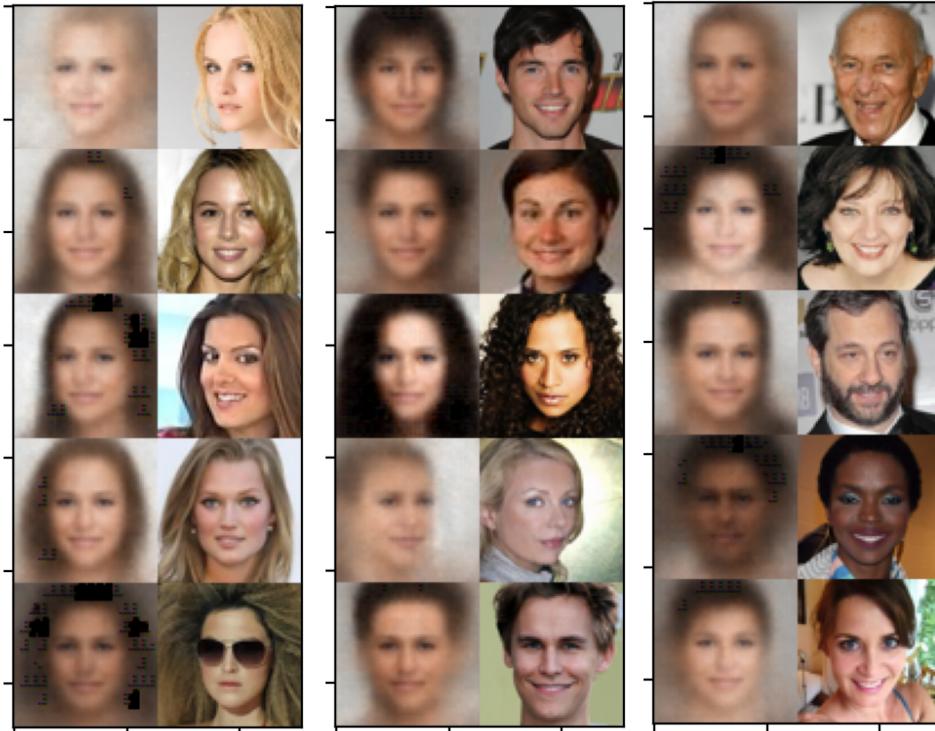


Non-Adversarial Generative Networks

1. בשאלת הראשונה בנוינו אוטואנקיודר לפי הארכיטקטורה הבאה:

האנקודר הכליל 3 שכבות קונבולוציה ולאחריה 2 שכבות FC, כאשר אחרי כל שכבה קונבולוציה הפעלה batch normalization, ואחרי כל שכבה באופן כללי הפעלה פונקציה אקטיבציה ReLU, הדיקודר הכליל בדוק את אותן שכבות רק בסדר הפוך באופן סימטרי, אימנו את המודל במשך 63 איפוקים.

אחרי האימון רצינו לבדוק עד כמה המודל מצליח לשחזר את התמונות המקוריות, התוצאות נראהות כר:



אפשר לראות שהמודל מצליח לשחזר צורות וצבעים כלליים, למשל במקבץ הימני בתמונה הרביעית, ניתן להבין מהשחזר שמדובר בדמות בעלת צבע עור שחור. בנוסף לצבע, המודל למד גם גם צורות - למשל שיער ארוך או קצר. דברים שהוא לא למד לשחזר זה למשל משקפי שמש (בתמונה התוחתונה משמאל) או זקן (בתמונה השלישית במקבץ הימני) וזה הגיוני כי אלה מאפיינים ייחסית נידירים.

השלב הבא היה לנקח שני באז"ם של תמונות מקוריות, לכל אחד מהם לחשב וקטור ב-latent space, לבצע אינטראפלציה בין שני הוקטוריים האלה ולראות את התוצאות של שחזר השילוב ביניהם:



הטור ה- i - בתמונה הקודמת מציג את השחזור של הווקטור: $z_2 \cdot \frac{1}{10} + z_1 \cdot \frac{1}{10} \cdot i$.
 עבור $[0 \dots 10] \in i$ כאשר z_1 זה הווקטור ב-*latent space* עבור באז' אחד של תמונות, ו- z_2 זה הווקטור עבור באז' נוסף.
 אכן ניתן לראות מעבר הדרגתני בפיצרים בתמונות, למשל בתמונה התחרתונה ובתמונה השלישית יש מעבר הדרגתני משיער כהה ואורח לשיער בהיר וקצר. (משמאלו לימין)

השלב האחרון היה ליצור תמונות חדשות לגמרי. המודל שאמנו בשלב הקודם העביר את התמונות ל-*latent space* לא ידוע, ולכן אם נקח את הדיקודר של המודל זהה ונתן לו וקטורים רנדומליים, או וקטורים מהתפלגות שאנו חיליט - זה לאו דווקא יהיה ה-*latent space* שהוא למד, لكن הוא לא בהכרח ידע ליצור ממה תמונות הגייניות. לכן אימנו מודל נוסף, שבו בנוסף למצער הטעות בין התמונות המקוריות לתמונות המשוחזרות, הוספנו גם מצער של המרחק בין הווקטור שהתקבל ב-*latent space* לבין וקטורי שנדגנו מהתפלגות ידועה שאנו קבענו ($\text{Normal with mean}=0, \text{variance}=2$, $\text{kurtosis}=3$) פועלה זו כופה על ה-*latent space* שנלמד להיות לפי הרתפלגות שאנו רוצים.

המודל החדש זהה במבנה למודל הקודם, אימנו אותו במשך אותו מספר איפוקים וכשאימנו דגמנו וקטור מהתפלגות שלימדנו אותו, ובדקנו איך נראהות התמונות שייצאו מהפעלת הדיקודר על הווקטור זהה, אלה התוצאות:

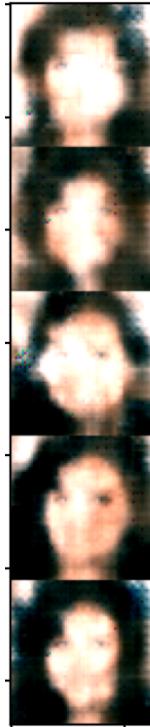


ניתן לראות שאמן איקות הפרטים בפרצופים נמוכה מזו שנוצרו בשחזור של תמונות אמיתיות, אך עדין ניתן להבחין באזור כללית של פנים ובתווי פנים גנריים כמו עיניים אף פה ושיער, ובפרט קיימות שונות בגובה בין התמונות השונות, מה שהגינוי מבחינה תיאורטית אחריה וה-*latent space* הנלמד הוא בעצם תת המרחב ששמור את השונות הגדולה ביותר.
 בנוסף רואים שאם אנחנו נתונים למודל, וקטור שהוא מאד רחוק מהתפלגות הזו, למשל כזה שנדגן מהתפלגות נורמלית עם תוחלת 100 הוא מייצר תמונות בהן כל הפרצופים נראים אותו דבר - ז"א שונות מאוד נמוכה.

התמונות שיצר המודל מוקטור שנדגם מההפלגות נורמלית עם תוחלת 00:



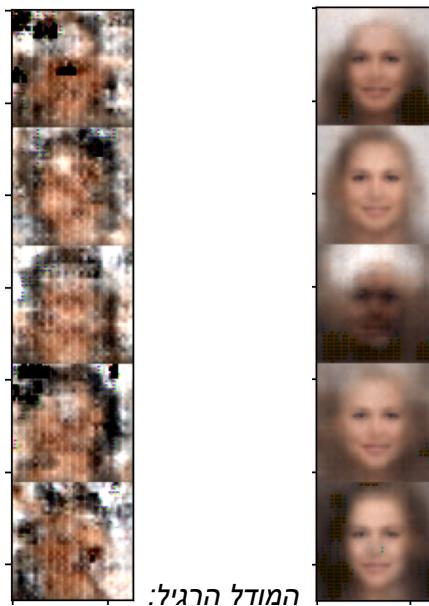
אמנם יש למידה של פנים, אך אין שונות, זו כל הפנים נראים בערך אותו דבר, - הדבר הגיוני מאחר ומדובר וקטור לא מה- *latent space* שנלמד ע"י המודל. תופעה מעניינת נוספת היא שכאשר נותנים למודל וקטור שנדגם מההפלגות עם תוחלת שלילית, נגד 00- הוא נצליח לחלווטין בשחזר מבנה של פנים (עיניים פה אף וכו...):



בנוסף, רצינו לראות שמודל שלא אומן על ייצרת ה- *latent space* זהה - ייתן תוצאות פחות טובות, אך להפתענו המודל הרגיל נתן תוצאות דומות:



לדעתי, הסיבה שיכולה להיות לכך היא שכשמדובר ברגיל התאמן, הוא יצר באופן טבעי סביב ה-0, ולכן התוצאות שהוא נותן הן יחסית טובות. כדי לבדוק את ההשערה שלנו אימנו מודל נוסף, הפעם כך שה- *latent space* שהוא לומד יהיה מההתפלגות הנורמלית עם תוחלת 3-. השווינו את התוצאות שהתקבלו בין המודל שהתאמן על ה- *latent space* זהה, לבין המודל הרגיל שלא אומן על שום *latent space* וקיבלנו:



המודל שהתאמן על *latent space*:

ניתן לראות שאשר אנחנו מייצרים פרצופים מוקטורים שנדגם מההתפלגות של ה-*latent space*, המודל הרגיל מייצר תמונות משובשות לגמרי והמודל שאומן ל- *latent space* זהה מצליח לייצר תמונות הגיוניות של פנים.