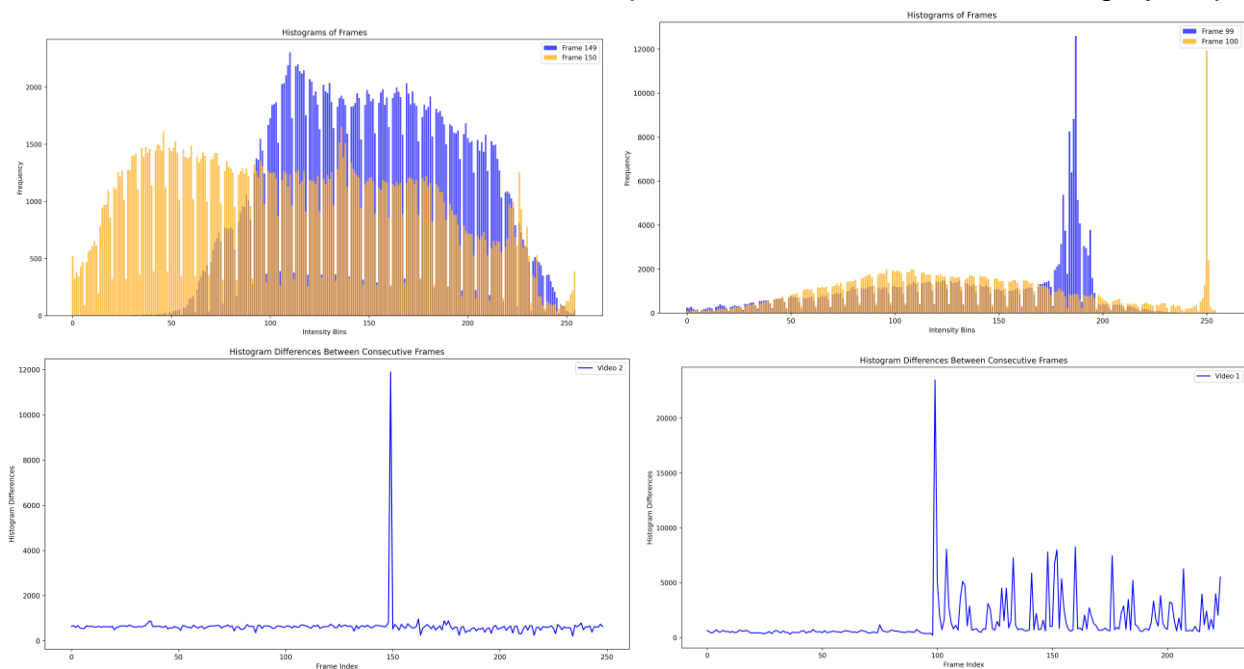
לגבי היפר-פרמטרים וספים, קבועי מטריצת ההמרה של RGB לגווני אפור, וההחלטה באיזה אלגוריתם להשתמש בהתבסס על קטגוריית הווידאו, משמשים כבחירות קריטיות. מטריצת RGB לגווני אפור (עליה למדנו בהרצאה) היא פרמטר קבוע, המבטיח המרה עקבית בין פריימים. פרמטר קטגוריית הווידאו מנחה את התנהגות האלגוריתם, וקובע אם ליישם את הגישה החלשה או החזקה יותר. בחירות אלו נעשות כדי ליצור איזון בין רגישות לשינויים בסצנה ויעילות חישובית.

במהלך היישום, נתקלו באתגרים בכווןן עדין של האלגוריתם לטיפול בתוכן וידאו מגוון ביעילות. בוצעו התאמות לפרמטרים האלגוריתמיים, כגון שיטות חיסור היסטוגרמה וחישובי נורמה, כדי להגביר את הרגישות לשינויים עדינים ובולטים כאחד. בדיקות ותיקוף איטרטיביות נערכו כדי להבטיח חוסן על פני תרחישי וידאו שונים, התמודדות עם אתגרים וחידוד האלגוריתם לשיפור הביצועים. באופן אישי אוסיף שלא לקחתי קורס כמו IML בעבר, ולכן היה לי מאתגר מאוד לעבוד עם numpy.

תוצאות קטגוריה 1

התוצאות עבור זיהוי שינוי הסצנה של קטגוריה 1 הן:

video1_category1.mp4 מתבצע שינוי בסצנה בין הפריימים 99-100
video2_category1.mp4 מתבצע שינוי בסצנה בין הפריימים 149-150

