למידה עמוקה – פרויקט סופי

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

## הבעיה

הפרויקט שבחרנו לעשות הוא צביעת תמונות בשחור לבן, וה-dataset שבחרנו לעבודה הוא "[Food-101](https://data.vision.ee.ethz.ch/cvl/datasets_extra/food-101/)" שבו יש 101,000 תמונות של אוכל, אבל לאימון לקחנו 10,000 תמונות (ועוד 1,500 בשביל ה-testset).

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.

## הארכיטקטורה

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.

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 .*

*ההוספה הזאת ל-loss נתנה ביצועים יותר טובים, ואת הרעיון לקחנו מהמאמר שידוע בתור “*[*pix2pix*](https://arxiv.org/abs/1611.07004)*”.*

*לבסוף, ה-objective הוא:*

*בשביל לחשב את , מימשנו GANLoss.py.*

**

### Generator

ה-generator שמימשנו הוא ה-Unet. הארכיטקטורה שמימשנו

### Discriminator

הארכיטקטורה של ה-discriminator היא: