

תאריך: 19/04/2021

מסמך אפיון פרויקט

הפרויקט		
Generative Deep Features		שם הפרויקט:
6308		מס' ב-LabAdmin:
סמסטר אביב		סמסטר:
חד סמסטריאלי		חד/דו סמסטריאלי:
הצוות		
תמר רוט-שחם		שם המנחה:
מקצוע רישום: פרויקט מיוחד	דע-אל קלנג	שם סטודנט 1:
מקצוע רישום: פרויקט ב'	הילה מנור	שם סטודנט 2:
<input type="checkbox"/> חברה מלווה		
		שם החברה:
		שם איש קשר:

1. מטרת הפרויקט

מטרת הפרויקט הינה חקירת היכולת לייצר תמונה חדשה לחלוטין בסגנון החזותי של תמונה נתונה כלשהי, דרך למידה מונחית של רשת עמוקה ללא שימוש ב-Generative Adversarial Networks. הסגנון החזותי מוגדר לפי סטטיסטיקת המאפיינים העמוקים של התמונה הנתונה, והדמיון שלו ליעד ייבדק לפי פונקציות מטרה שונות. כדי לייצר תמונה ריאלית ולא רק תמונות טקסטורה, הפרויקט ישתמש בפרקטיקות שהוצגו במאמר SinGAN (Rott Shaham et al. 2019), כמו שימוש בפירמידת סקאלות, אך ללא שימוש ב-adversarial loss.

2. פירוט הנחות ודרישות

אנו מניחים, בעקבות מחקרים שונים שהוצגו בשנים האחרונות, כי בפיצ'רים העמוקים של רשתות קלסיפיקציה מאומנות טמון מידע המאפשר להבין את המהות הסמנטית של התמונה.

אנו מניחים גם כי מאחר ומקור הפיצ'רים הללו הוא ברשתות קונבולוציה, הרי שעדיין יש קשר לוקאלי בשכבות, ועבודה במסגרת פירמידת סקאלות תאפשר להבין את ההתפלגות הסטטיסטית ללא תלות מרחבית.

במידה והנחות אלו יתבררו כנכונות, ניתן להשתמש במאפיינים עמוקים אלו של תמונות למשימה של גנרציה של תמונות חדשות.

3. פתרונות אפשריים וסיכום קצר של סקר הספרות

בשנים האחרונות גובר השימוש בפיצ'רים שחולצו משכבות עמוקות ברשתות קלסיפיקציה מוכרות (VGG19/16) לטובת משימות של העברת סטיל [1][2][3] או שיפור תמונה. כשמנסים להעביר סגנון חזותי של תמונה, ההצעות מנסות לשמר גם דמיון פיקסלי של המוצא לתמונה המקורית, וגם דמיון לסגנון החזותי הרצוי באמצעות השוואה של הפיצ'רים העמוקים במוצא לפיצ'רים העמוקים של תמונת היעד מבחינת סגנון חזותי. הוצעו מספר שיטות להשוואת הדמיון בין הפיצ'רים העמוקים.

ההצעות החלו מהשוואה ישירה בין הפיצ'רים העמוקים של תמונת היעד (זו שתכווין את הסגנון החזותי) לתמונת המוצא (זו שרוצים לשנות לה את הסגנון), באמצעות מטריצת גראם [1].

הצעות נוספות ניסו להשוות בין זוגות הפיצ'רים העמוקים הכי קרובים של תמונת היעד ותמונת המוצא. בפועל, זהו קירוב של KL divergence בין התמונות [2].

אחת ההצעות האחרונות משתמשת בממד מרחק וואסרשטיין. מדד זה הוא רב-ממדי, אך ניתן לקרבו באמצעות שימוש בממד וואסרשטיין-חתוך, שתחת הנחות נאיביות של חוסר-תלות בין הפיצ'רים ניתן לקרב למדידת המרחק בין ההתפלגויות השוליות של הפיצ'רים [3].

בדרך-כלל העברת הסגנון התבצעה מתמונת מקור לסגנון של תמונת יעד כלשהי, וכאשר הוכנס רעש לבן למערכת נוצרה טקסטורה כלשהי בסגנון תמונת היעד, במקום תמונה טבעית חדשה. תופעה זו מוכרת ואף צפויה שכן השימוש ברשתות קונבולוציה כופה קשרים מרחביים. אחת הדרכים להתמודד עם תופעה זו היא שימוש בפירמידת סקאלות [4]. כפי שהוצג במאמר, כל סקאלה אחראית על אזור הולך וקטן בתמונה, ובכך ניתן להשיג "טקסטורות" במגוון מימדים, שמאפשרים יצירת תמונה שלמה טבעית. באמצעות שימוש בפירמידת סקאלות ייתכן וניתן לייצר תמונה חדשה לחלוטין.

