

Deep Learning- Assignment 4

GAN\C-GAN

רקע מקדים :

בעבודה זו התעסקנו עם רשתות GAN נדרשנו לממש בחלק א' את הרשת במימוש הרגיל, לקרוא.

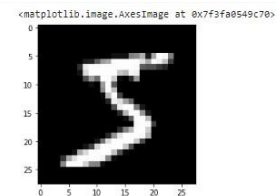
מטרתנו היתה שה-Generator ילמד לייצר Fake Images כך שה-Discriminator יתקשה בזיהוי תמונה אמיתית למזויפת.

בשתי החלקים השתמשנו בסט הנתונים של MNIST, מכיוון שלא היינו צריכים לפצל ל- train\test השתמשנו בדאטה כולו לצורך אימון.

הדאטה מכיל : 60,000 תמונות

```
len(mnist_train)
60000
```

כאשר כל תמונה נראית באופן הבא:



```
mnist_train[0][0].shape
torch.Size([1, 28, 28])
```

וכל תמונה מגיעה ב-tuple כאשר במקום ה-0 המערך המייצג תמונה :

```
mnist_train[0][1]
5
```

ובמקום ה-0 הלייבל של אותה התמונה:

בנוסף אתחלנו את המשתנים אשר שימשו בשתי החלקים באופן הבא –

Num Of Epochs= 150

Batch_size= 256

אופן בחירת ה-Loss Function בשתי חלקי העבודה -

השתמשנו בפונקציית ההפסד הבינארית BCEWithLogitsLoss , מכיוון שהתייחסות לתמונה אמיתית\מזויפת בעלת שתי אופציות בלבד כאשר אנחנו החלטנו ש-1 יסמן תמונה אמיתית, ו-0 תמונה מזויפת.

כאשר אנו מחשבים Loss ל-Discriminator זה מתגמל את התמונות באופן המוצג מעלה, והמטרה היא להקטין את ההפסד בכל פעם, משמע לזהות באופן יותר טוב בין אמיתית למזויף.

אך כאשר אנו מחשבים Loss ל-Generator, דבר זה עובד בצורה הפוכה . מכיוון שאנו רוצים לייצר תמונות מזויפות בצורה שיהיה קשה ל-Discriminator לזהות אנו משווים את הערך אשר יצא רק עם הערך 1.

חלק 1 : Vanilla GAN

בחלק זה מימשנו את רשת ה-GAN הרגילה אשר נקראת גם Vanilla GAN .

ארכיטקטורת הרשת נעשתה בדומה למה שלמדנו בתרגול, באופן הבא:

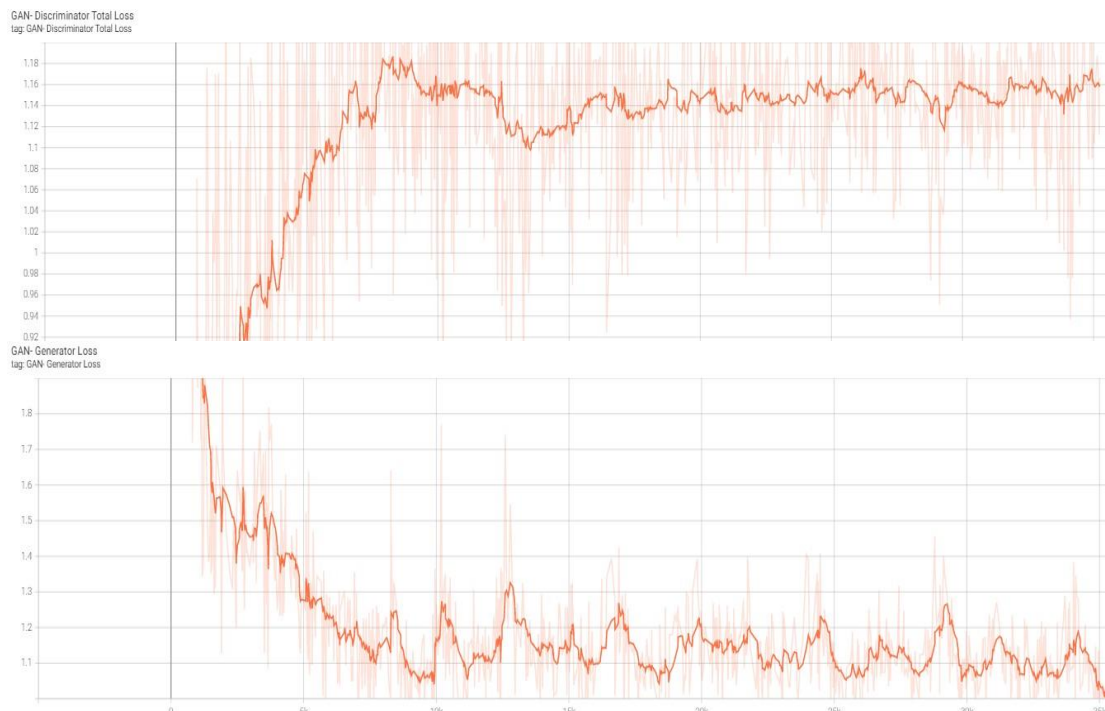
Generator:

- Conv(in=256 (batch), out=128) - Kernel=5, stride=1, paddin=2
- Betch Norm
- Leaky Relu
- Conv(in=128, out=64)- Kernel=5, stride=2, paddin=2
- Betch Norm
- LeakyRelu
- Conv(in=64, out=1)- Kernel=5, stride=2, paddin=2
- Tanh

Discriminator :

- Conv(in=1, out=64)- Kernel=5, stride=2, paddin=2
- Leaky Relu
- Dropout
- Conv(in=64, out=128)- Kernel=5, stride=2, paddin=5
- Leaky Relu
- Dropout
- Flatten
- Linear(10*10*128,1)

גרפי ה- Loss :



מהגרפים אנו רואים מה שציפינו לו, שה-Loss של ה-Discriminator עלה מכיוון שהיה לו קשה בזיהוי ומנגד מי שהקשה עליו ויצר תמונות ברמה טובה יותר זה ה-Generator אשר ה-Loss שלו ירד.

התמונות שיצאו מה- Generator :

הצגת התמונות נעשתה בשערך מחזור של 40 Epochs מלבד ההתחלה והסוף בו נעשה במרווח צפוף יותר על מנת להראות את השינויים באופן טוב יותר.

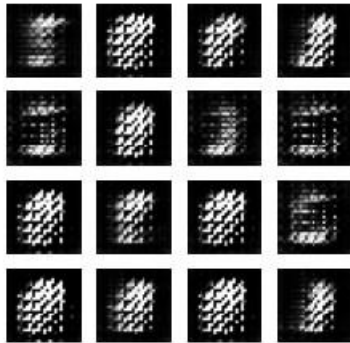
GAN- Generated Samples

Jan26_13-02-44_af8300b1cf76GAN_DEEP_4_HW

tag: GAN- Generated Samples

step 1

Thu Jan 26 2023 16:09:39 GMT+0200 (Israel Standard Time)



GAN- Generated Samples

Jan30_22-18-53_3a82ba50aca5GAN_DEEP_4_HW

tag: GAN- Generated Samples

step 20

Tue Jan 31 2023 00:26:44 GMT+0200 (Israel Standard Time)



GAN- Generated Samples

Jan30_22-18-53_3a82ba50aca5GAN_DEEP_4_HW

tag: GAN- Generated Samples

step 42

Tue Jan 31 2023 00:34:35 GMT+0200 (Israel Standard Time)



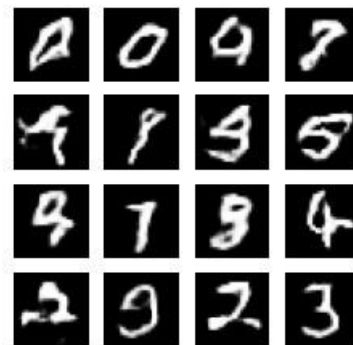
GAN- Generated Samples

Jan30_22-18-53_3a82ba50aca5GAN_DEEP_4_HW

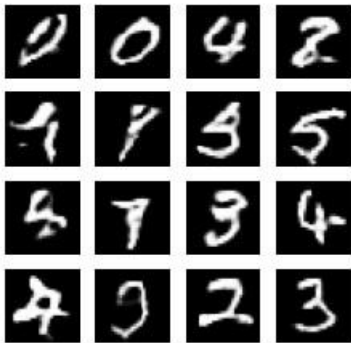
tag: GAN- Generated Samples

step 81

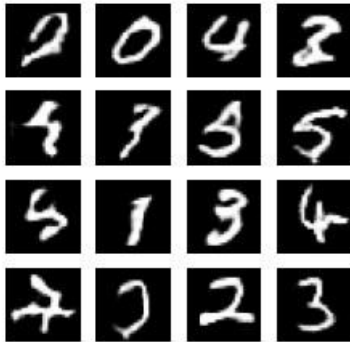
Tue Jan 31 2023 00:48:26 GMT+0200 (Israel Standard Time)



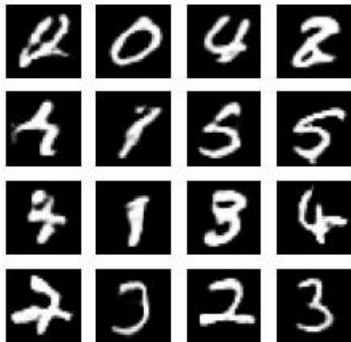
GAN- Generated Samples
Jan30_22-18-53_3a82ba50aca5GAN_DEEP_4_HW
tag: GAN- Generated Samples
step 120
Tue Jan 31 2023 01:02:08 GMT+0200 (Israel Standard Time)



GAN- Generated Samples
Jan30_22-18-53_3a82ba50aca5GAN_DEEP_4_HW
tag: GAN- Generated Samples
step 129
Tue Jan 31 2023 01:05:17 GMT+0200 (Israel Standard Time)



GAN- Generated Samples
Jan30_22-18-53_3a82ba50aca5GAN_DEEP_4_HW
tag: GAN- Generated Samples
step 149
Tue Jan 31 2023 01:12:20 GMT+0200 (Israel Standard Time)



מסקנות מרשת ה-GAN :

- על פי האיטרציות ניתן לראות כי המודל מתחיל באיטרציה הראשונה בתמונות עם רעש רנדומלי אשר מתמקד במרכז התמונה .
באיטרציה ה-20 אנו כבר יכולים להבחין בניצנים של ספרות כמו 0,4, הרעש הופך למשהו מרוכז וברור יותר ובחלקם אפילו כבר ספרה. באיטרציה ה-42 כבר ממש ניתן לראות תמונות חלקות יותר כאשר במרכזם ספרות ומופיעות ספרות כמו 2,3 ו-9 .
בין איטרציות 82-42 המודל מתחיל להתייצב .
כאשר מגיעים לאיטרציה ה-120 כבר המודל התאמן מספיק בכדי לזהות את רוב הלייבלים שקיימים בצורה טובה.
- אפשר להבחין כי המודל לא מייצר תמונות של כלל הספרות, זאת בעקבות כך שאנו לא מציינים לו מידע נוסף או התפלגות מסוימת ממנה צריך לייצר את הסמפלים.
הספרות אותן המודל מייצר באופן טוב – 0,1,2,3,4,5,7.
- הספרות אותן המודל לא מייצר או מייצר אך לא באיכות כזו שניתן לזהות – 6,8,9.
נקודה מעניינת היא שהמודל התחיל מסיווג עיגולים וקווים בודדים ולכן ניתן לראות את הספרות 0,1 שנראה שדבר זה נובע כי אלה הצורות הקלות לייצור בהתחלה.
לאחר מכן ה-0 מתחיל להפוך לספרות 8,9 וה-1 ל-3 או 7.

C-GAN : 2 חלק

בחלק זה מימשנו את רשת ה-C-GAN על פי מאמר אותו התבקשנו לקרוא. מה ששונה במימוש רשת זו מהרשת הרגילה הוא שהקלט של ה-Discriminator וה-Generator משלב גם את התיוגים של כל תמונה. וכאן אנו מייצרים דוגמא מזויפת עבור כל סיפרה בנפרד, כאשר ה-Generator מטרתו ליצר תמונות של הספרות בהתאם לתיוג באופן שה-Discriminator יתקשה בהבחנת תמונה אמיתית\מזויפת.

ארכיטקטורת הרשת נעשתה בדומה למה שלמדנו בתרגול, באופן הבא:

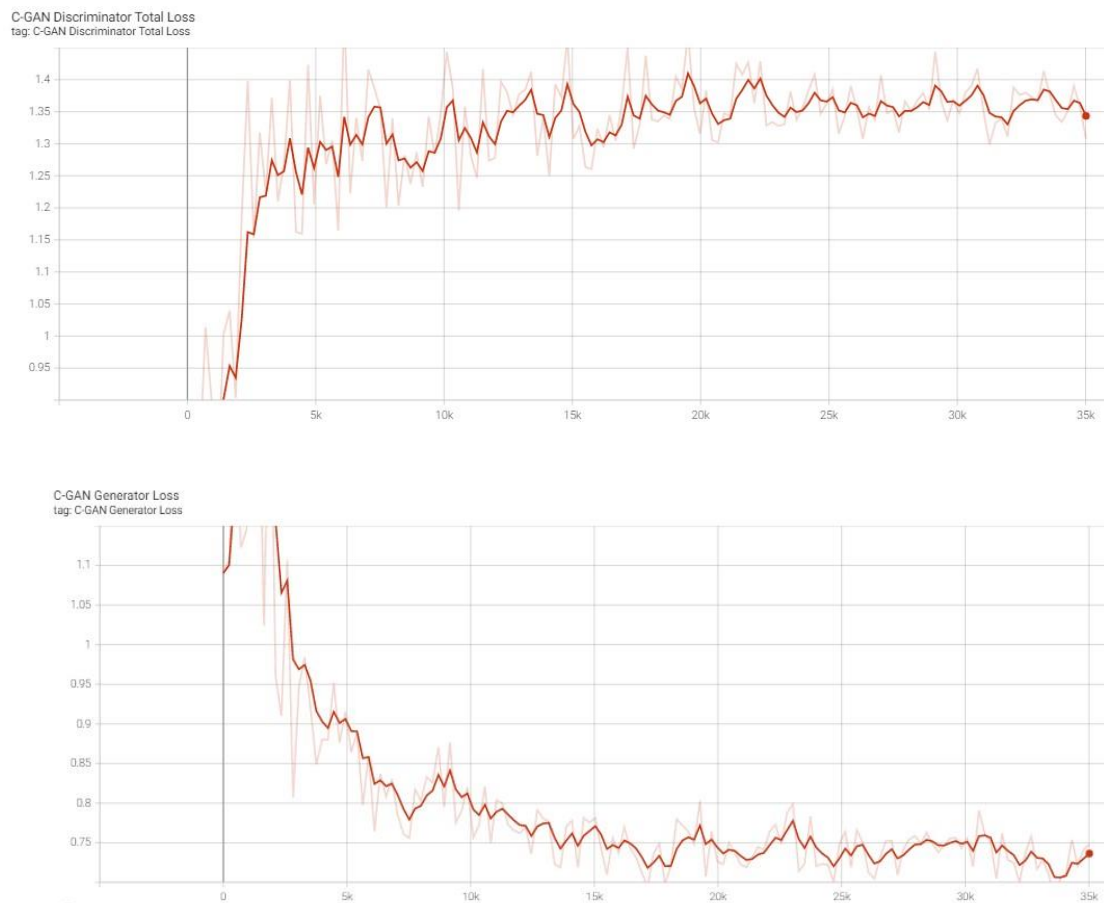
C Generator:

- Linear()
- Conv(in=256 (batch), out=128) -
Kernel=5, stride=1, paddin=2
- Betch Norm
- Leaky Relu
- Conv(in=128, out=64)-
Kernel=5, stride=2, paddin=2
- Betch Norm
- LeakyRelu
- Conv(in=64, out=1)-
Kernel=5, stride=2, paddin=2
- Tanh

C Discriminator :

- Flatten(794)
- Linear(in=28*28+10, out=256)
*10= One hot encoding of labels.
- Leaky Relu
- Dropout
- Linear (in=256, out=128)-
Kernel=5, stride=2, paddin=5
- Leaky Relu
- Dropout
- Linear(128,1)
- Leaky Relu
- Dropout

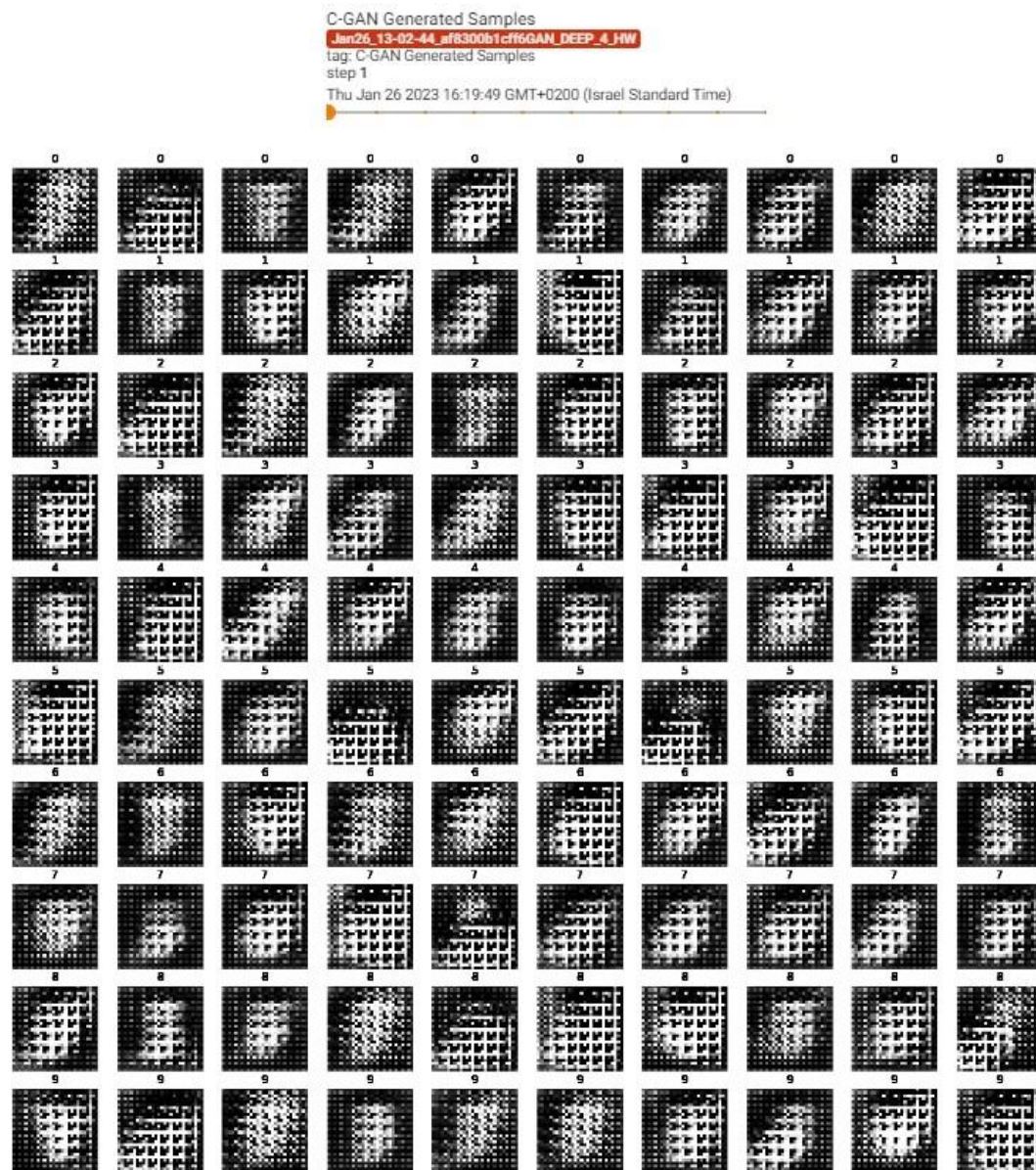
גרפי ה-Loss :



מהגרפים אנו רואים מה שציפינו לו, שה-Loss של ה-Discriminator עלה מכיוון שהיה לו קשה בזיהוי ומנגד מי שהקשה עליו ויצר תמונות ברמה טובה יותר זה ה-Generator אשר ה-Loss שלו ירד.

התמונות שיצאו מה- Generator :

הצגת התמונות נעשתה בשערוך מחזור של 40 Epochs מלבד ההתחלה והסוף בו נעשה במרווח צפוף יותר על מנת להראות את השינויים באופן טוב יותר.



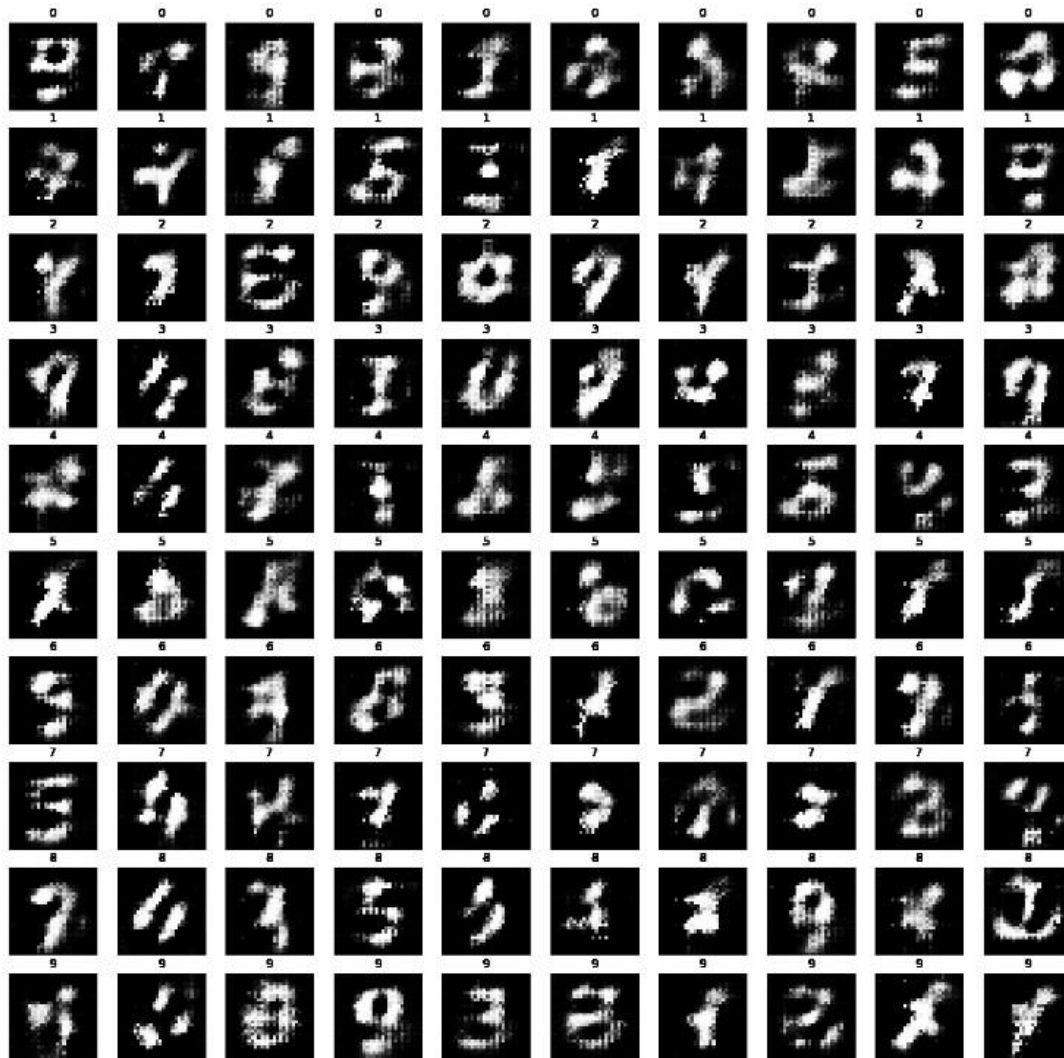
C-GAN Generated Samples

Jan26_13-02-44_af8300b1cf76GAN_DEEP_4_HW

tag: C-GAN Generated Samples

step 13

Thu Jan 26 2023 16:22:29 GMT+0200 (Israel Standard Time)