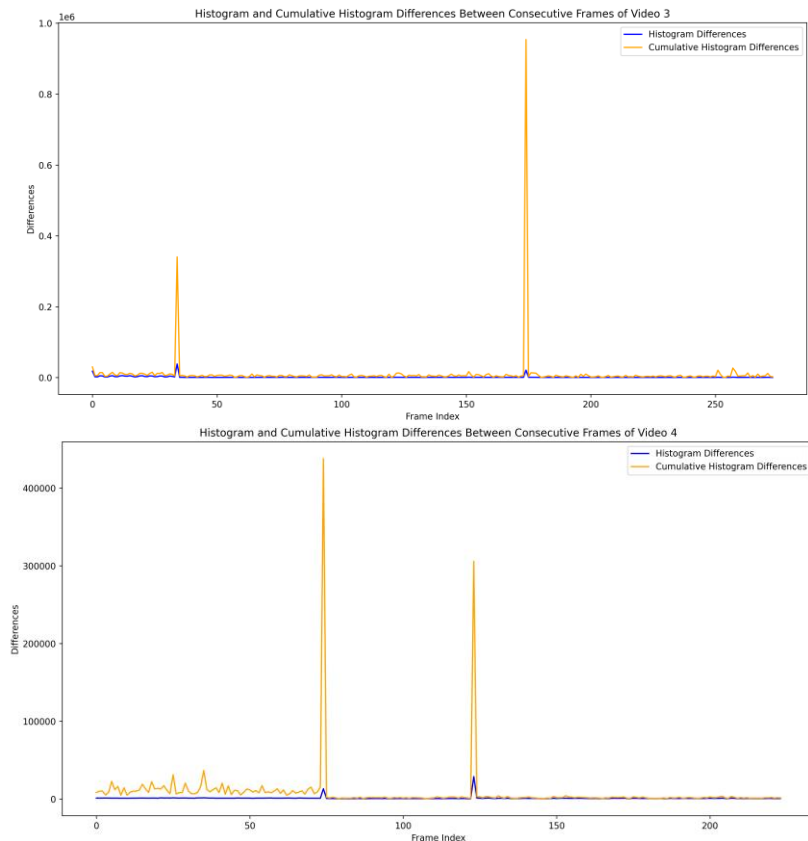


בקטגוריה 1, הן בוידאו 1 והן בוידאו 2, הניתוח של פריימים עוקבים חושף דפוס עקבי של הבדלים נמוכים יחסית בעוצמות הפיקסלים, מה שמצביע על מעבר הדרגתי בתוך הסצנה. עם זאת, נצפית חריגה משמעותית בפריימים 99-100 של סרטון 1 ובפריימים 149-150 של סרטון 2. בפריימים אלו, יש עלייה בולטת בערכי ההפרש. קפיצה זו מרמזת על הבדל מהותי בהיסטוגרמות, מה שמצביע על שינוי סצנה. בזכות ההבטחה כי ישנו שינוי אחד בלבד בין שתי סצנות, ובהסתמך על חומר שהועבר בהרצאה, נוכל להגיד בוודאות כי נקודת המקסימום היא הנקודה בה מתרחש שינוי סצנה.

תוצאות קטגוריה 2

התוצאות עבור זיהוי שינוי הסצנה של קטגוריה 2 הן:

video3_category2.mp4 מתבצע שינוי בסצנה בין הפריימים 174-175
video4_category2.mp4 מתבצע שינוי בסצנה בין הפריימים 74-75



בקטגוריה 2, מתגלות שתי נקודות גבוהות בגובה דומה, מה שמבדיל אותה מקטגוריה 1. ניתן לראות בתמונה הימנית התחתונה שתי נקודות גבוהות, ב-175 ו-174 וכן ב-35 ו-34. יש לציין שהעלייה הקטנה הזו, המשקפת שינויים בעומק הצבע מבלי להשפיע על הסצנה הכללית, מביאה להשפעה פחות בולטת על הנורמה שחישבנו בין פריימים עוקבים. למעשה, הנקודה של 34-35 היא גבוהה יותר מ-174-175, ולכן אם היינו מסתמכים על האלגוריתם מקטגוריה 1, היינו טועים (וידאו 3). הדבר נכון גם לוידאו 4, שכן במקום לקבל 74-75 כנדרש, היינו מקבלים 123-124. כפי שלמדנו בהרצאה 1, האלגוריתם ה"חלש" יותר לא מספיק כשמדובר בתמונות חלקות (אחידות) יותר. וזו הבעיה בקטגוריה 2.

מסקנות

במהלך תרגיל 1 בקורס, המוקד שלנו היה פיתוח, מימוש וחידוד של אלגוריתם לזיהוי שינוי סצנה המותאם הן לסרטוני קטגוריה 1 והן לקטגוריה 2. האלגוריתם, המתאים בצורה מיומנת למאפיינים הייחודיים של כל קטגוריה, רתם את הכוח של הבדלי ההיסטוגרמה עבור קטגוריה 1 והפגין תחכום בקטגוריה 2 על ידי הבחנה יעילה בין חיתוכים גדולים ושינויים עדינים יותר של צבע/רמת אפור באמצעות היסטוגרמות מצטברות. עזרים חזותיים, לרבות גרפי עמודות וגרפים קווים, מילאו תפקיד מכריע באימות ביצועי האלגוריתם (חלק התוצאות), בעוד שתצפיות בעוצמת הפיקסלים לפני ואחרי חיתוכים של הסצנה סיפקו תובנות חשובות. באופן כללי, נהייתי לממש את החומר ולגעת בפעם הראשונה בחיי ב-numpy.