Style

שאלות

1. איך קובעים את המשקלים?

סיכום:

1. להעתיק את התוכן הסמנטי בתמונה זו בעיה קשה, ומגבלה גדולה לכך הינה החוסר בייצוג תמונות.
2. במאמר זה, משתמשים ברשתות קונבולציה ששמשו לזיהוי עצמים בתמונות על מנת לחלץ פיצ'רים מרמה גבוהה, ועליהם להעתיק את הסטייל.
3. העברת סטייל מתמונה אחת לשניה יכולה להיחשב כבעית של העברת טקסטורות. קיימים מספר רב של אלגוריתמים לא פרמטרים המבצעים פעולה זו ע"י דגימה מחדש של הפיקסלים בהינתן טקסטורה. אלגוריתמים אלו משתמשים בפיצ'רים מרמה נמוכה.
4. באופן אידאלי, העברת סטייל תעשה שימוש בפי'צרים מרמה גבוהה כל שתוכל לחלץ מידע סמנטי.
5. הפרדה בין התוכן לסטייל עדיין בעיה קשה, אך רשתות קונבולציה עמוקות עזרו להתקדם בנושא.
6. הראו שרשתות קונבולציה עמוקות מאומנות על סט גדול מספיק, הצליחו לחלץ מידע מתקדם על הפיצ'רים. במאמר זה מתבססים על מנת לקחת אתת הפיצ'ירים ממימד גבוהה.
7. עקרונית, האלגוירתם המוצע הוא סוג של העברת טקסטורות, המגביל סינטזה של טקסטורה ע"י שימוש רק בפי'צרים של התמונה.
8. יוצרים תמונה חדשה ע"י חיפוש ייצוג מתאים במרחב הפיצ'רים. על גביהן מפעילים את האלגוריתם ?
9. במאמר עשו שימוש ברשת VGG מאומנת לביצוע לזיהוי אובייקטים ומיקומם. המודיפיקציה שעשו לרשת במאמר זה הינה נרמול שכבות הקונבולציה, כך שממוצע האקטיבציה של כל פילטר יהיה 1. פעולה זו לא משנה את מוצא הרשת כי המעברים לינארים ואין פולינג.
10. כל שכבה ברשת מגדירה למעשה פילוטר לא לינארי, שמורכבתו הולכת וגדלה ככל שמתקדמים. אפשר לייצג את תגובתה של שכבה מסויימת ע"י מטריצה.
11. נראה שהצליחו לשחזר טקסטורות ע"י שימוש ברעש לבן ו-GD. (זה מאוד דומה למה שאנחנו רוצים לעשות?)
12. הגדירו לוס ריבועי לתוכן על ההבדל בין מטריצות האקטיבציה של התמונה המקורית לשחזור, על כל שכבה, וכך ניתן לייצר תמונת תוכן חדשה.
13. במהלך ההיררכיה של הרשת, ישנה רגישות גבוה לתוכן התמונה, אך יש חוסר רגישות לדיוק ההופעה.
14. השכבות הגבוהות ברשת קולטות תוכן מרמה גבוה (כלומר, אובייקטים והסדר שלהם בתמונת המקור) אבל ללא הגבלה על הערכים של פיקסל מסויים. שימוש בשכבות מוכות דווקא יוצא תמונה דומה מבחינת ערכי פיקסלים.
15. על מנת לקבל ייצוג סטייל של תמונה, משתמשים במרחב פיצ'רים הנועד לזהות טקסטורה. מרחב זה יכול להיבנות על תגובות הפילטרים מכל שכבה ברשת ע"י מטריצת גרהאם. כאשראר מתחשבים בקורלציית הפיצ'רים ממספר שכבות, מקבלת ייצוג סטציונרי של תמונת המקור.
16. ע"י ביצוע GD למחיר הרביעוי בין מטריצות הגראהם של הפיצ'רים, ניתן למזער את ההבדלים בסטייל. יש שימוש בסכום משוקלל בין השכבות השונות כדי להגיע לתוצאה מיטבית.
17. על מנת להעביר את הסטייל מתמונה אחת לאחרת, אנו מסנטזים תמונה חדשה כך שתכיל את התוכן הרצוי והסטייל הרצוי, כאשר ממשקלים כל אחד מהם. מזעור המחיר נעשה באופן נומרי. במאמר, תמיד שינו את גודל תמונת הסטייל כך שיתאים לגודל תמונת התוכן.
18. הפיצ'רים מרמות נמוכות יכולים להוות prior לסטייל של התמונה. בנוסף שינויים בסינטזה יכולים להיווצר כתוצאה מאלגוריתם שונה.
19. אבן הייסוד לתוצאות המאמר: גילו שהייצוג של התוכן והסטייל הם פרידים. לכן ניתן לפתח כך את האלגוריתם – אבל ייצוג זה אינו נפרד לחלוטין, וישנו טרייד אוף בין המשקלים בפונקציית המחיר.
20. ייצוג הסטייל נעשה ע"י מספר שכבות שונות של הרשת. מיקומן ומספרן משפיע את ההשפעה של הגודל הלוקאלי של התאמת הסטייל. מצאו כי התאמת סטייל בשכבות גבוהות שומר על מבנים מקומיים בסקאלה גדולה, ומאפשר תמונה חלקה ורציפה יותר.
21. כאשר משווים את התוכן של שכבות גבוהות, המידע של הפיקסלים לא מוגבל, והטקסטורה של התמונות 'מתאחדת'.
22. במקום שימוש ברעש, ניתן לאתחל גם את הרשת באמצעות שימוש בתמונת התוכן \ תמונת הסטייל הרצויה. עם זאת, שימוש ברעש מאפשר ליצור כמות בלתי מוגבלת של תמונות, לעומת שימוש שתוצרי אותה תמונה, שגורם להתכנסות לתוצר מסויים.
23. המאפיין המגביל ביותר הוא רזולציית התמונה – היא משפיעה על ממד בעיית האופטימזציה וכן על גודל הקונבולציה.
24. תמונות שנוצרו מהרשת נתונות לרעש, המשפיע יותר על התוכן, או על התמונה המשולבת של הסטייל והתוכן. אבל למזלנו, נראה שלרעש יש תבנית מסויימת הנשענת על הפילטרים ברשת, וכנראה ניתן לנקות אותו באמצעות פילטר מתאים.