**דו"ח מסכם למידה עמוקה תרגיל בית 2**

מגישים: רועי זהר 209896174 רועי מזן 319042800

בתרגיל 2 נתבקשנו לאמן רשת נוירונים על CIFAR10, מאגר תמונות מוכר עם 10 מחלקות, כך שמטרת הרשת תהיה לסווג את התמונות בדיוק של 78%, במינימום פרמטרים.

**הרשת המקורית**

התבקשנו לחשב את גודל הרשת המקורית. להלן החישוב של כמות הפרמטרים ברשת המקורית:

1. שכבה ראשונה – 896
   1. שכבת קונבולוציה של 32\*3\*3\*3
   2. 32 biases
2. שכבה שניה – 18432
   1. שכבת קונבולוציה של 64\*32\*3\*3 = 18432
   2. 64 biases
3. שכבה שלישית – 73856
   1. שכבת קונבולוציה של 64\*128\*3\*3 = 73744
   2. 128 biases
4. שכבה ריביעית – 65664
   1. שכבת קונבולוציה של 64\*128\*3\*3 = 65664
5. 3 שכבות Fully connected
   1. 1500\*128\*4\*4 = 3072000 + 1500 biases
   2. 1000\*1500 = 1500000 + 1000
   3. 1000\*10 = 10000

סך הכל יוצא 4743358

**הרצה**

כדי לשחזר את התוצאות שלנו, יש להריץ את הקובץ python cifar10\_train.py. כדי לעלות את הרשת המאומנת שלנו (צירפנו שתי רשתות בתיקיית models), יש להריץ את הקובץ python cifar10\_eval.py, והוא מדפיס את הדיוק והloss.

**מה עשינו**

תוצאה סופית: הצלחנו להגיע לדיוק של 79.14% עם 86564 פרמטרים.

אפשר לחלק את העבודה שלנו לשני חלקים, שהם

**1) Data Augmentation**

בחלק של עיוות התמונה, ביצענו טרנספומרציות רבות על התמונות לפני שהשתמשנו בהן בשביל אימון. ביצענו חיתוך של ריבוע קטן מגודל 24 על 24 מתוך התמונה הגדולה, הפכנו אותו, שיחקנו עם הקונטראסט והבהירות וממוצעים של התמונה. קיבלנו תוצאות יפות עם השינוי הזה, והוא הביא לשיפור האימון. קיבלנו גם תוצאות עוד יותר טובו כששיחקנו עם סדר הטרנספומרציות, לאחר שראינו שהפעולות שאנו מפעילים אינו קומוטטיבי.

**2) משחק עם ארכיטקטורת הרשת**

שיחקנו הרבה עם מבנה הרשת, בהתחלה עבדנו עם מבנה הרשת שניתן כדוגמא בתרגיל, ולאחר מכן עברנו לעבוד עם ארכיטקטורה המוצעת באתר tensorflow. ראינו שהתוצאות הכי טובות התקבלו כשהרצנו עם מספר פרמטרים שונה, תוך שימור על המבנה הכללי של הרשת.

הרצנו 2 רשתות במקביל, ואלו התוצאות שקיבלנו:

|  |  |  |
| --- | --- | --- |
| **Accuracy:** | **Parameters:** | **Index:** |
| 79.14% | 86564 | 1 |
| 81.5% | 225000 | 2 |

**ארכיטקטורה סופית**

הארכיטקטורה שבסוף הלכנו איתה דומה לזו שמצאנו באתר של tensorflow, והיא הניבה רשת המגיעה לדיוק של 78.4 עם 86000 פרמטרים.

שכבה 1) קונבולוציה ראשונית, עם אאוטפוט של 30 נוירונים. לאחר מכן RELU. לאחר מכן POOLING ונרמול.

שכבה 2) קונבולוציה שניה, שוב אאוטפוט של 30 נוירונים. לאחר מכן RELU. נרמול וPOOLING

שכבה 3) ו4) שתי שכבות נוספות, שוב אאוטפוט של 30 נוירונים, מכפלה לינארית ואז RELU בשתיהן.

שכבה 5) מכפלה לינארית אחרונה + SOFTMAX, עם 10

הערה: שיחקנו עם פונקציות אקטיבציה נוספות אך לא הגענו לתוצאות טובות יותר מאשר בRELU.

**גרפים של LOSS וACCURACY:**

שיחקנו גם עם הoptimizers, והשווינו בין התוצאות. השווינו בין ProximalGradientDescentOptimizer, לבין AdamOptimier, וקיבלנו תוצאות מעט יותר טובות עבור הAdam Optimizer.

הערה: הייתה לנו בעיה לייצא את הtraining accuracy לתוך הגרפים.

**Adam Optimizer:**

עבור הoptimizer הזה, הגענו לדיוק מירבי של 79.14% לאחר 200 epochs. את הסף של 78% דיוק עברנו בepoch 113.

**A screenshot of a map

Description generated with very high confidenceA picture containing screenshot

Description generated with very high confidence**

**Proximal Gradient Descent Optimizer:**

עבור הoptimizer הזה, הגענו לדיוק מירבי של 78.97% לאחר 200 epochs. את הסף של 78% דיוק עברנו בepoch 129.

**A screenshot of a cell phone

Description generated with very high confidenceA picture containing screenshot

Description generated with very high confidence**