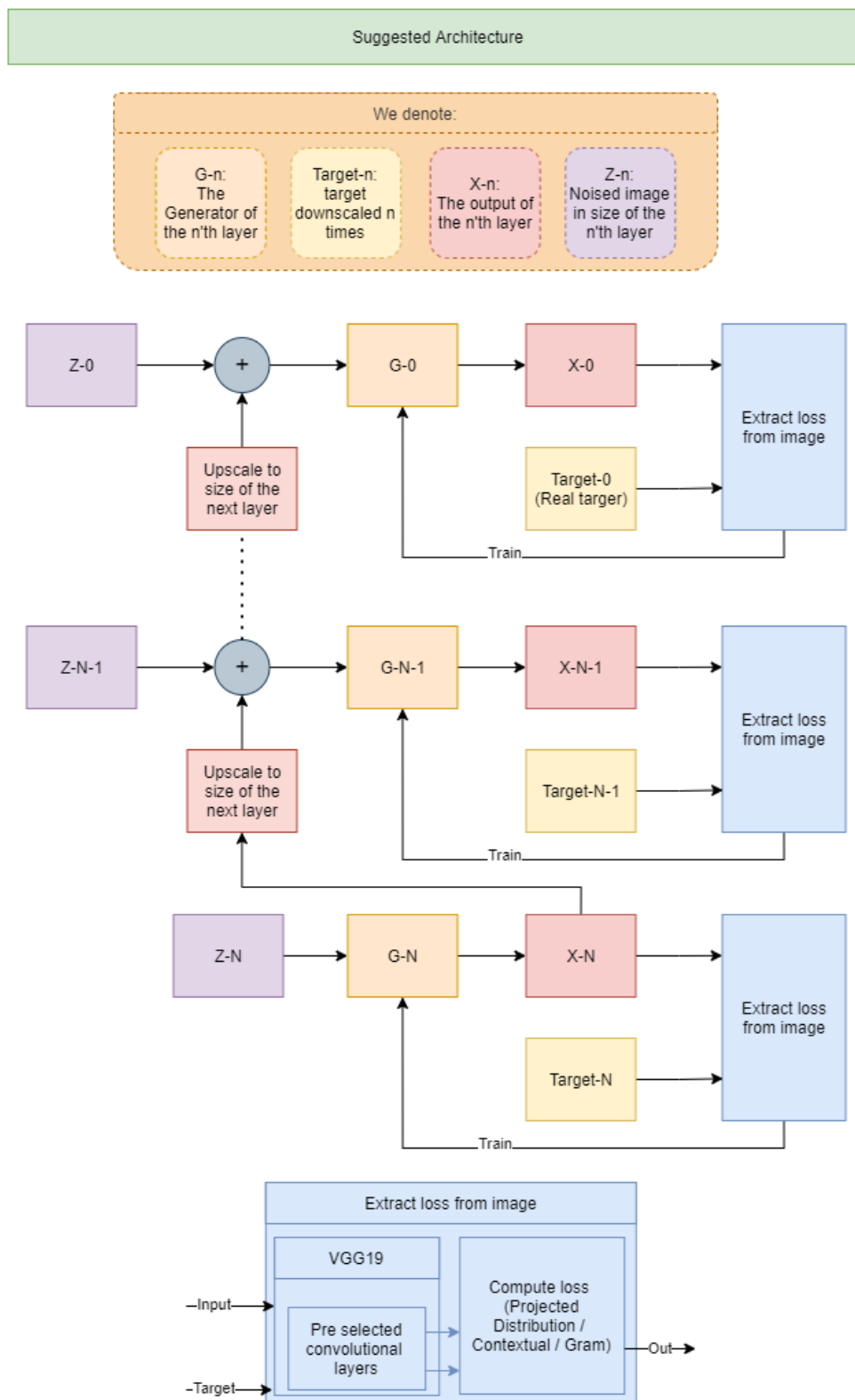
[1] L. A. Gatys, et al. "Image Style Transfer Using Convolutional Neural Networks", IEEE conference on computer vision and pattern recognition, 2016 pp. 2414-2423

[2] R. Mechrez, et al. "The Contextual Loss for Image Transformation with Non-Aligned Data", European Conference on Computer Vision (ECCV), 2018, pp. 768-783

[3] M. Delbracio, et al. "Projected Distribution Loss for Image Enhancement", arXiv preprint arXiv:2012.09289, 2020

[4] T. R. Shaham, et al. "SinGAN: Learning a Generative Model from a Single Natural Image", IEEE International Conference on Computer vision (ICCV), pp. 4570-4580, 2019.

