

רשת נוירונים בארכיטקטורת UNET:

מודל רשת נוירונים שנעשה בו שימוש נרחב באפליקציות של סיווג תמונות ומניפולציות על תמונות וכן בעיות סיווג שונות.

מבנה ה - UNET:

ה - UNET בנוי ממבנה של "U" עם שני חלקים עיקריים:

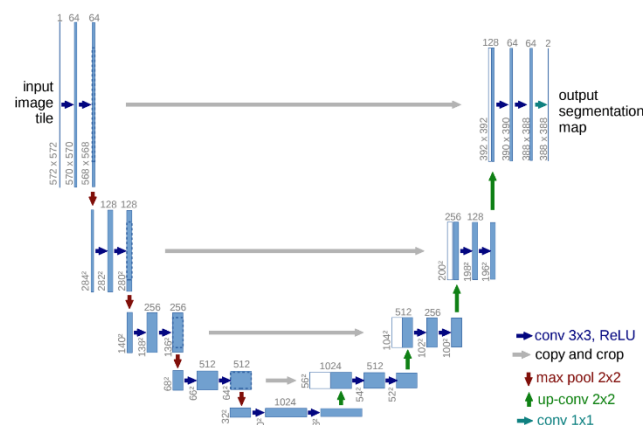


Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

Encoder – צד ירידה:

זהו החלק הראשון של הרשת, שבו הפלט המקורי עובר סדרה של שכבות קונבולוציה, ואחריהן שכבות של מקס-פולינג (max – pooling).

מטרה: להקטין את ממדי הקלט ולשלוף תכונות חשובות. כל שכבת קונבולוציה מיישמת פילטרים על הקלט כדי לחלץ תכונות, והפולינג מצמצם את המימדים.

Decoder – צד עליה:

החלק השני כולל את הפיכת הקלט חזרה למימדים המקוריים שלה. זה נעשה ע"י השימוש בשכבות קונבולוציה שנקראות "שכבות של הפעלה" (upsampling) שמגדילות את הממדים של התמונה.

מטרה: להחזיר את התמונה לממדים המקוריים שלה. תוך שמירה על התכונות שנלמדו בשלב הראשון.

Encoder:

שכבת קונבולוציה:

שכבה שתפקידה להציג ולחדד כל מיני נתונים עיקריים בדאטה, הפחתת מימדי הקלט וסיווג מידע לא יעיל,

מהות השכבה:

השכבה מבצעת פעולת קונבולוציה על נתונים באמצעות סנן (פילטר). הסנן הוא מטריצה קטנה שמחליקה על המידע הנכנס ומביאה ליצירת מפת תכונה חדשה.

איך זה עובד:

בחירת סנן (Filter) – סנן הוא מטריצה בגודל קבוע לדוגמא 3×3 , 5×5 ... כל ערך בסנן הוא משקל שנלמד במהלך האימון.

לדוגמא, אם הסנן הוא בגודל 3×3 , הוא יחפש דפוסים בגודל 3×3 בתמונה.

החלקת הקלט על הקלט – הסנן עובר על הקלט בתנועות מחליקות (strides) ובכל עמדת סלייד (sliding position), הסנן מבצע פעולה שנקראת "קונבולוציה".

פעולת הקונבולוציה היא חיבור מכפלה של הערכים בסנן עם הערכים המתאימים בקלט. כתוצאה מפעולת כפל מטריצות וחיבור המכפלות

יצירת מפת תכונות (feature map) - התוצאה של כל פעולה היא ערך אחד במפת התכונות. התהליך חוזר על עצמו לאורך נתוני הקלט.

התוצאה הסופית היא מפת תכונות חדשה שמתארת את הדפוסים שנמצאו בקלט המקורי.

פרמטרים עיקריים בשכבת קונבולוציה –

גודל הסנן (filter size) – הגודל של הסנן, למשל 3×3 , קובע את האזור בקלט שנבדק בכל שלב של הקונבולוציה.

Stride- הסטפ או המרחק שבו הסנן עובר על הקלט. סטרייד של 1 אומר שהסנן עובר תמונה פיקסל אחרי פיקסל, בעוד ש-2 אומר שהסנן קופץ כל שני פיקסלים.

Padding- הוספת ערכים (לרוב אפסים) סביב התמונה לפני ביצוע הקונבולוציה. זה עוזר לשמור על מידות התמונה או הקלט או להבטיח שהקלט לא יצטמצם מדי.

מספר הסננים (num of filters) – מספר הסננים שברשות שכבת הקונבולוציה. כל סנן לומד לחלץ תכונה שונה. מספר הסננים קובע את מספר הערוצים (scannals) במפת התכונות.

