עיבוד תמונה – תרגיל בית 3

אוריין יוסף חסידים | oryan.hassidim | 319131579

# מבוא

בתרגיל זה נדרש לערבל תמונות ע"פ מסיכה וליצור תמונה היברידית (שממרחקים שונים רואים תמונות שונות).

דוגמא לערבול תמונות לפי מסכה דוגמא לתמונה היברידית

נשתמש בטכניקת פירמידות – פירמידות גאוסיאן ופירמידות לפלסיאן כדי לבצע את הנ"ל. פירמידות נותנות לנו לפרק את התדרים בתמונה למספר רמות.

# אלגוריתם

## ייצור הפירמידות

### ייצור קרנל גאוסיאני חד ממדי

באמצעות קונבולוציה רפטטיבית עם עד ההגעה לארוך הרצוי.

### ירידת רמה

קונבולוציה עם הקרנל הגאוסיאני החד ממדי הן לאורך והן לרוחב, ואז דגימה לסירוגין באורך וברוחב (כרבע מהפיקסלים, תלוי בזוגיות הממדים).

### עליית רמה

יצירת מטריצת אפסים בגודל הרצוי (שהוא לא תמיד בדיוק פי שניים בכל ממד), שימת התמונה עם הרזולוציה הנמוכה בפיקסלים האי זוגיים באורך וברוחב, שוב קונבולוציה עם הקרנל החד ממדי לאורך ולרוחב והכפלת העוצמה לפי יחס הגדלים מהתמונה המקורית לחדשה.

### ייצור פירמידת גאוסיאן

באופן רפטטיבי לרדת רמה ולשמור את כל הרמות ברשימה.

### ייצור פירמידת לפלסיאן

לקבל פירמידת גאסיאן ולחשב הפרשים בין כל צמד סמוכות (כמובן אחרי עליית רמה של הקטנה מביניהן) ולשמור במקום הגדולה.

**❣️**הדבר **החשוב ביותר בכל התהליך!** – המרה למטריצות של int16 ממטריצה של uint8, כדי שנוכל לקבל מספרים מכוונים ואת כל הטווח שצריך (שהוא ).

### בניית תמונה מפירמידת לפלסיאן

באופן איטרטיבי החל מהרמה הגבוהה ביותר לעלות רמה ולהוסיף את הרמה הנמוכה יותר (שבה שמרנו מקודם את ההפרש). המרה חזרה ל-uint8.

## ערבול תמונות לפי מסיכה

פרמטרים: תמונה 1, תמונה 2, מסיכה, מספר רמות לפירמידות, אורך הקרנל הגאוסיאני לטשטוש.

1. לייצר פירמידות לפלסיאן לשתי התמונות – lpA, lpB.
2. לייצר פירמידת גאוסיאן למסכה – lpM.
3. ליצור פירמידה משולבת: .
4. לבנות מהפירמידה המשולבת תמונה.

האלגוריתם הועבר בשלמותו בהרצאה (ואף כתוב בשקפים) – הטשטוש (קונבולוציה עם גאוסיאן בשני ממדים), פירוק לפירמידת גאוסיאן ובנייה חזרה, בניית פירמידה למסכה.

ערכי שני הפרמטרים האחרונים – מספר רמות לפירמידות ואורך הקרנל הגאוסיאני לטשטוש – משפיעים על התוצאה ואף משתנים בין תמונה לתמונה, אך ערכי ברירת המחדל שנתתי להם, שעבדו עבור התמונות שלי, הם 6 ו-5 בהתאמה.

## תמונה היברידית

פרמטרים: תמונה 1, תמונה 2, מספר רמות לפירמידות, אורך הקרנל הגאוסיאני לטשטוש.

1. לייצר פירמידת לפלסיאן בגובה המבוקש לשתי התמונות.
2. לקחת את הפירמידה של התמונה שרואים מקרוב, ולהחליף את הרמה הגבוהה ביותר (הווי אומר הקטנה ביותר, שאינה הפרשים) ברמה הגבוהה ביותר של התמונה שנראה מרחוק.
3. לבנות מהפירמדה הנ"ל את התמונה ההיברידית.

ההיגיון מאחורי האלגוריתם הוא שכמקטינים את התמונה אז המסך בעצם דוגם את התמונה, ולכן יציג רק את התדרים הנמוכים ביותר, וכאשר מגדילים אותה אז המסך יציג גם תדרים גבוהים.