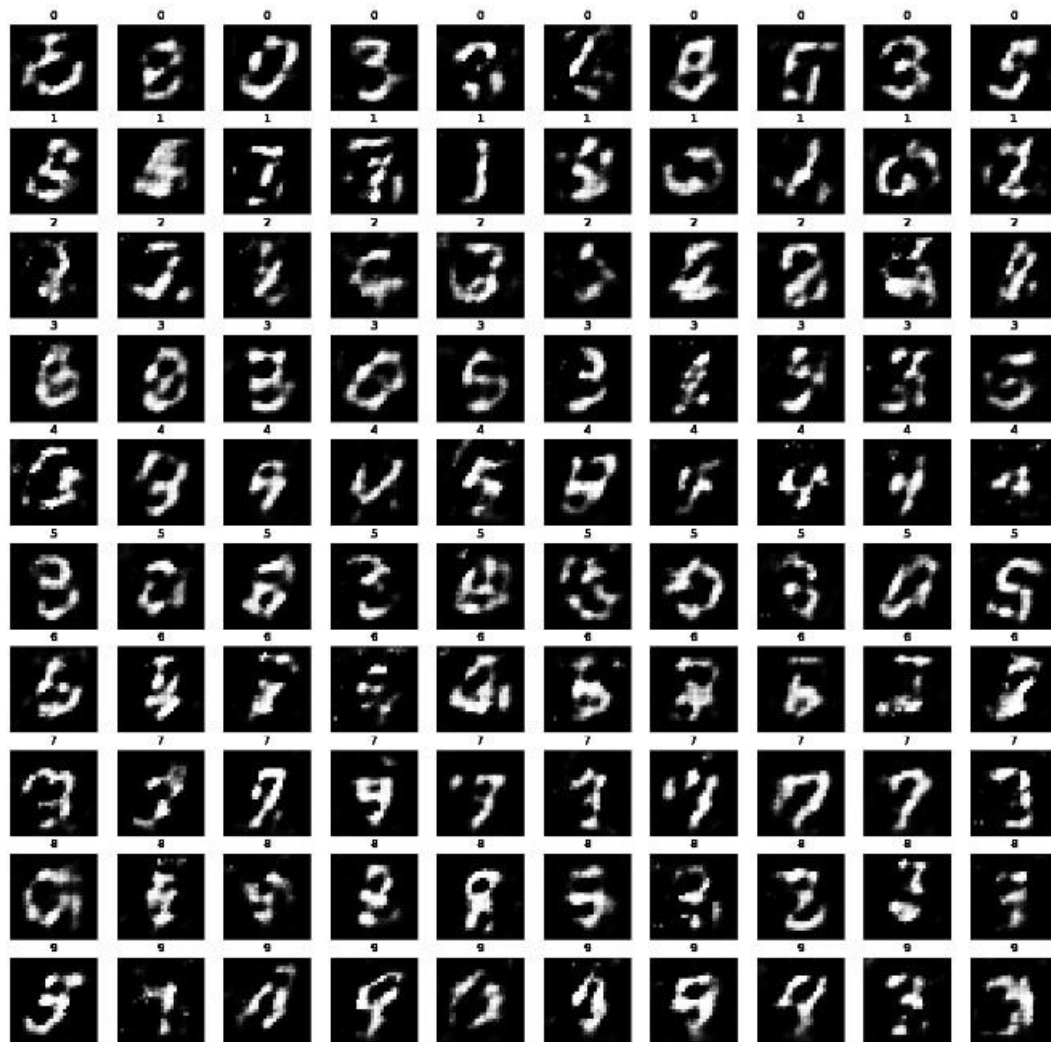
C-GAN Generated Samples

Jan26_13-02-44_sr8300b1cf16GAN_DEEP_4_HW

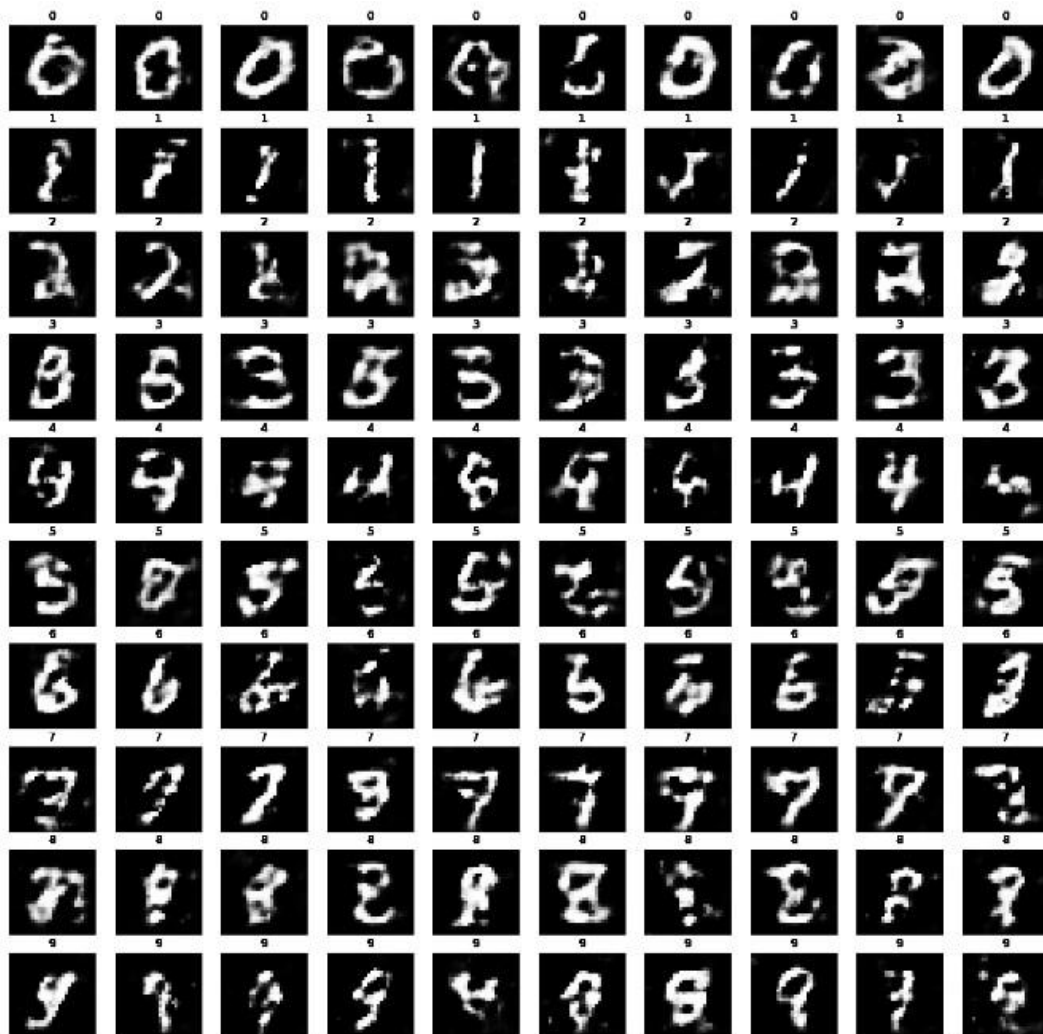
tag: C-GAN Generated Samples

step 42

Thu Jan 26 2023 15:12:55 GMT+0200 (Israel Standard Time)



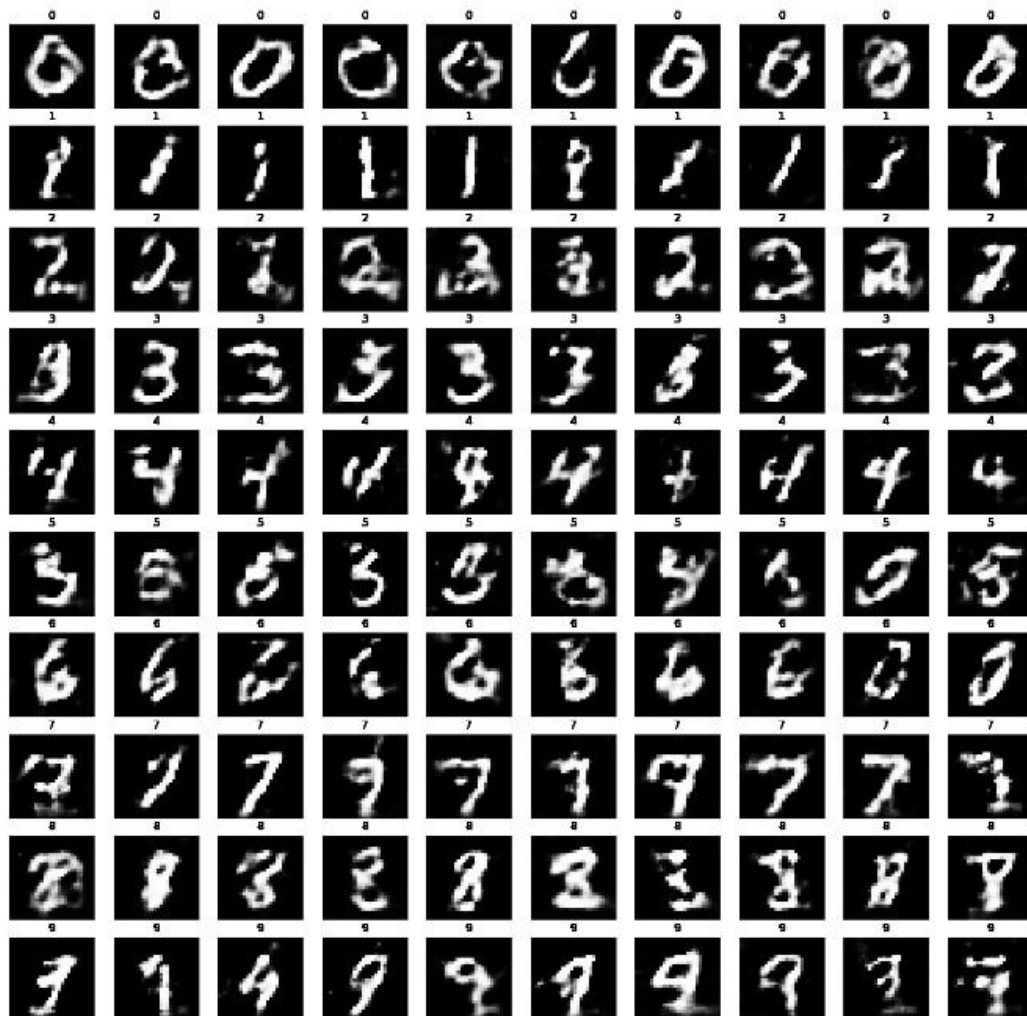
C-GAN Generated Samples
Jan26_13-02-44_af8300b1cf6GAN_DEEP_4_HW
tag: C-GAN Generated Samples
step 81
Thu Jan 26 2023 15:21:40 GMT+0200 (Israel Standard Time)



