

## למידה עמוקה בראייה ממוחשבת

### פרויקט הקורס – תשפ"ב – סמסטר א'

סטודנטים יקרים,  
זה פרויקט סוף הקורס בלמידה עמוקה בראייה ממוחשבת.

### כללי ואדמיניסטרציה:

מועדי הגנות:  
מועד א' - 17/02/2022  
מועד ב' - 27/03/2022

### עפ"י נהלי הפקולטה:

- יש להעלות את הפרויקט למודל עד יום לפני מועד ההגנה אליו אתם נרשמים (שעה 23:59).
- סטודנט שלא יעלה את הפרויקט בזמן יורדו לו 20 נקודות מהציון הסופי
- סטודנט יכול להירשם למועד הגנה אחד בלבד. אין "מועד ב"
- בלוח ההגנות/בחינות יקבע מראש מועד מיוחד להגנה אליו יוכלו להירשם רק סטודנטים אשר לא נבחנו בשני המועדים וקיבלו אישור למועד מיוחד מהממונה לענייני סטודנטים. שימו לב שגם מי שיאושר לו לגשת למועד הגנה מיוחד – הגשת הפרויקט למודל הינה עד ה-26/3.
- סטודנט שלא נרשם לאף אחד משני המועדים ו/או שלא יאושר לו מועד ג', יקבל ציון נכשל בקורס ויהיה עליו להירשם לקורס חוזר בתשלום.
- כל הסטודנטים בקבוצה חייבים להירשם יחד לאותו מועד הגנה
- לא ניתן להירשם רישום כפול לשני מועדי הגנה
- ההגנה תתקיים בזום (אא"כ יוחלט אחרת...)
- הציון הינו פרטני – כל סטודנט יקבל ציון בהתאם לידע שיפגין

בימים הקרובים נעלה לכם קובץ לרישום להגנות

הגשה בקבוצות של ארבעה.

ההגנה הינה אישית. על כל אחד מחברי הקבוצה לשלוט ולהכיר את החומר במלואו.

## תיאור הפרויקט:

הפרויקט יעסוק בבעיית שחזור חלק חסר בתמונה, או בשמה המקצועי Image Inpainting. זהו פרויקט עם "עושר" גדול מאוד בניסויים ובאפשרויות השונות למימוש.

בתחום האפליקציות של ראייה ממוחשבת, קיים שימוש תדיר בפונקציונליות של Image Inpainting. ניתן לפגוש בפונקציונליות הנ"ל בתצורה של הסרת אובייקט, השלמת תמונה ועוד מקרים רבים ומגוונים.

בעיית ה-"Image Inpainting" היא אתגר קלאסי בתחום הראייה ממוחשבת שהראה המון התפתחות בשנים האחרונות.

בפרויקט הזה, אתם תחקרו את בעיית ה-Image Inpainting בהתבסס על המאמר הבא:

### **Context Encoders: Feature Learning by Inpainting**

Deepak Pathak   Philipp Krähenbühl   Jeff Donahue   Trevor Darrell   Alexei A. Efros  
University of California, Berkeley  
{pathak, philkr, jdonahue, trevor, efros}@cs.berkeley.edu

<https://arxiv.org/pdf/1604.07379.pdf>

דוגמא מתוך המאמר:



Input

Output

## הסבר על הדאטה:

הדאטה שתקבלו ימצא בשתי תיקיות שונות. התיקייה הראשונה (missing\_cats), היא תכיל 5652 תמונות שונות של חתולים כאשר בכל אחת מהן יופיע ריבוע לבן שיסתיר חלק מסוים מהתמונה. התיקייה השנייה (cats), היא תכיל 5652 תמונות שיציגו את הגרסה השלמה של תמונת החתול. לכל זוג תמונות יהיה את אותו השם שיהיה ייחודי רק עבורן. \*\* התמונות יגיעו בגדלים שונים כך שתצטרכו לשנות את כולן לגודל אחיד (לדוגמא: 256X256).