כל סנן עבר על כל הקלט ומבצע קונבולוציה כדי לחלץ תכונה שונה מהקלט. התוצאה מכל סנן נשמרת במפת תכונות נפרדת, ומפות התכונות הללו נערמות יחד כדי ליצור את פלט השכבה.

שכבת מקס פולינג (max pooling):

שכבת מקס פולינג היא רכיב חשוב ברשתות נוירונים במיוחד ברשתות קונבולוציה.

מטרת השכבה היא אלצמצם את ממדי הקלט ולשמור תכונות משמעותיות תוך צמצום מספר הפרמטרים הנדרשים.

איך זה עובד:

שכבת מקס פולינג פועלת על הקלט על ידי חלוקתו לאזורים קטנים ומציאת הערך המקסימלי בכל אזור.

שלבי הביצוע:

חלוקה לאזורים – הקלט מחולק לרשת של תתי אזורים בגודל קבוע הנקראים "פיסות". לדוגמא אזור בגודל 2×2 או 3×3 .

חישוב המקסימום – עבור כל תת-אזור, השכבה מוצאת את הערך המקסימלי מבין כל הערכים באותו תת-אזור.

הערך המקסימלי הזה הופך להיות הערך החדש בקלט המוקטן.

צמצום מימדים – התוצאה היא קלט עם ממדים קטנים יותר, אך עם שמירה על התכונות המשמעותיות ביותר מהקלט המקורי.

יתרונות השכבה: צמצום מימדים, שיפור יכולת שינויי הנתונים, הפחתת עודף מידע.

פרמטרים עיקריים:

גודל החלון (kernel size) – גודל החלון קובע את גודל תת האזור שנחפש בו את הערך המקסימלי.

Stride (קפיצה) – הפרמטר קובע באיזו קפיצה נעבור בין האזורים בקלט. אם ה-stride הוא 2, נחפוף על הקלט בקפיצות של 2 תאים.

Padding (ריפוד) – ריפוד מוסיף גבולות לקלט המקורי. כך שהחלונות יוכלו לכסות את כל הקלט או את חלקו הבלתי מכוסה במקרים של קפיצות שאינן מכסות את כל הקלט.

Valid (ללא ריפוד) או same (ריפוד כך שהממדים יישמרו) הם הגדרות נפוצות.

המודל הבסיסי הוא מחזור של 2 שכבות קונבולוציה ושכבת מקס פולינג.

Decoder:**Transpose convolutions (deconvolutions)**

שכבה בסיסית שמבצעת את השטח של הקלט. היא מאפשרת להגדיל את ממדי הקלט ולהשיב אותו למידותיו המקוריות או קרוב לכך.

Concatenation:

בתהליכים כמו UNET – decoder לא רק משחזר את התמונה אלא גם משתמש במידע שנשמר מה – Concatenation.

פעולות לא ליניאריות:

כדי לשפר את הקלט, ה – decoder כולל פעמים רבות פעולות כמו relu או sigmoid כדי להוסיף את הגמישות הנדרשת ולשפר את האיכות הסופית של הקלט.

סוגים שונים של דה – קונבולוציה:

ConvTranspose2d**פרמטרים עיקריים של השכבה:**

In_channels – מספר ערוצי הקלט שהשכבה מקבלת

Out_channels – מספר ערוצי הפלט של השכבה, כלומר מספר התכנים שהשכבה תפיק.

Kernel_size – גודל המסנן של השכבה. לדוגמא, אם הוא 2×2 , זה אומר שהשכבה תבצע את ההרחבה על בסיס חלון בגודל 2×2 .

Stride – גובה ורוחב הקפיצה של המסנן. לדוגמא, קפיצה של 2 תגרום לכך שהתמונה המוגדלת תהיה פי 2 מהתמונה הקודמת.

Padding – מספר הפיקסלים שיתווספו סביב סביב הקלט כדי לשמור על גודל הקלט או להתאים לגודל הנדרש.

Upsample**פרמטרים עיקריים של השכבה:**

Scale_factor – מפרט את כמות ההכפלה בגובה וברוחב. לדוגמא 2 יכפיל את גובה ורוחב התמונה פי 2.

Mode – מציין את שיטת האינטרפולציה:

Nearest – אינטרפולציה ע"י שימוש בערך הפיקסל הקרוב ביותר.

Bilinear – אינטרפולציה בי לינארית (בהנחה שמדובר בקלטים בעלי ערוץ אחד או שניים, לדוגמא RGB).

Bicubic – אינטרפולציה בי-קיובית. (שיטה יותר מתקדמת)

Align_corners – קובע האם פיקסלי הפינות בקלט המקורי יתאימו לפיקסלים בפינות של הקלט המורחב.

קישור לגיט: <https://paperswithcode.com/method/u-net>