ערכי שני הפרמטרים האחרונים – מספר רמות לפירמידות ואורך הקרנל הגאוסיאני לטשטוש – משפיעים על התוצאה ואף משתנים בין תמונה לתמונה, אך ערכי ברירת המחדל שנתתי להם, שעבדו עבור התמונות שלי, הם 3 ו-5 בהתאמה.

## פרטי מימוש

אין לי מה להוסיף על הכתוב בנ"ל מבחינת פרטי המימוש של כל צעד (מלבד ערכי הפרמטרים האופציונליים שפירטתי לעיל). כאמור, הדבר החשוב היה לשמור מספרים מכוונים בפירמידת הלפלסיאן.

מלבד הקונבולוציה עם התמונה שהשתמשתי במובנה של OpenCV, את כל שאר התהליכים בניתי בעצמי מ-0 (כמובן עם שימוש ב-NumPY).

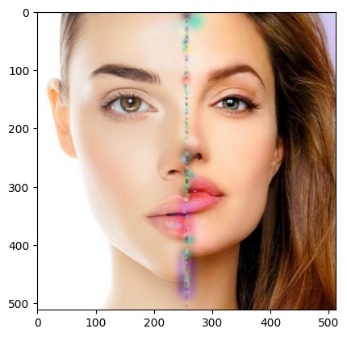
בחרתי להשתמש ב-OpenCV וב-NumPY בגלל הדוקומנטציה הנוחה שלהם, הפופולריות שלהם ושנוח מאוד להשתמש בהם.

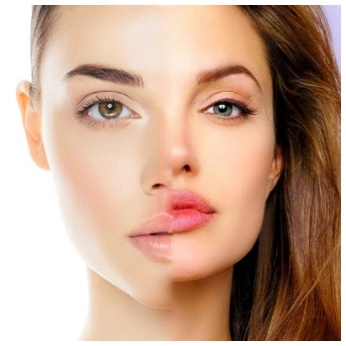
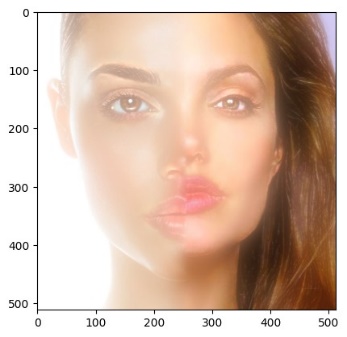
## אתגרים

### מספר הקסם 4?

ע"פ הדוקומנטציה באתר של OpenCV, בעליית רמה בפירמידה מכפילים את הערכים ב-4. ניסיתי להבין את ההגיון בזה ובעיקר באיזה פרמטרים זה תלוי. בהתחלה חשבתי שזה תלוי באורך הגאוסיאן, אבל ערכים שונים לא נתנו את התוצאה המיוחלת, עד שהבנתי שבעצם זו החזרת העוצמה לגודל התמונה החדשה ולכן הפקטור הוא יחס ממדי התמונות.

### האתגר הגדול – מספרים מכוונים

תמונה שמכילה פני אדם, ריס, עור, גבה

התיאור נוצר באופן אוטומטי

כמנהגי בקודש בשני התרגילים הקודמים, די בקלות ובמהירות כתבתי את האלגוריתמים, אבל נתקעתי על איזו בעיה טכנית. הבעיה הטכנית הייתה שקיבלתי באיזורי החיתוך תגובה לא טובה.

ניסיתי לשנות אולי למסכה צריך פירמידה ללא טשטוש בין הרמות אלא רק דגימה. זה הביא לתוצאה טובה יותר אך לא מספקת – עדיין הייתה תגובה לא טובה.

בדוקומנטציה של OpenCV הייתה דוג', שהביאה לתוצאה גם לא טובה מבחינתי משום שהוסיפה בהירות לתמונה שאני לא רציתי. אפילו יצירה של פירמידת לפלסיאן והחזרה מייד לתמונה לא החזירה לתמונה המקורית.

מנבירה בדוקומנטציה הבנתי שהם משתמשים בפונקציה cv.substruct ו-cv.add במקום פשוט להשתמש בסימני פלוס ומינוס, אז ניסיתי להבין מה הפונקציות האלו עושות.

בהמשך המסע, גיליתי שהפונקציות הנ"ל מבצעות clipping לערכים להיות בטווח , ואז אמרתי לעצמי שזה לא טוב כי אני רוצה הפרשים והם לפעמים שליליים.

אז בדקתי ואכן גם במימוש שלי היה עדיין clipping בגלל הטיפוס של המטריצות שהוא טיפוס לא מכוון של 8 ביט (uint8), אז המרתי לטיפוס מכוון של 16 ביט (int16) וסוף סוף קיבלתי את התוצאה הרצויה.

אבל המשמעות היא – שבעצם אנחנו מכפילים את המקום שנדרש לתמונה להישמר פי שניים, כי ב-8 ביט אי אפשר לשמור את כל הטווח שאנחנו צריכים להפרשים – – בפחות משני בתים.

## הבדלים בין האלגוריתמים

בעוד בערבול, החיבור הוא רוחבי בין שתי הפירמידות – הווי אומר חיבור בין שתי הפירמידות בכל רמה, בתמונה ההיברידית החיבור הוא בין רמות.

# תוצאות

## ערבול תמונות

תמונה שמכילה פני אדם, ריס, גבה, שפתון

התיאור נוצר באופן אוטומטי

מובאות לעיל דוגמאות ביניים עד שהצלחתי להגיע לתוצאה הזו.

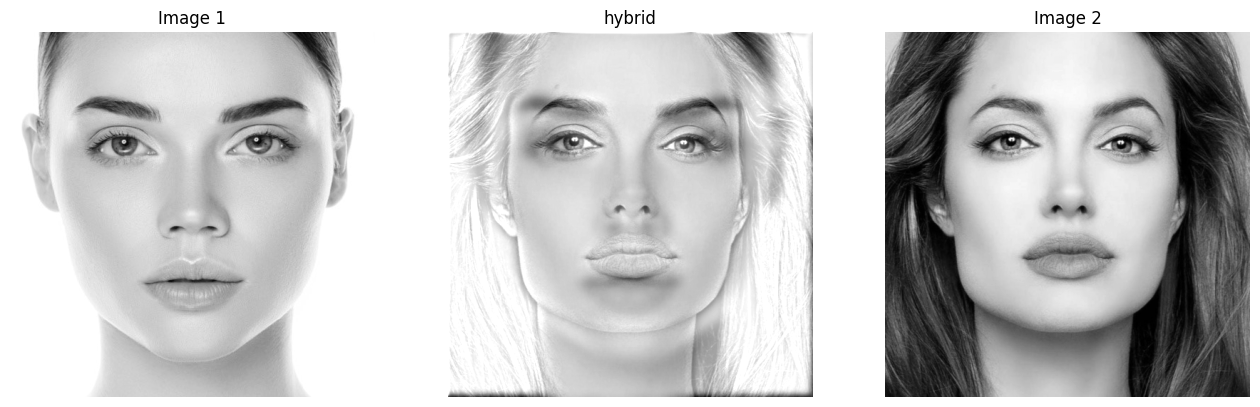
## תמונה היברידית

ניסיתי להשתמש באותן התמונות לעיל (עם פרמטרים 5, 5):

