**תרגיל בית 2 - רשתות קונבולוציה**

# הקדמה

במסגרת תרגיל בית זה מימשנו מספר רשתות קונבולוציה ביחד לסיווג תמונות ממאגר הנתונים Cifar10. הרשתות השונות אומנו במספר קונפיגורציות, כאשר בסופו של דבר נבחרה הרשת שהניבה את הביצועים הטובים ביותר - שגיאה של ??? על סט הבוחן. בנוסף לכך בוצעה השוואה לבחינת השפעת העיבוד המקדים (augmentation) על טיב התוצאות.

# תהליך הלמידה

לצורך בחירת המודל המיטבי אומנו 3 רשתות. עבור כל רשת נבדקה השפעת קצב הלימוד (lr = [???,???]) וגודל תת-קבוצת האימון (bs = [???,???]), כאשר השוואת הביצועים בוצעה לאחר אימון של 300 epochs. האימון התבצע כנדרש על קבוצת האימון המכילה 50k תמונות מתוייגות. אתחול משקלי שכבות הקונבולוציה בוצע על ידי שימוש באתחול Xavier, כהנחה שזהו אתחול מושכל יותר (הנובע משיקולי variance). פונקציית ההפסד שהוגדרה היא NLLL (שיחד עם Soft-max מהווה הפסד Cross-entropy) ואלגוריתם ADAM שימש לצורך אופטימיזציה. עבור כל ההרצות הללו בוצע עיבוד מקדים לתמונות בסעיף 3.א.

נבחנו 3 ארכיטקטורות רשת שונות:

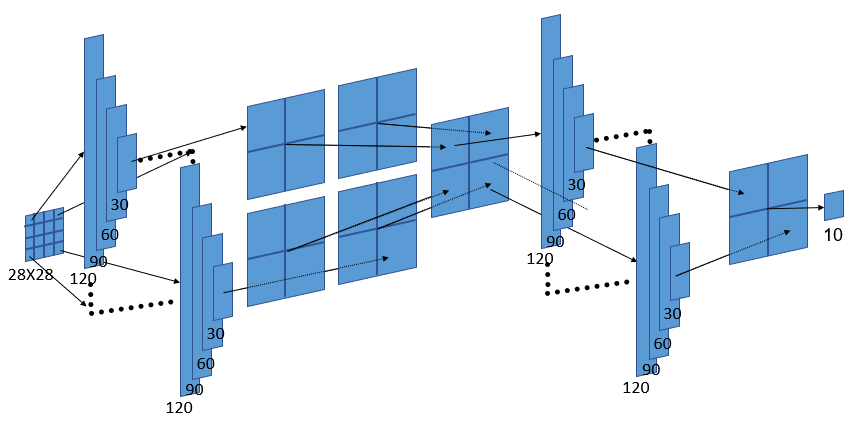
* מודל 1 (VGG 4) - רשת בעלת 3 שכבות קונבולוציה עם [20,40,80] פילטרים בגודל 3x3 ופסיעה 1. לאחר כל שכבת קונבולוציה ישנה פונקציית אקטיבציה ReLU ו- MaxPooling בגודל 2x2 עם פסיעה 2. לאחר שכבות הקונבולוציה ישנה שכבת FC עם 1280 נוירונים ולאחר מכן הסיווג מבוצע באמצעות Soft-max. לרשת סה"כ 49,490 משקלים. ראה איור 1.
* מודל 2 (mini VGG) - רשת בעלת 4 שכבות קונבולוציה עם [5,5,10,10] פילטרים בגודל 3x3 ופסיעה 1. לאחר כל שכבת קונבולוציה ישנה פונקציית אקטיבציה ReLU. אחרי שכבות הקונבולוציה השנייה והרביעית מבוצע MaxPooling בגודל 2x2 עם פסיעה 2. לאחר שכבות הקונבולוציה ישנן שתי שכבות FC, הראשונה עם 640 נוירונים והשנייה עם 70 נוירונים המלוות בפונקציות אקטיבציה ReLU ורגולריזציה על ידי DropOuts בהסתברות 0.5. לאחר השכבה האחרונה מבוצע Soft-max לצורך סיווג. סה"כ 47,320 משקלים במודל. ראה איור 2.
* מודל 3(DenseNet) - רשת שיוצרת יחידות של מספר שכבות קונבולוציה יחד עם Append בין השכבות. בתחילת הרשת יש קונבולוציה אחת רגילה עם מספר פילטרים [3,24]. לאחר מכן יש שלושת יחידות של קונבולוציה עם Append בתוכן המופרדות ע״י BatchNorm, ReLU ו-AvgPool. לאחר מכן יש FC עם [48,10]. לאחר השכבה האחרונה הסיווג מבוצע באמצעות Soft-max. סה"כ 44,410 משקלים במודל. ראה איור 3.

