

## פעולת הקונבולוציה עם גרעין

הפעולה הבסיסית המשמשת לבניית שכבות קונבולוציה אינה המצאה חדשה של מומחי הלמידה העמוקה, אלא פונקציה ידועה בתחום עיבוד התמונה – היא הקונבולוציה של תמונת הקלט עם גרעין (Kernel). בתחום עיבוד התמונה פעולה זו נקראת לעתים הפעלת פילטר, ולה מגוון שימושים שונים כגון טשטוש תמונה, ניקוי רעש, או זיהוי קצוות של אובייקטים.

גרעין הקונבולוציה הוא מטריצה קטנה, אשר ערכיה נקבעים בהתאם לאפקט הרצוי על תמונת הפלט. למשל השימוש במטריצה

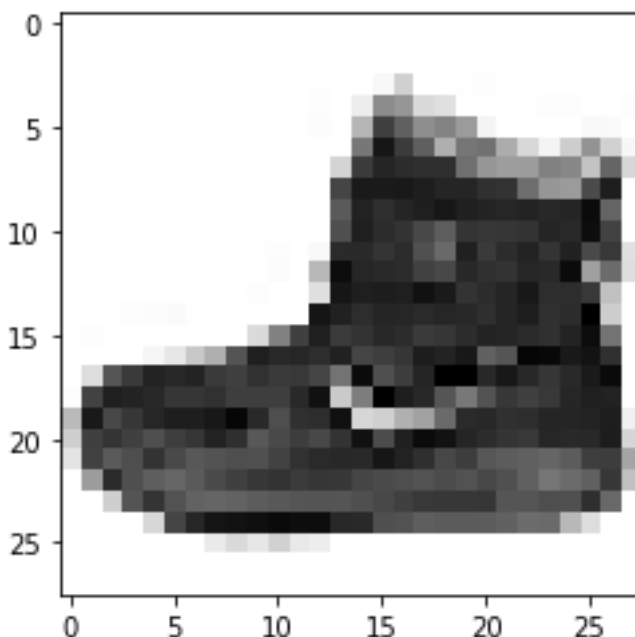
$$K = \begin{pmatrix} \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \end{pmatrix}$$

יוביל לטשטוש התמונה כפי שנראה מיד. להשלמת הדוגמה נתבונן באחת התמונות מהאוסף Fashion-MNIST, תמונה בגודל 28X28 פיקסלים אשר כל אחד מהם מיוצג על ידי מספר טבעי בטווח 0 (לבן) עד 255 (שחור) כמפורט בקטע הקוד הבא.

```
img, _ = next(iter(train_dataloader))
img=torch.squeeze(img)
print(img.dtype,img.size(),sep='\n')
print(img[9:15,9:15])
plt.imshow(img,cmap='Greys');

torch.uint8
torch.Size([28, 28])
tensor([[ 0,  0,  0,  0, 183, 225],
        [ 0,  0,  0,  0, 193, 228],
        [ 1,  3,  0, 12, 219, 220],
        [ 0,  6,  0, 99, 244, 222],
        [ 4,  0,  0, 55, 236, 228],
        [ 0,  0,  0, 237, 226, 217]], dtype=torch.uint8)
```

פלט:



נגדיר בזכרון את גרעין הקונבולוציה הנ"ל, ונחשב בהמשך את הקונבולוציה של תמונת המגף עם גרעין זה.

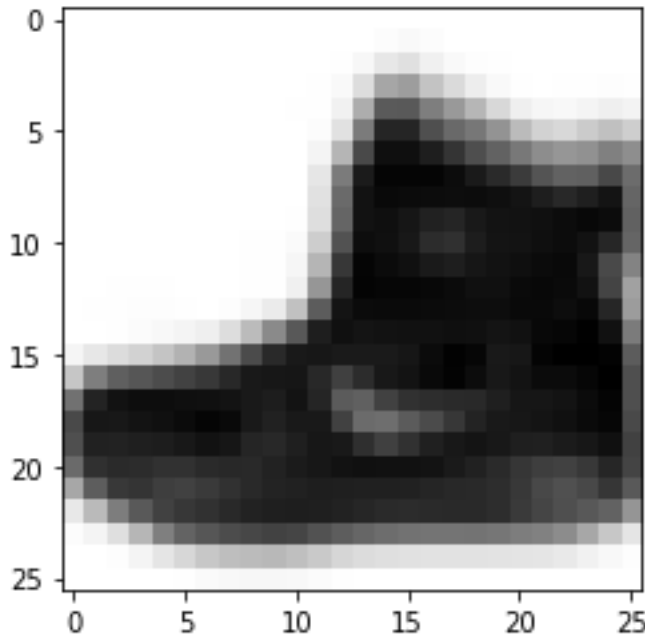
```
kern=torch.tensor([[1/16,1/8,1/16],
                    [1/8,1/4,1/8],
                    [1/16,1/8,1/16]])
```

כעת נסרוק את מטריצת תמונת הקלט משמאל לימין ומלמעלה למטה, ועבור כל תת מטריצה בגודל  $3 \times 3$  נחשב פיקסל יחיד של תמונת הפלט ע"י:

1. כפל חלק התמונה הרלוונטי עם הגרעין איבר-איבר.
2. סכימת התוצאה.

```
output=torch.empty(26,26)
for i in range(26):
    for j in range(26):
        sub_img=img[i:i+3,j:j+3]
        output[i,j]=(sub_img*kern).sum()
plt.imshow(output,cmap='Greys');
```

פלט:



שימו לב בקוד הנ"ל שהמשתנים  $i, j$  מסמנים את הפינה השמאלית העליונה של תת המטריצה  $sub\_img$ . בהתאם ערכם המקסימלי הוא 25, בעוד שאינדקס השורה/עמודה המקסימלי של תמונת הקלט הוא 27. אין זו טעות שכן בכדי ליצור את תת המטריצה  $sub\_img$  יש צורך בקיומם של עוד שני איברים מימין ומתחת ל- $i, j$  במטריצת הקלט. ניתן לראות אם כן שפלט פעולת הקונבולוציה הוא תמונה קטנה יותר, בגודל  $26 \times 26$ . לעתים הקטנת המימד תהיה שימושית עבורנו, ולעתים נרצה לשמור על גודל התמונה המקורי – בהמשך היחידה נדון בשיטות לשליטה במימד הפלט.

אין זה הכרחי שגרעין הפעולה יהיה ריבועי, ראו למשל דוגמה נוספת בה נבצע קונבולוציה של התמונה עם הגרעין  $K = \begin{pmatrix} -1 & 1 \end{pmatrix}$ :

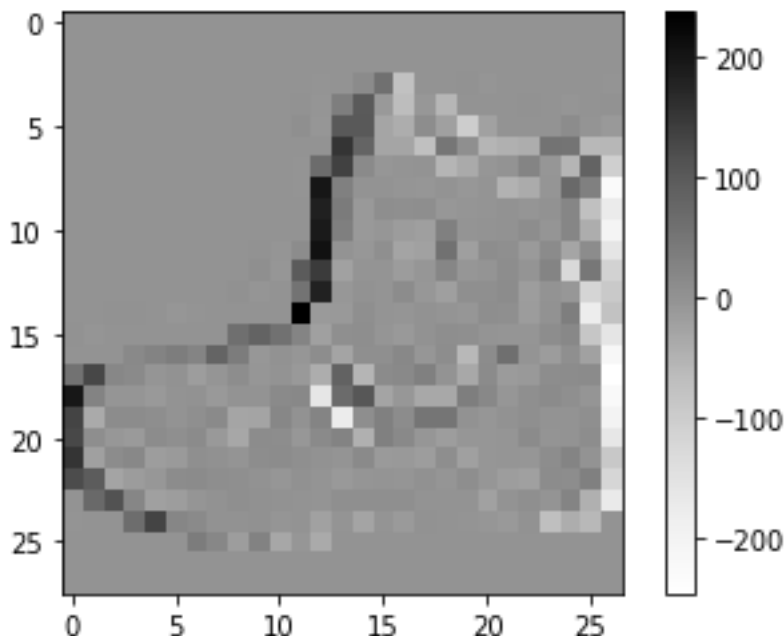
```

kern=torch.tensor([-1,1])
output=torch.empty(28,27)
for i in range(28):
    for j in range(27):
        sub_img=img[i,j:j+2]
        output[i,j]=(sub_img*kern).sum()
print(output.size())
plt.imshow(output,cmap='Greys');
plt.colorbar();

torch.Size([28, 27])