Jan26_13-02-44_af8300b1cff6GAN_DEEP_4_HW

step 120

Thu Jan 26 2023 15:30:29 GMT+0200 (Israel Standard Time)

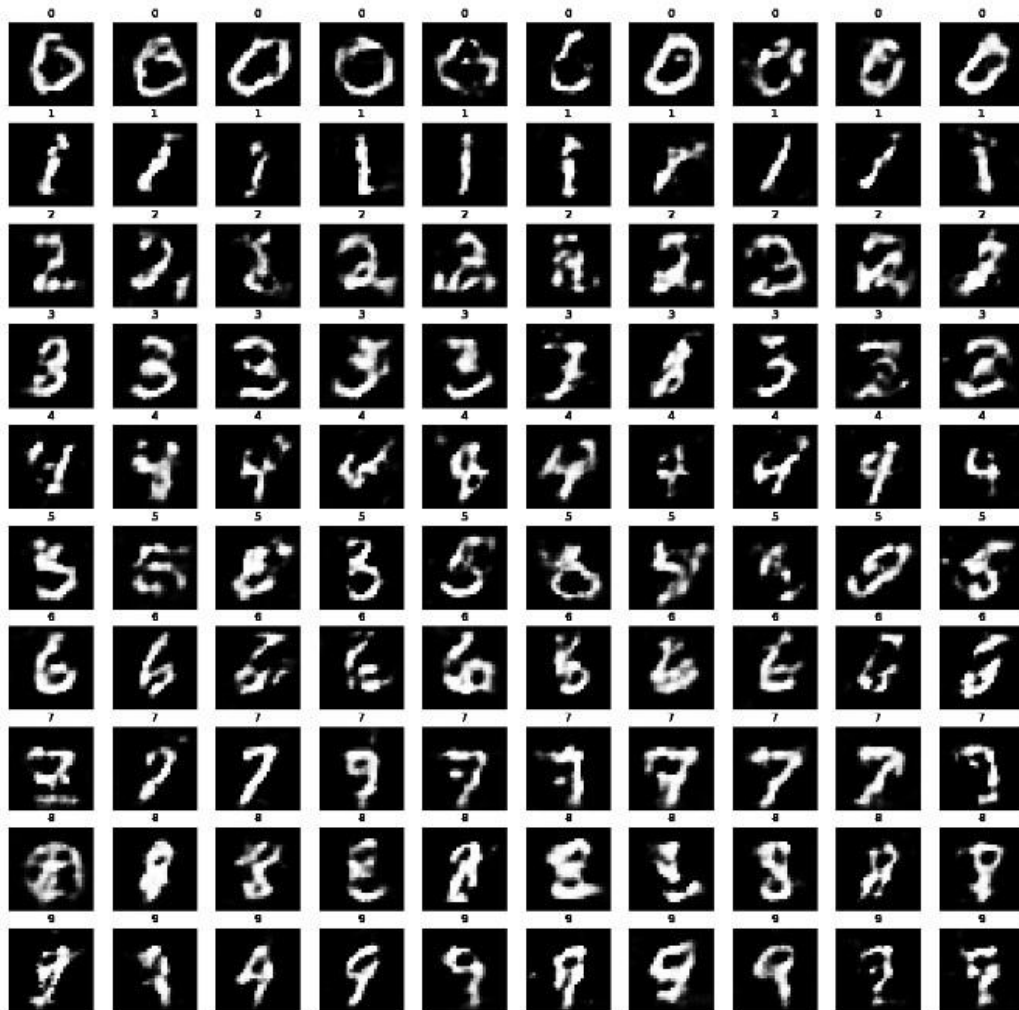


Jan26_13-02-44_af8300b1cff6GAN_DEEP_4_HW

tag: C-GAN Generated Samples
step: 100

step 129

Thu Jan 26 2023 15:32:29 GMT+0200 (Israel Standard Time)



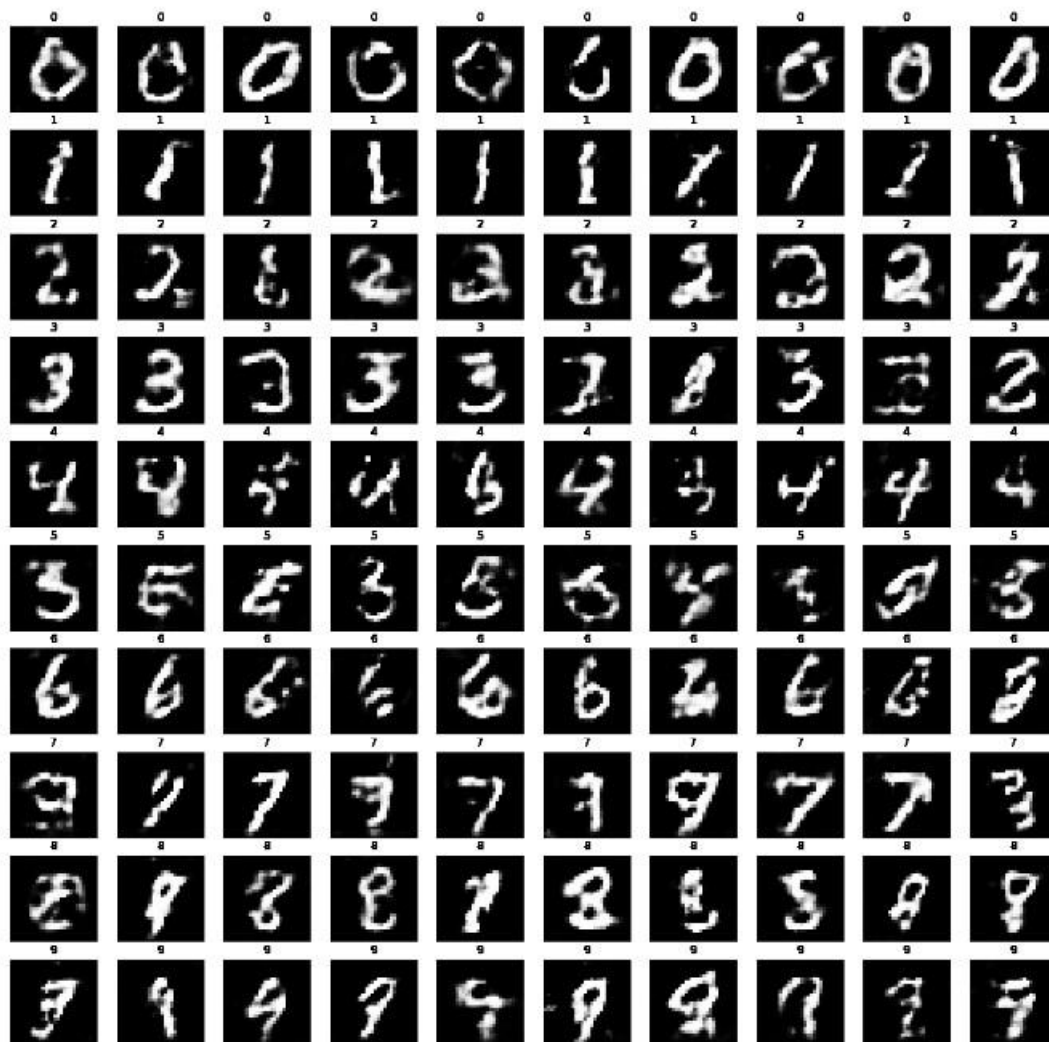
C-GAN Generated Samples

Jan26_13-02-44_af8300b1cf16GAN_DEEP_4_HW

tag: C-GAN Generated Samples

step 149

Thu Jan 26 2023 15:37:08 GMT+0200 (Israel Standard Time)



מסקנות מרשת ה-C-GAN :

- על פי האיטרציות ניתן לראות כי המודל מתחיל באיטרציה הראשונה בתמונות עם רעש רנדומלי בו אין אפשרות לזהות עדיין משהו ברור. באיטרציה ה-13 כבר מתחילים לזהות רקע נקי וספרות שמתחילות להיווצר במרכז אך עדיין ללא זיהוי משמעותי. באיטרציה ה-42 כבר ממש ניתן לראות חלק מהספרות, ומהאיטרציה ה-82 אנחנו כבר מזהים חלוקה ברורה על פי הלייבלים של הספרות וכן משתפרת איכות התמונה. כאשר מגיעים לאיטרציה ה-120 כבר המודל התאמן מספיק בכדי לזהות את רוב הלייבלים בצורה טובה.
- המודל למד בצורה טובה את הספרות – 0,1,3,8,9.
- ישנן ספרות שסיווג בצורה טובה אך היה גם בלבול בהרבה מהתמונות כמו הספרות- 5,7 בהן יש זיהויים של 5 כ-6 ו-7 כ-9.
- ההשערה שדבר זה קורה מכיוון שהספרות 5,6 די דומות בכתיבתן והספרות 7,9 תלוי בזווית של ובאופן הכתיבה בהן יכול להתפרש כ-9, מכאן שזוהי משימה לא קלה ל-Generator.
- יש תמונות אותן המודל כלל לא למד בצורה טובה כמו הספרות 4,2 שאכן ניתנות לזיהוי בחלק מהתמונות אך הטעויות רבות, טעות הספרה 2 יכולה לנבוע מהדימיון המרובה לכמה ספרות ומטכניקת הכתיבה, וטעות ספרה 4 במורכבות הכתיבה ובשוני שלה.

השוואת הרשתות :

Domain	GAN	C-GAN
Target	משמש למשימות Unsupervised.	יכול לשמש למשימות Supervised\Semi-Supervised.
Input	מייצר סמפלים מרעש.	מייצר סמפלים מרעש + תוספת תנאי.
Image Quality	איכות טובה יותר בד"כ.	איכות הניתנת לזיהוי אך מלווה ביותר רעש.
Training Speed	בד"כ מהיר יותר.	איטי יותר- זאת בעקבות שילוב קלט נוסף ולמידת פרמטרים נוספים. יש מקרים בהם התנאים שהוספנו מביאים להתכנסות מהירה יותר ומכאן לזמן אימון מהיר יותר.
Epochs to converges	60	80
Divergences	לא מכיל את כלל הספרות.	מכיל את כלל הספרות.