SINGAN

1. הקשר בין patch size לn
2. איך מאמנים – במיוחד שלב 0. כמה איטרציות? מפעפעים למעלה בפירמדה כל הזמן?
3. Markovian discriminator? למדנו? זה משנה?
4. כשמשנים גודל של תוצר – איך זה קורה? מה צריך לשנות ברשת? רק את הרעש? סקייל מתאים לתמונת מקור? אם משנים רשת, לא דורש אימון מחדש? דקה 6:30

שאלות:

1. לשאול על SFID
2. שינוי גודל תמונה – איך מתבצע? ללא אימון חוזר? התמודדות עם מריחה? מגבלות?
3. למה batchnorm לא גם בסוף? Tanh מספיק? למה אין batchnorm בD הראשון

סיכומון:

1. sinGan – ללא התנייה, ויודע לדמות מאפיינים של התמונה בעזרת פירמדת של GAN קונבולוציונים. מאפשר לשחזר מבנים גדולים וכן פרטים עדינים.
2. GAN עזרו למדל פילוגים של מידע חזותי, אך לרוב היה צורך בהתנייה על מנת להצליח במשימה. sinGan מצליח לפתור את רוב הבעיות ללא התנייה.
3. תמונה אחת נושאת מספיק מידע כדי לבצע ניתוח, ולכן בעזרת פירמדת של GAN קלים ניתן לבצע ניתוח.
4. sinGan טובה בהעלאת רזולציה וכו' ומשמרת נתונים פנימיים של התמונה
5. בעבודות קודמות לא יצרו תמונות חדשות לחלוטין, אלא בעיקר השוו תמונה לתמונה.הסט אפ שלנו בוא גנרטיבי טהור, ואינו מוגבל לטקסטורה בלבד.
6. לgan יש שדה ראייה קטן, וכך ניתן להימנע מזיכרון של כל התמונה (התאמת יתר).
7. Singan בנוי בצורת פירמדה – כאשר בשכבה הכי נמוכה N יש את הרזולוציה הכי נמוכה, ושם מתבצעת גנרציה של תמונה אקראית. משם זה מפועפע למעלה בכל פעם עם התמונה בDS מתאים, כאשר עושים US לתוצר מהשכבה הקודמת.
8. שדה הראייה בשכבה הכי גסה (N) הוא בערך חצי מגובה התמונה, וכך לומדים להתייחס למאפיינים גדולים.
9. הLOSS מוגדר כשגיאה מה-adv של ה-GAN וגם משגיאת השחזור. (יש הרבה נקודות שלא הבנו עד הסוף – לחכות להבהרות – פרק 2.2)
10. שגיאת החשזור מוגדרת ביחס לרעש בשכבה ה-N עבורו הגענו לשחזור טוב כלשהו, ומשאירים אותו קבוע. מפעפעים את השגיאה למעלה.
11. שונות הרעש פרופורציונלית לRMSE.
12. השכבה התחתונה תמיד נקבעה לגודל מינימלי של 25 פיקסלים, ובגל שלב הקטנו ביחס של 4/3. שינו את גודל תמונת האימון ל-250 פיקסלים.
13. השיטה מאפשרת לשחזר מאפיינים איכותיים של התמונה שהתאמנו עליה בגנרציה רנדומלית. גם כללים וגם קטנים ועדינים.
14. ככל שמזריקים בשכבה גסה יותר, ההשפעה גדולה יותר על מאפיינים גלובליים ולהפך.
15. ככל שיש פחות שינויי גודל, אנחנו רגישים רק לשינויים קטנים.
16. מדד SFID – מדד חדש שהגדירו כדי למדוד FID לבאטצ' ספציפי שיתאים לSINGAN

תשובות תמר:

1. על הoverfit- ברשת סיווג למידה על תמונה אחת יכולה להיות הרבה יותר מאתגרת. בשימושים אחרים כן למדו על תמונה אחת. בסופר רזולוציה הם מאמנים על זוג תמונות (באיכות נמוכה וגבוהה) בתקווה שהוא יצליח גם בסט אימון. הOF נשמע רצוי, כי אנחנו רוצים ללמוד את הפילוג הספציפי יותר.  
   אנחנו כן מנסים ללמוד ולבחון את התמונה (?)
   1. אצלנו כנראה שגם נזין רעש. אצלנו כנראה נתחיל בתמונת אפסים וכו'. המטרה הסופית לגנרט. צעד קודם זה לפתור זו בעיית האופטימיזציה על הפיקסלים של התמונה, כדי להבין אם הLOSS עושים את מה שאנחנו רוצים, הרובד של הרשת מוסיף עוד מורכבות.
2. חיבור במוצא הרשת – הסיבה שמוסיפים את התמונה הפריטים מהשלב הקודם הוא כדי להקל על הרשת. לכן, רשת הקונבולציה מאומנת על הוספת פרטים בלבד, ולא לייצר תמונה. אנחנו מקבלים תמונה קצת מורעשת, מלמדים את הרשת איפה הפרטים המעניינים בה, ומוסיפים את התמונה הקודמת. לכן החיבור עם התמונה ממוקדם מאלץ את הרשת ללמוד רק איך להוסיף פרטים.  
   אם מסתכלים רק על מוצא הרשת, זה לא נראה כמו תמונה, אלא כמו םפרטים כמו גבולות וכו'.
3. מבנה רשת – כל קוד בלוק מכיל קרנל ב- 3X3 \* 5 בלוקים, לכן התמך הכללי הוא 32. (כי קונבולציה על קונבולציה).
4. Markovian discriminator – הD מוציא פלט לפי חלקים נפרדים בתוך התמונה, והציון הוא לפ חלקים בתמונה. לכן הDATASET שלנו הופך להיות הפאטצ'ים של תמונה אחת.
5. Boundary condition - קונבולציה 'אוכלת' את השפות של התמונה. אנחנו לא רוצים שזה יקרה. לכן כל שכבת קונבולציה נהוג לרפד באפסים. לכן מוסיפים רעש מרופד לתמונה במרכז.
6. אפסים ברעש – הסקאלה הראשונה היא מגנרטת ממש, ושאר הסקלות רק מוסיפות פרטים. לכן בסקאלה הראשונה אנו חייבים להכניס רעש (או שיהיה קשה מאוד אחרת). בשאר הסקלות, רצינו לשלוט רק בשוני בין התמונות. לכן כאשר רצינו לשחזר, העדפנו לשמור את המאפיינים של התמונה. רק ברקנסטרקציה מקבעים אפסים, בADV זה לא אפסים.
7. הסיבה ל4/3 היא נטו אמפירית.
8. להראות תמונה לשניה 1 – זה פרוטוקול ידוע ומקובל. תמר עושה מחקר על על זמני זיהוי וכו'.
9. FID – לא צריך להיכנס לזה לעומק, רק אם נרצה להיכנס לזה. בגדול, בGAN יש אוסף תמונות אמיתיות ואוסף מזויפות. לכאורה נרצה לבדוק את הפילוג של התמונות ולהשוות. בגלל שהמימד מאוד גבוה, אז לפני שהופכים את זה להחלטה סופית soft—max, אז מסתכלים על הווקטורים האלו, ולומדים את הפילוג עליהן בעזרת mean+var. נוסחה פשוטה יחסית.
10. סופר רזולציה – על מה חוזרים K פעמים כדי לקבל סופר רזולציה – כשיש singGAN מאומן, אז משתמשים רק בשכבה הכי עליונה. לכן מגדילים את התמונה G0 וומכניסים שוב לרשת עד שמגיעים לגודל הרצוי. בגלל שיש את פקטור r, הרשת יודעת להתמודד רק עם שינויים בגודל כזה. לכן נאמן גם עם r שמתאים לסופר רזולוציה שרוצים. S זה פקטור ההגדלה הסופי. בגלל שזו רשת קונבולציה, אז אין בעיה עם תמונות גדולות יותר.