015.jpg



015.jpg

## מטרה:

המטרה הראשונה והעיקרית היא כמובן לבנות בעצמכם מערכת גנרטיבית היודעת להשלים בעצמה תמונת חתול.

המטרה השנייה היא לקרוא את המאמר הנ"ל, להבין אותו ולהשתמש בטכניקות המוצגות במאמר.

יש לציין כי אנו מצפים לראות אזכורים מתוך המאמר בתאי הטקסט של מחברת הפרויקט.

## הוראות טכניות:

כלל הפעולות שקשורות למודל עצמו תבצעו בעזרת שימוש בספרייה של TensorFlow Keras בלבד. מותר לכם לעשות שימוש בספריות עזר כגון, os, pandas, numpy, matplotlib וכו'. מעבר למודל עובד, אנו שמים דגש גדול על שלב הניסויים והדו"ח, גם במידה והמודל מוצלח, מחברת שתוגש ללא הסברים מפורטים תוביל לציין נכשל באופן אוטומטי. אנא התייחסו לדו"ח כמרכז הפרויקט.

**בדו"ח עצמו**, מאוד חשוב לראות את ההבנה שלכם והיכולת שלכם הן בבניית מע' למידה עמוקה על ההיבטים השונים והן בדיווח תוצאות מול הניסויים השונים. דוגמאות: אוגמנטציה, מה זה LOSS, על מה יכולה להשפיע שיטת אופטימיזציה שונה, מה עושות שכבות קונבולוציה ומה עושות שכבות ה-fully, מהו מקדם למידה (learning-rate) וכיצד הוא משפיע, למה צריך את קבוצת התיקוף וכו'. ככלל, בפרויקט זה יש "חופש פעולה" גדול מבחינת הניסויים. מה גודל תמונות כקלט, ארכיטקטורת רשת ועוד.

כמובן השתמשו בכלים שלמדנו כגון אוגמנטציות, dropout, regularization, stopping early, validation, activation, אתחול משקולות, norm batch, גודל תמונות כקלט למערכת, עם צבע/ללא, אין צורך לבחון את כל הצירופים האפשריים אולם יש להציג עושר בניסויים.

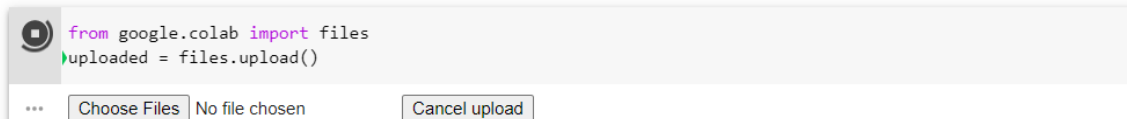
**\*\*** כאשר אתם כותבים את הדוח, **אנא שימו דגש על ה"ארכיטקטורה הפנימית"** של המודל שלכם (כגון: Encoder-decoder , Generator-discriminator).

במהלך הדו"ח תציגו **ארבעה** ניסויים שונים.  
אחד מתוך ארבעת הניסויים שתציגו יוגדר כ"הארכיטקטורה המנצחת".  
עבור כל ניסוי תציגו את הדברים הבאים:

- תיאור מילולי של הניסוי
- היפר פרמטרים
- אוגמנטציה
- ארכיטקטורה
- תוצאות ויזואליות של הניסויים שלהם כגון plot שמציג את פונקציית ה-loss של קבוצת האימון מול פונקציית ה-loss של קבוצת התיקוף.
- שימוש בכלים ויזואליים יצירתיים להערכת ביצועי מודל (ניתן כמובן לקבל השראה מהתוצאות שהציגו כותבי המאמר).
- ביצועי מערכת שונים
- משמעויות של בחירות שונות (לצורך העניין, שיקול בחירת מודל עבור Transfer-Learning)
- תוצאות ויזואליות של הניסויים השונים (בסופו של הדו"ח גם השוואה ביניהם)