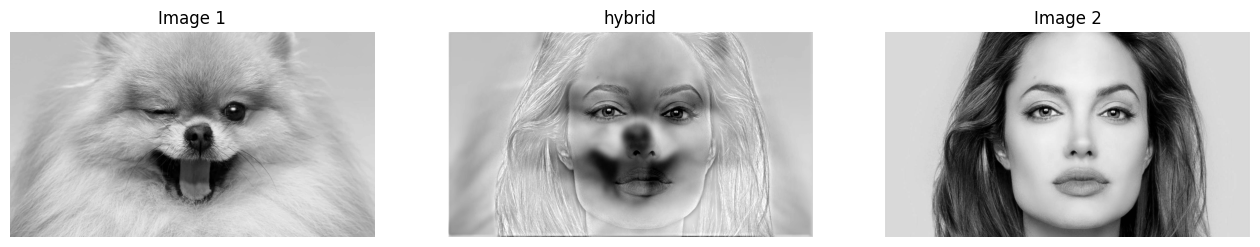
תמונה שמכילה פני אדם, גבה, אדם, ריס

התיאור נוצר באופן אוטומטי

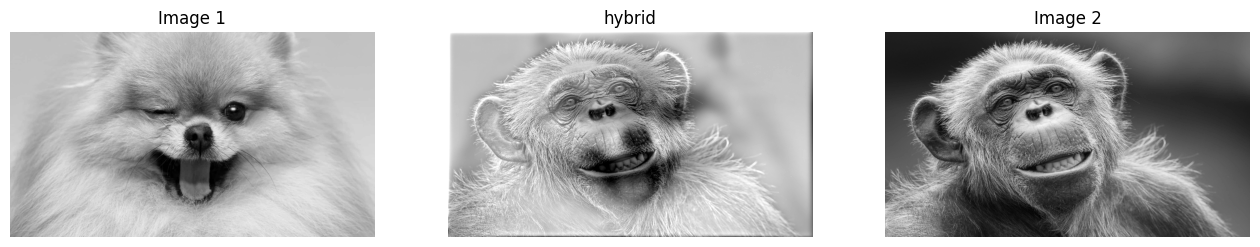
לא מושלם אבל קרוב. בכיוון השני:



גם לא אידיאלי. השיער מאוד מאוד משפיע ודומיננטי, והתמונות לא אידיאליות למשימה... נסיונות נוספים גם לא עלו יפה...

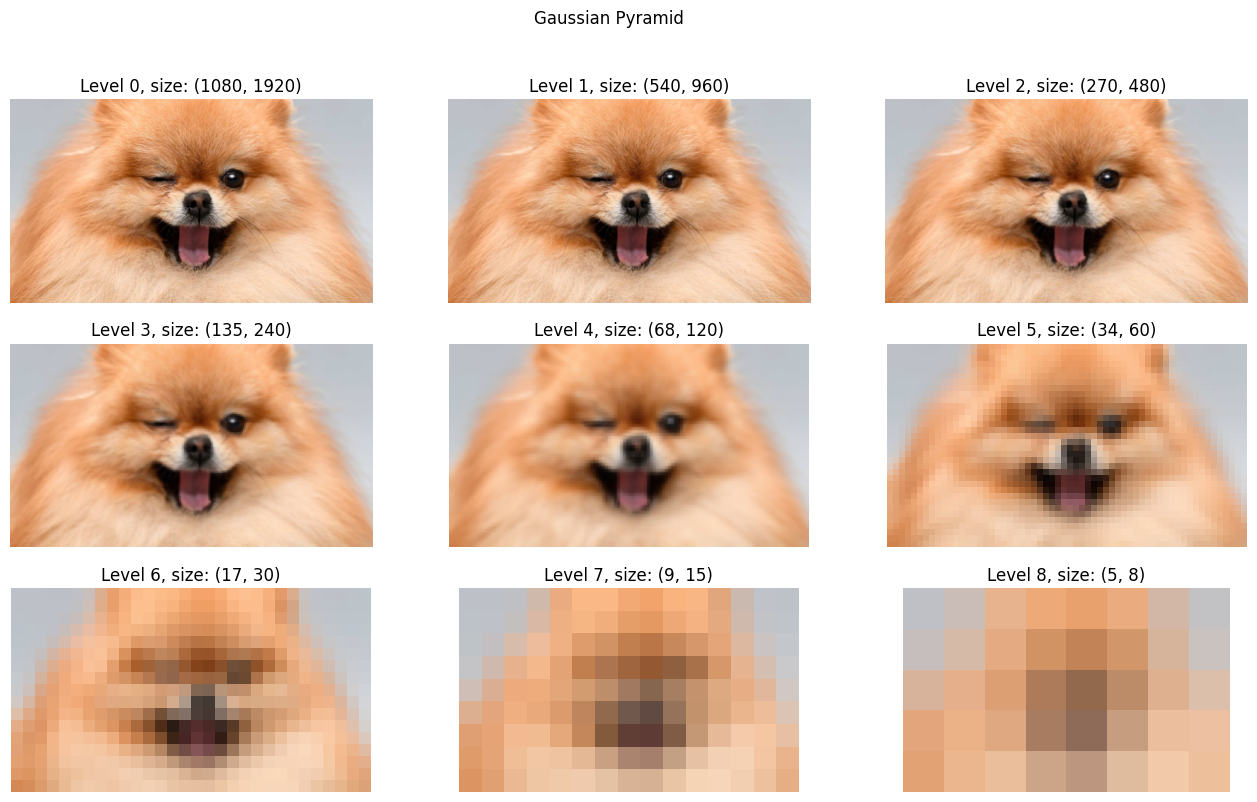
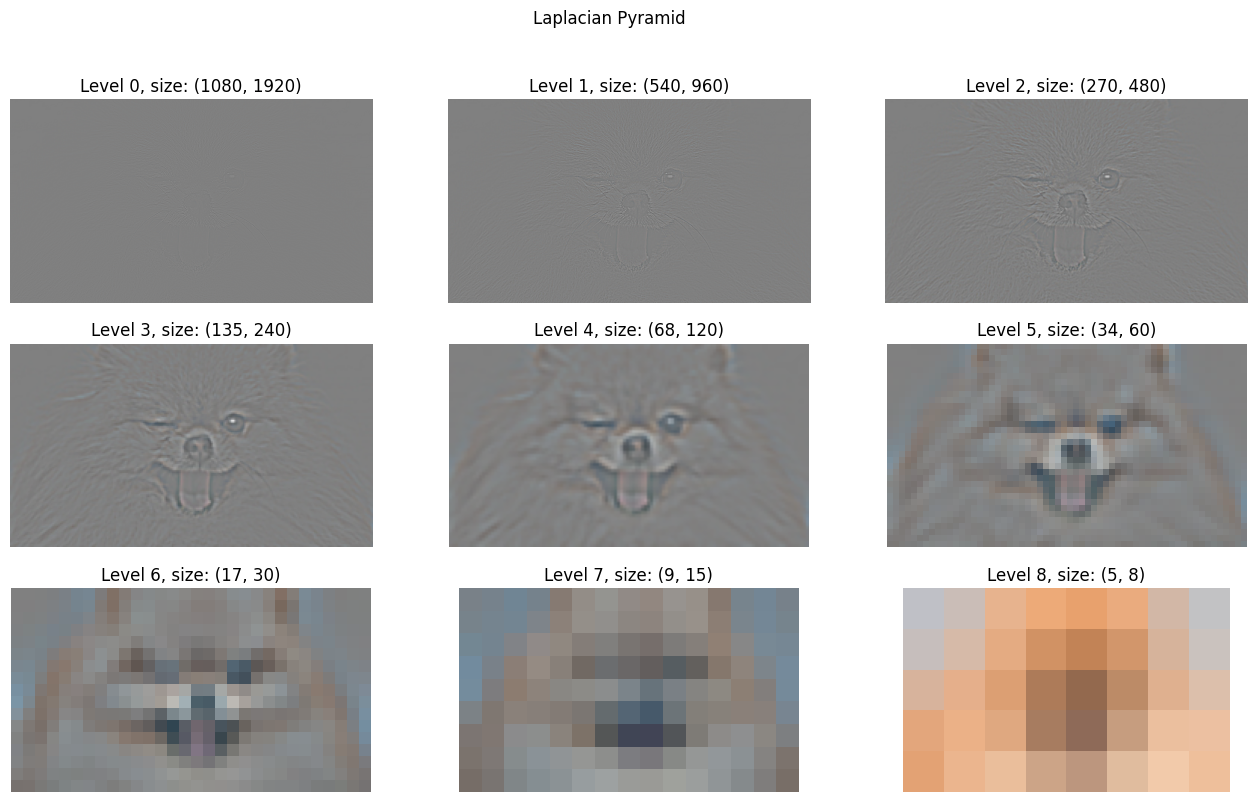


אבל:



כך שנראה לי חשוב גם לדעת את המגבלות של תמונה היברידית. זה לא באמת ייראה ממש אותו הדבר כמו אף אחת מהתמונות.

# פירמידות

אין לי הרבה מה להגיד על פירמידת הגאוסיאן.

מפירמידת הלפלסיאן ניתן לראות איך שהשכבות הראשונות כמעט ללא הבדלים. רלוונטי לצרכי כיווץ – לקחת רק את השינויים המשמעותיים. בנוסף, אפשר לראות שממש אפשר לעשות אומנות מהמשחקים האלו... נגיד רמה 3 היא יכולה להיות אפקט אומנותי ממוחשב ל-Office לתמונות.

# סיכום

הממצא העיקרי שלי מהעבודה הוא השימושים המעניינים של פירמידות. הם פילטרים נהדרים לתדרים וניתן לבצע מניפולציות מעניינות בקלות יחסית. בהרצאות ראינו עוד המון שימושים לפירמידות.

דבר נוסף הוא החשיבות של הטיפוסים והבנה של הצדדים הטכניים מעבר לצדדים המתמטיים והאלגוריתמיים בבואנו לפתח ולממש מוצר. ז"א – אע"פ שפייתון מנסה לצמצם את הפער ככל הניתן, כשאנחנו משתמשים ב-NumPY אנחנו חייבים להיות עם שליטה והבנה בסיסית ב-low level כדי לממש את האלגוריתם.