4. תרשים מלבנים (block diagram) של הפתרון הנבחר או הנבדק



5. מודולים שנידרש לפתח

בפרויקט זה ייתכן ונצטרך לפתח את מבנה הרשת, ולמצוא את הפרמטרים הטובים ביותר עבורה. בנוסף, נצטרך לממש את פונקציות המחיר הנדרשות.

6. מודולים מוכנים שניתן להיעזר בהם

נוכל להתבסס על הגנרטור מתוך sinGAN [4], וכן בפונקציות מחיר שפותחו בעבר.

7. סביבת עבודה וכלי פיתוח שיהיו בשימוש

★ פייתון

Pytorch ○

Torchvision ○

Matplotlib/Seaborn ○

★ רשת VGG-19 מאומנת מראש

★ מחשבים בעלי GPU חזק להרצות

8. שיטת הבדיקה שתידרש בסיום הפרויקט

1. נוכל להשוות את זמני האימון והמשאבים הנדרשים בין אלגוריתמים דומים.

2. ניתן לבצע השוואה בין ביצועי המודל המוצע שלנו לבין רשתות גנרטיביות דומות.

a. באמצעות מדדים כמותיים בסגנון SIFID [4].

b. באמצעות התבוננות עצמאית [1]:

In our work we consider style transfer to be successful if the generated image 'looks like' the style image but shows the objects and scenery of the content image. We are fully aware though that this evaluation criterion is neither mathematically precise nor universally agreed upon.

~ Gatys et al., 2016[1]

9. רשימת משימות:

מס'	שם המשימה	תיאור המשימה	משך ביצוע משוער
1	סקר ספרות	קריאת מאמרים על פונקציות מחיר שונות, העתקת סגנון ו-SinGan	שבועיים
2	התנסות בספריות מוכנות של GAN והעברת סטייל	ביצוע ניסויים עם מימושים קיימים ל-style transfer DC-GAN כדי לנסות להבין את השפעת השכבות השונות של הפיצ'רים והיחסים ביניהם	שבוע
3	מימוש פונקציות מחיר נוספות להעתקת סגנון	מימוש פונקציות מחיר Contextual Loss ו-Projected Distribution Loss - העתקת סגנון בלבד מתוך תמונת מקור לתוך תמונת רעש.	שבוע
4	התאמת ארכיטקטורת sinGAN לסקאלה בודדת	הטמעת רשת הקובולוציה של הגנרטור מתוך SinGAN	שבועיים
5	אימון הסקאלה הבודדת תחת פונקציות מחיר שונות	אימון הארכיטקטורה החדשה באמצעות פונקציות המחיר PDL/CXL/Gram עבור סקאלה בודדת, והבנת השפעת כל פונקציות מחיר	שבועיים
6	אימון multi-scale	ממימוש גנרטורים ואימונים במבנה פירמדה	שבועיים
7	אימון multi-scale תחת פונקציות מחיר שונות	הבנת השפעת כל פונקציית מחיר במבנה הפירמידלי, וכוונון היפר פרמטרים לתוצאות מיטביות	שלושה שבועות
8			
9			
10			

10. תרשים גאנט (התקדמות הפרויקט):

שם משימה	מספר חודשים מתחילת הסמסטר								
	1	2	3	4	5	6	7	8	9
סקר ספרות	X			הקופת מבחנים					
התנסות בספריות מוכנות של GAN והעברת סטיל	X								
מימוש פונקציות מחיר נוספות להעתקת סגנון	X								
התאמת ארכיטקטורת sinGAN לסקאלה בודדת		X							
אימון הסקאלה הבודדת תחת פונקציות מחיר שונות		X							
אימון multi-scale			X						
אימון ה-multi-scale תחת פונקציות מחיר שונות			X						