

## דוח תרגיל 5

אורי דאבוש ואיתי גרדנויץ

בתרגיל זה מימשנו מודל שמבוסס על ResNet. המודל היה מורכב מכמה שכבות שהגדרנו כבלוק. הבלוק שלנו מורכב מהשכבות הבאות:

שכבת קונבולוציה

batch normalization

פונקציית אקטיבציה (relu)

שכבת dropout

שכבת קונבולוציה

batch normalization

שכבת קונבולוציה

batch normalization

שכבות הקונבולוציה משתמשות בפרמטר kernel\_depth שהגדרנו שהוא מספר ה-kernel'ים שאנחנו משתמשים בהם בשכבה.

לאחר מכן מימשנו פונקציית make\_layer שמקבלת כפרמטרים kernel\_size, kernel\_depth שהם פרמטרים של שכבות הקונבולוציה וה-, batch normalization in\_channels שהם ה-in channels בשכבת הקונבולוציה הראשונה, stride ו-padding בהם משתמשים בשכבות הקונבולוציה, ו-p שהוא ההסתברות לאיפוס איבר ב-dropout. הפונקציה בונה שכבה שמורכבת מהחלקים הבאים:

שכבת קונבולוציה

batch normalization

פונקציית אקטיבציה (relu)

ו-2 בלוקים כפי שתיארנו קודם

במודל עצמו מורכב מהשכבות הבאות:

שכבת קונבולוציה ראשונה (שמקבלת כ-in channels את ה-in channels ב-input, כלומר 1)

batch normalization

פונקציית אקטיבציה (relu)

שכבה שמתקבלת מהפונקציה make\_layer

שכבת MaxPool

שכבה שמתקבלת מהפונקציה make\_layer

שכבת MaxPool

שכבה שמתקבלת מהפונקציה make\_layer

שכבת MaxPool

שכבת fully connected

שכבת dropout

שכבת fully connected

שכבת dropout

את הפרמטרים לשכבות השונות ניתן לראות במילון הנ"ל:

```
{
  "conv0": {"kernel_depth": 16,
            "kernel_size": 3,
            "stride": 1,
            "padding": 1,
            "p": 0.1
            },
  "conv1": {"kernel_depth": 16,
            "kernel_size": 3,
            "stride": 1,
            "padding": 1,
            "p": 0.1
            },
  "conv2": {"kernel_depth": 32,
            "kernel_size": 3,
            "stride": 1,
            "padding": 1,
            "p": 0.1
            },
  "conv3": {"kernel_depth": 64,
            "kernel_size": 3,
            "stride": 1,
            "padding": 2,
            "p": 0.1
            },
  "maxpool1": {
    "kernel_size": 2,
```

```

        "stride": 3},
    "maxpool2": {
        "kernel_size": 2,
        "stride": 2},
    "maxpool3": {
        "kernel_size": 3,
        "stride": 2},
    "linear1": {
        "out": 512,
        "p": 0.4},
    "linear2": {
        "out": 30,
        "p": 0.3
    }
}

```

ההיפר פרמטרים שהשתמשנו בהם הם האופטימיזר, שהוא Adam במקרה שלנו, עם הפרמטרים הבאים:

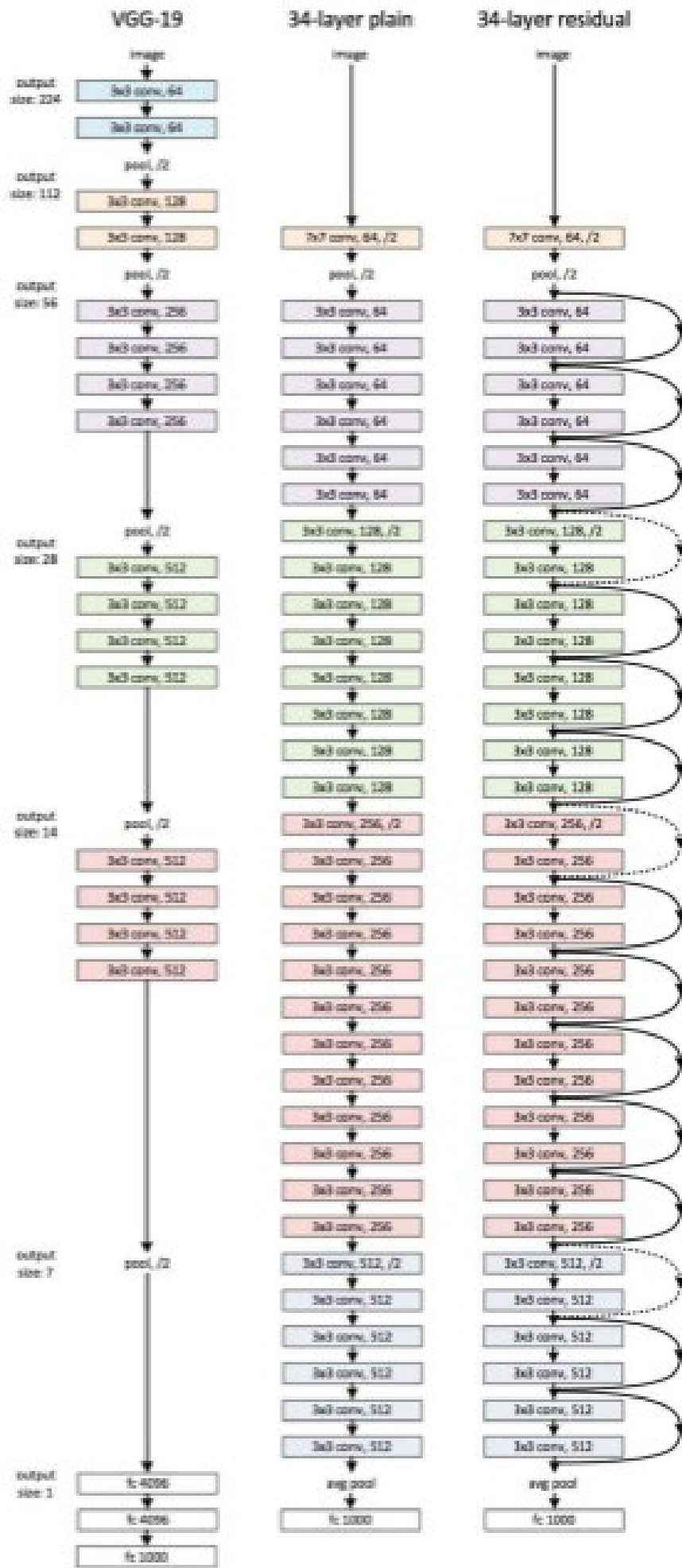
```

lr=0.0004, betas=(0.999, 0.999), eps=1e-08, weight_decay=0,
amsgrad=True

```

את האימון הרצנו במשך 10 איפוקים, כשבכל איפוק אנחנו בודקים האם המודל שלנו מפחית את הממוצע של הלוס על סט הולידציה ושומרים את ה-state של המודל בהתאם. לאחר האימון אנחנו טוענים את המודל הטוב ביותר ומריצים אותו על ה-test, וכותבים את התוצאות לקובץ test\_y.

# ResNet



(תראו ארקין!!!)



[OBJ]