**יישומים ברשתות נוירונים עמוקות**

**תרגיל בית 4 חלק א**

**הרשת שלי מרוכבת מכמה שכבות קונבולוציה ואחריה כמה שכבות fully connected**

**כל שכבת קונבולוציה מרכבות מ-**

1. **Convolution עם גודל קירנל של 3\*3 עם מספק פילטרים מוכפל ב-2 כל פעם עם צעד אחת ובלי ריפוד באפסים.**
2. **batch normalization בגודל מספר הפילטרים.**
3. **Max pooling בגודל 2\*2**
4. **activation function : ReLU**

**צורת כל שכבה :**

self.<layer\_number>= nn.Sequential(  
 nn.Conv2d(input\_depth, filters\_number, kernel\_size=3, padding=0, stride=1),  
 nn.BatchNorm2d(filters\_number),  
 nn.ReLU(),  
 nn.MaxPool2d(2)  
)

**צורת כל שכבת fully connected :**

self.<fully connected layer\_number>= nn.Linear(in\_features ,out\_features)

**ה- Weights Initialization משתמש ב- Xavier weight initialization עם אתחול ל- Batch Normalization.**

**השתמשתי גם ב- nn.Dropout(0.5) כאשר עוברים בין שכבות הקונבולוציה לשכבות ה fully connected-**

**תוצאות הקוד:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Accuracy** | **Learning Rate** | **Optimizer**  **Decay for ADAM Weight +**  **Momentum for SGD +** | **Batch Size** | **epochs** | **Image Size** | **Fully Connected**  **Number** | **Convolution Number** |
| **63%** | **0.0001** | **Adam, weight decay=1e-4** | **32** | **50** | **128x128** | **2** | **3** |
| **47%** | **0.001** | **Adam, weight decay=1e-4** | **32** | **50** | **128x128** | **2** | **3** |
| **72%** | **0.0001** | **Adam, weight decay=1e-4** | **32** | **50** | **256x256** | **2** | **3** |
| **68%** | **0.0001** | **Adam, weight decay=1e-4** | **32** | **50** | **400x400** | **2** | **3** |
| **62%** | **0.0001** | **SGD, momentum=0.9** | **32** | **50** | **400x400** | **2** | **3** |
| **71%** | **0.0001** | **Adam, weight decay=1e-4** | **64** | **70** | **400x400** | **3** | **7** |
| **78%** | **0.0001** | **Adam, weight decay=1e-4** | **128** | **70** | **400x400** | **3** | **7** |
| **79%** | **0.0001** | **SGD, momentum=0.9** | **128** | **70** | **400x400** | **3** | **7** |

**ניתוח הקוד:**

1. **הגדלת מספר שבות הקונבולציה : אם הגדלתי את שכבות הקוהבלוציה ברשתות נוירונים עמוקות, נכון לעכשיו יצא לפני שמעתנו שתוכל לקבוע סדר מעקב יותר מדויק לנתונים כדי לפחות את הקוהבלוציה ולמקסם את היעילות. אך, ישנם גם מגמות לא יעילות בגדלת שכבות הקוהבלוציה. למשל, יכול להתרחש שכאשר יש יותר שכבות, יתרון הגדרת סדר מעקב מדויק לא יהיה מאוד משמעותי בעוד שכבות נוספות נוספות יוסיפו עוד יותר קוהבלוציה וזמן המעבר לנתונים יארך. לכן, יש לעשות שימוש עצבני בגדלת שכבות הקוהבלוציה ברשתות נוירונים עמוקות כדי לוודא שנתונים המעבר דרך הרשת יעבדו באופן מאוד יעיל.**
2. **הגדלת Batch\_Size : הגדלת גודל הקבצים (Batch\_Size) ברשתות נוירונים עמוקות יכולה ליצור משמעות רצינית על היעילות של הרשת. כאשר גודל הקבצים גדול יותר, ניתן לעבד יותר נתונים בכל איטרציה וכך להקטין את זמן הלמידה של הרשת. עם זאת, ישנם גם מגבלות כאשר גודל הקבצים גדול מדי כך שעלול לגרום לעומס על המעבד וכך לפחות את היעילות במקום להגדילה. כך שחשוב למצוא את הגודל המתאים לפעולות הספציפיות של הרשת.**
3. **הורדת Learning Rate: הורדת ה Learning Rate ברשתות עמוקות יכולה להאט את התהליך הלמידה בעת האחזור של הרשת. ה Learning Rate נועד לקבוע את הצעד הכי גדול שנוכל לעשות בכל מרכז למדע כאשר אנחנו נעבור דרך העדכון של הפרמטרים של הרשת. אם הורדתי את ה Learning Rate, אני אוסיף עצור לרשת לעדכן את הפרמטרים באופן מהיר יותר ואז הרשת תקח יותר זמן כדי ללמוד את הפתרון הטוב ביותר.**
4. **הגדלת את Image Size :** **אם הגדלתי Image Size ברשתות נוירוניות עמוקות, הפעולה תאפשר לנו לנהל את גודל התמונה המוצגת ברשת נוירונית העמוקה. בעקבות הגדלת גודל התמונה, הנתונים המוצגים בתמונה יהיו יותר מפורטים וניתן יהיה לראות פרטים נוספים בתמונה. עם זאת, הגדלת גודל התמונה עשויה לגרום לפעילות אחדית לאטה יותר ברשת נוירונית העמוקה, כיוון שכך יש צורך במשאבים נוספים כדי לנהל את התמונה הגדולה ברשת נוירונית העמוקה.**
5. **הגדל מספר epochs : אם הגדלתי את מספר ה-epochs ברשתות נוירוניות עמוקות, זה אומר שאני ארצה לעשות עוד מספר מעגלים נוספים של ריצת הנתונים על הרשת, כדי לסדר את הנתונים עמוק יותר. הפעולה הזאת תעשה כי הרשת תעסוק יותר זמן בלמידה ותתרגם בטווח הנתונים של הרשת שתהיה יותר טובה. עם זאת, אם אני מגדיל מדי את מספר ה-epochs, זה עשוי לגרום לפנייה והרשת תתחיל לדעת את הנתונים באופן שגוי ולא יספק תוצאות טובות.**