התוצאות הטובות ביותר עבור כל מודל, כלומר - השגיאה היחסית הקטנה ביותר על פני סט הבוחן, מוצגות בטבלה 1 יחד עם הפרמטרים המתאימים (lr, wd, bs, epoch).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **מודל** | **מספר משקלים ברשת** | **שגיאת וולידציה יחסית מינימלית במהלך האימון [%]** | **lr** | **wd** | **bs** | **epoch** |
| **1** | 64,711 | 2.240 | 1e-3 | 1e-4 | 32 | 63 |
| **2** | 42,310 | 2.940 | 1e-3 | 1e-5 | 32 | 93 |
| **3** | 57,520 | 1.458 | 1e-3 | 1e-4 | 64 | 77 |

טבלה 1 - השוואת טיב המודלים

|  |  |  |
| --- | --- | --- |
| Image  [3, 32x32]  Conv1  [20, 32x32]  Conv2  [40, 16x16]  Conv3  [80, 8x8]  Flatten  FC  [1280]  איור 1 - מודל 1 | | Conv1  [5, 32x32]  Conv2  [5, 16x16]  Conv3  [10, 16x16]  Conv4  [10, 8x8]  Flatten  FC1  [640]  Image  [3, 32x32]  FC2  [70]  איור 2 - מודל 2 |
|  |



איור 3 - מודל 3

# השפעת העיבוד המקדים והמשך אימון המודל המיטבי

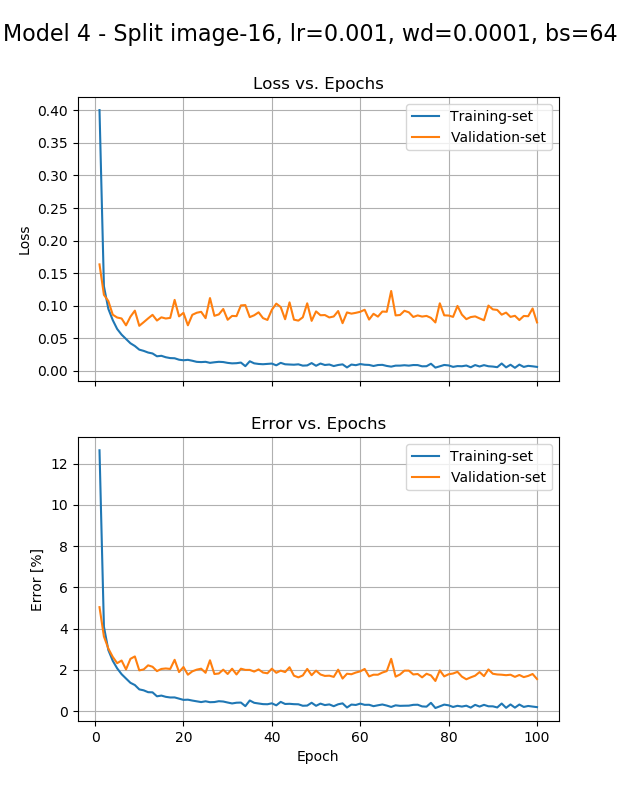
המודל שהניב את התוצאות הטובות ביותר הוא מודל 3, כפי שניתן לראות בטבלה 1. ערכי פונקציית ההפסד והשגיאה היחסית לפי ה- epoch על פני סט האימון מוצגים על סט האימון באיור 8. ניתן לראות שערך השגיאה ממשיך לרדת - לא נראה לי נכון...

לאחר בחירת המודל המיטבי בחנו את השפעת העיבוד המקדים על טיב התוצאות. נבחנו העיבודים המקדימים הבאים:

1. ריפוד הדפנות של התמונה ב-2 פיקסלים לכל כיוון ובחירה רנדומלית של איזורי 32x32. בנוסף לכך היפוך אופקי בהסתברות 0.5.
2. אופציה א' ובנוסף שינוי בהירות התמונה ב- ±50% בהסתברות 0.5.
3. אופציה א' ובנוסף שינוי ניגודיות התמונה ב- ±50% בהסתברות 0.5.
4. איחוד של אופציה א', ב' ו-ג'.

השפעות העיבוד המקדים מוצגות בטבלה/גרף ???.

המשכנו לאמן את המודל הטוב ביותר למשך 1000 epochs על סט האימון. הדיוק שהתקבל לבסוף על סט הבוחן(test set) הוא 98.87. ניתן להתרשם מתהליך האימון בגרף ????.



*איור 8 - הפסדים ושגיאות בזמן הלמידה*

# סיכום ומסקנות

התחלנו עם ניסיון להתמודד עם הבעיה בעזרת מודלים המוכוונים לרעיון עיבוד התמונה ע״י דחיסת התמונה ומיצוע המידע בה ע״י מספר גדול של פילטרים על פני מימד קטן, כמו מודל הVGG. אך התוצאות של המודלים הללו לא סיפקו את דרישתנו. לאחר מכן החלטנו לעבור למודל שמשתמש בעקרון דחיסת שכבות הקונבולוציה עם Append מעל השכבות. המודל החדש אפשר לנו להוסיף יותר שכבות על חשבון מספר פילטרים קטן יותר בכל שכבה. המודל השלישי הקטין לנו את אחוז השגיאה בכ8%. ההקטנה אכן משמעותית אם אנחנו לוקחים בחשבון את הגבלת מספר הפרמטרים(50000 פרמטרים).

לאחר שהחלטנו על המודל השלישי התחלנו בניסיון למצות את המידע בתמונות על ידי שינוי של התמונות עם שינוי בהירות, היפוך או ניגודיות של התמונות. השינויים הללו בנוסף לאפשרות להתבונן אחרת על התמונות, הרחיבו את כמות התמונות השונות שהמודל אומן עליהם, מכיוון שכל Batch התמונות נדגמות עם הסתברות לשינוי כל תמונה לאחד האפשרויות(עם אפשרות להפעלה של מספר שינויים יחדיו). הוספת השינוי סט האימון הקטין עוד טיפה את אחוז השגיאה של המודל לאחוז הסופי שקיבלנו.