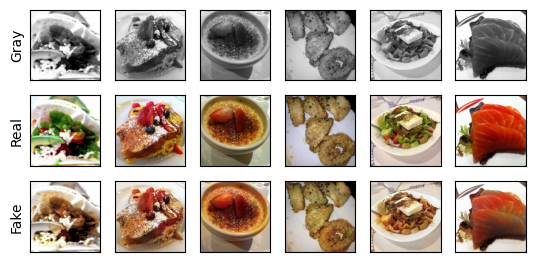
*פרויקט סופי בלמידה עמוקה – צביעת תמונות*

**אלברט כיאט – 212747026, דוד אפיק – 318267721**

תיאור הבעיה

עלינו ליצור מודל שמקבל תמונות בשחור-לבן וצובע אותן.

דוגמה לפלט הסופי של המודל:



דאטה

בחרנו ב-database "[Food-101](https://data.vision.ee.ethz.ch/cvl/datasets_extra/food-101/)" שבו יש 101,000 תמונות של אוכל, אבל לאימון לקחנו 8,000 תמונות באקראי (ועוד 2000 בשביל ה-testset).

ארכיטקטורה

פורמט התמונות:

בחרנו להמיר את התמונות למרחב LAB מכמה סיבות.

במרחב LAB, הערוץ L הוא פונקציה (לא לינארית) של התמונה בשחור לבן, והערוצים AB הם ערוצי הצבע. במרחב הזה אנחנו יכולים להפריד את התמונה הצבועה לערוץ L, ה-feature, ולערוצים AB, ה-label, ובכך להפוך את הבעיה להיות: בהינתן תמונה באפור (הערוץ L), המודל צריך לייצר את הצבע (הערוצים AB).

סיבה נוספת לשימוש במרחב הזה היא שמרחק בין תמונות מייצגת את ההבדל בנראות של הצבעים יותר מאשר המרחב RGB.

הרשת cGAN

הארכיטקטורה שמימשנו היא Conditional GAN (cGAN). בשונה מ-GAN רגיל שמקבל , ב-cGAN ה-generator מקבל feature בנוסף לרעש. -=-=-=--=-=-=-=-=-=-=-=-=-

cGAN מתאים לבעיה שלנו כי ה-generator מקבל תמונת אפור שהוא צריך לצבוע, ולא מייצר תמונה חדשה מרעש כמו ב-GAN. בנוסף, ה-discriminator מקבל את ה-feature ואת ה-label כדי שהוא יסווג את התמונה המלאה, ולא רק את הצבע.

GENERATOR – UNET

השתמשנו ברשת UNET דומה לזאת שראינו בקורס.

להלן תרשים הרשת:

מקרא מודולי הרשת:

A screenshot of a computer

Description automatically generated

רשת ה-UNET

A diagram of a computer

Description automatically generated

הרשת מקבלת תמונה עם ערוץ אחד (L – רמות האפור של התמונה), מוסיפה לה רעש ומחזירה את שני הערוצים של הצבע (A ו-B).

DISCRIMINATOR

השתמשנו ברשת fully-convolutional בעלת 6 שכבות:

A group of rectangular labels

Description automatically generated

הרשת מקבלת תמונה בעלת 3 ערוצים L, A, B, ומחזירה ערך בין 1 ל-0 שאומר אם התמונה אמיתית (1) או מזויפת (0).

אימון

Loss

פונקציית ה-loss ב-GAN היא:

כאשר מתפלג מתוך ה-data האמיתי שמייצרים, ו- הוא רעש שמתפלג בצורה כלשהיא.

פונקציית ה-loss שלנו (cGAN) היא:

כאשר הוא התמונה בשחור לבן, הוא הצבע של התמונה ו- רעש.

*בנוסף ל-cGAN רגיל, הוספנו ל-loss את פונקציית ה-MAE (mean absolute error), או , בין הצבעים שה-generator מייצר לבין הצבע האמיתי, והוא מוכפל בערך . עם ההוספה הזאת, כשעושים צעד בגרדיאנט של ה-generator, הוא מנסה לא רק לעבוד על ה-discriminator, אבל גם להתקרב לצבע האמיתי עם משקל .*

*לבסוף, ה-objective הוא:*

איך מחשבים את ה-loss:

דגדג

תוצאות: