

## - Report -

### מבנה המודל:

המודל שלנו מתבסס על ארכיטקטורת CNN עם 5 שכבות של קונבולוציה (מ-1 ל-8, מ-8 ל-16, מ-16 ל-32, מ-32 ל-48 ומ-48 ל-64) ו-5 שכבות של max-pooling בגודל 2x2 כאשר הפילטר שלנו הוא בגודל 5x5 וה-stride בגודל 1x1 וה-padding הוא 2.

נדגיש שבחרנו בארכיטקטורה זו לאחר שבדקנו כמה ארכיטקטורות שלמדנו (RNN, שכבות fully connected בלבד ועוד) וקיבלנו תוצאות פחות טובות מהארכיטקטורה הסופית שלנו.

את הסיווג עצמו אנחנו מבצעים באמצעות רשת fully connected הבנויה באופן הבא - הקלט של הרשת הוא 960 נוירונים. החישוב הוא 64, שזה הגודל של שכבת הקונבולוציה האחרונה, כפול 5, שזה גודל הפילטר, כפול 3 – סה"כ 960. יש לנו סה"כ 4 שכבות רשת fully connected כאשר הראשונה היא מ-960 ל-512, השנייה מ-512 ל-256, השלישית מ-256 ל-128 ומ-128 ל-30.

פעולת האקטיבציה של הרשת בה השתמשנו היא ReLU.

על מנת להימנע מ-overfitting הוספנו dropout על 0.2.

### בחירת היפר-פרמטרים:

בדקנו כל מיני אופציות ואלו שהניבו את התוצאות הטובות ביותר הם אלו שבחרנו, להלן הערכים הסופיים –

1. **epochs** – 25

2. **eta** – 0.0005

3. **batch size** – 100

4. **אופטימיזר** – האופטימיזר ADAM הניב את התוצאות הטובות ביותר בפער מכל האחרים.