**\*\*** את המודל המאומן עם ה"ארכיטקטורה המנצחת" אתם תשמרו בדרייב שלכם כקובץ עם סיומת h5. כמו כן, תבנו פונקציה ייעודית שטוענת את המודל המאומן ומאפשרת לעשות בו שימוש כ-Single-Prediction. הכוונה היא, שנוכל בכל זמן נתון לטעון את המודל המאומן, להעלות תמונה שהמודל יוכל לשחזר.  
לטובת העניין תעשו גם שימוש בפיצ'ר הבא שמאפשר להעלות קובץ.

#### ▼ Uploading Files



## הגנה:

ההגנה כמו הגשת הפרויקט תתבצע בקבוצה של רביעיות.

בהתאם להוראות נוספות שתמצאו בתיבת ההגשה, תתבצע הגנה בעל פה. כאשר כל אחד מחברי הקבוצה צריך לשלוט בכל היבטי הפרויקט.

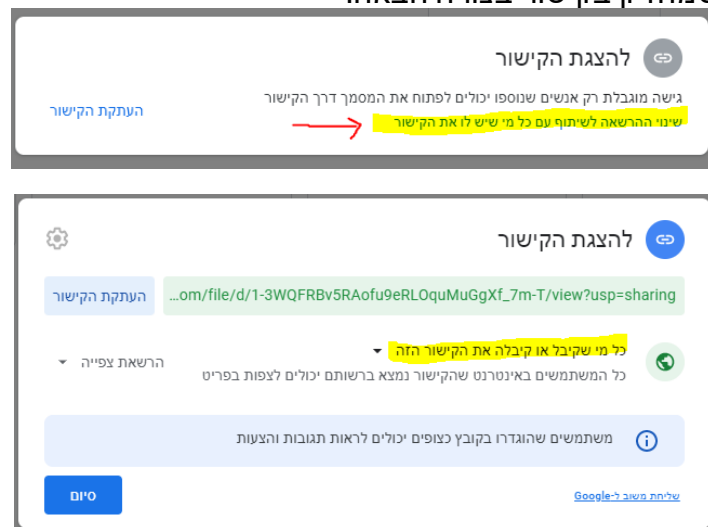
## הגשה:

ההגשה עצמה תתבצע על גבי מחברת Google Colab. המחברת תהיה מחולקת בצורה מסודרת, תכיל תאי קוד נפרדים ותאי טקסט המסבירים על הפעולות שנעשו.

**\*\*חשוב מאוד – בעת ההגשה, המחברות יכילו את כל הפלטים הרלוונטיים לתוצאות האימון\*\***

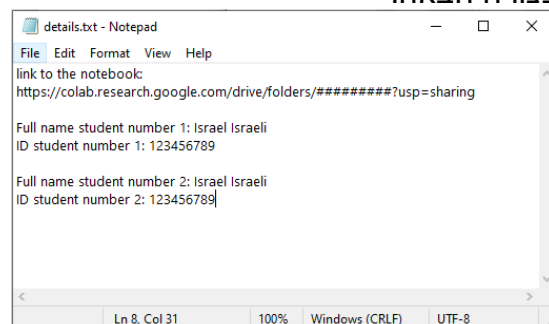
**\*\*חשוב מאוד – המחברת תכיל את שמות הסטודנטים המגישים ואת מספרי תעודת הזהות שלהם בתא טקסט שימוקם בחלק העליון של המחברת\*\***

את המחברת אתם תשתפו מתוך חשבון ה"Google Drive" שלכם, ניתן לייצר שיתוף לכל מי שמחזיק בקישור בצורה הבאה:



יש לציין בקובץ טקסט ששמו details.txt את הכתובות למחברות שלכם ואת ת.ז המגישים

בצורה הבאה:



בהצלחה!