```

פלט:



לעיל ניכר כי רוב מטריצת הפלט מלאה בערכים סביב האפס (אפור) אך יש ערכים גבוהים היכן שבתמונה המקורית היה מעבר מלבן לשחור (בתנועה משמאל לימין) וערכים נמוכים במעבר הפוך. לאור זאת ניתן להגיד שהגרעין הנ"ל הוא למעשה מזהה קצוות אופקי. ראו גם כי פלט פעולת הקונבולוציה הפעם הוא ממימד  $28 \times 27$ , שכן לא בוצע כל חישוב עבור העמודה האחרונה בתמונה המקורית: אין פיקסלים מימין לעמודה זו שבעזרתם ניתן להגדיר את תת המטריצה  $sub\_img$ .

מדוגמאות אלו נקיש את הכלל. בהנתן מטריצת קלט בגודל  $H \times W$  (המייצגת תמונה בעלת ערוץ צבע יחיד: שחור/לבן),

$$X = \begin{pmatrix} x_{1,1} & \cdots & x_{1,W} \\ \vdots & \ddots & \vdots \\ x_{H,1} & \cdots & x_{H,W} \end{pmatrix}$$

ומטריצה קטנה מגודל  $p \times q$ , אשר תשמש כגרעין,

$$K = \begin{pmatrix} k_{1,1} & \cdots & k_{1,q} \\ \vdots & \ddots & \vdots \\ k_{p,1} & \cdots & k_{p,q} \end{pmatrix}$$

תוצאת פעולת הקונבולוציה של  $X$  עם  $K$  היא מטריצה  $Y$  בגודל  $(H-p+1) \times (W-q+1)$  בעלת הערכים הבאים:

$$y_{r,s} = \sum_{i=1}^p \sum_{j=1}^q x_{r+i-1,s+j-1} k_{i,j}$$

שימו לב שלפי נוסחה זו, בכדי לחשב את  $y_{r,s}$ , יש:

1. לחתוך מ- $X$  תת מטריצה בגודל  $p \times q$  אשר הפינה השמאלית העליונה שלה היא  $x_{r,s}$ ,
  2. לכפול תת-מטריצה זו איבר-איבר במטריצת הגרעין,
  3. ולבסוף לסכום את התוצאה.
- מהגדרת הפעולה נבין גם את מימד הפלט: ניתן לחשב את  $y_{r,s}$  כל עוד מתחת ל- $x_{r,s}$  יש עוד  $p-1$  שורות ומימין לו  $q-1$  עמודות.

לסיום פרק זה נציין שבעוד שבתחומי עיבוד התמונה והלמידה העמוקה הפעולה המתוארת לעיל קרויה קונבולוציה, שמה המתמטי הוא קורלציה צולבת (Cross-Correlation) והפעולה המתמטית הקרויה קונבולוציה שונה במעט. למרות זאת נמשיך, כמקובל בתחום, להשתמש בשם קונבולוציה עבור הפעולה הרלוונטית לשימושינו.

### שאלות לתרגול

1. כתבו פונקציית פייתון המקבלת מטריצת קלט וגרעין כלשהו ומחשבת את הקונבולוציה שלהם. על הפונקציה לבדוק את תקינות הקלט, כלומר שבכלל ניתן לבצע את הקונבולוציה (יש לבדוק שגרעין הקונבולוציה אינו גדול מדי).
2. תכננו גרעין  $3 \times 3$  המזהה קצוות אלכסוניים והפעילו אותו על התמונה של המגף הנ"ל.