

# אלגוריתמים בראייה ממוחשבת

046746

Quiz 4

דניאל טייטלמן – 207734088

[Daniel.tei@campus.technion.ac.il](mailto:Daniel.tei@campus.technion.ac.il)

יאיר נחום – 034462796

[nahum.yair@campus.technion.ac.il](mailto:nahum.yair@campus.technion.ac.il)

1. נחשב את גדלי המוצא לפי  $W_{out} = (W_{in} - \text{KernelSize} + 2 * \text{Padding}) / \text{Stride} + 1$ .  
 לדוגמא עבור מוצא השכבה הראשונה  $W_{out} = (64 - 7 - 2 * 2) / 1 + 1 = 62$ .  
 את מספר המקדמים נחשב לפי  $\text{KernelSize} * \text{KernelSize} * \text{InputDepth} * \text{OutputDepth}$  כאשר אנו כאמור לא מתחשבים ב bias.

שכבה	מימדי המוצא	מספר הפרמטרים
INPUT	64x64x3	0
CONV7-16	62x62x16	7x7x3x16
POOL2	31x31x16	0
CONV7-32	29x29x32	7x7x16x32
POOL2	14x14x32	0
FC-3	3	14x14x32x3

2. ניתן לבצע את השיפורים הבאים במבנה הארכיטקטורה והאימון:
- 2.1. להוסיף עוד שכבת/ות FC לפני שכבת ה FC האחרונה עם 3 ניוונים. כך הייצוג הנלמד של הפיצ'רים המורכבים נלמד לוקטור קידוד ולא ישר משכבת הקונוולוציה שלומדת דברים מרחביים (reception fields) יותר. בדומה לעיבוד מידע במוח אחרי שקלטנו מה אנו רואים וצריך לתרגם למילה בדומה ל BOW.
- 2.2. לשים מסווג ליניארי עם softmax במוצא במקום שכבת FC רגילה שעושה פעולות לא לינאריות על המוצא.
- 2.3. לבצע cross validation על ה hyper parameters ולחפש בצורה רנדומלית על כל אחד עם כל אחד, כך מכסים יותר נקודות ב grid של החיפוש עבור ה hyper parameters המתאימים ( learning rate, batch size, dropput). כמו כן, להוסיף שכבות של batch normalization.
- 2.4. להקטין את גדלי הפילטרים אך להוסיף שכבת conv נוספת. כלומר ה  $7 \times 7$  יוחלף ב  $3 \times 3$ . אך נוסיף עוד שכבות conv עם  $3 \times 3$ . כך שה receptive field יישמר. בכך הרשת תהיה פחות מסובכת לאמן ועם נטייה ל overfit (פחות פרמטרים ללמוד). כל 3 שכבות של  $3 \times 3$  ידרשו  $9+9+9=27$  לעומת  $7 \times 7=49$  (בהנחה שבמצטבר העומקים של שכבת המוצא יהיו כמו העומק של שכבת המוצא  $7 \times 7$  אחרת נוסיף הרבה מאוד משקלים)