

האוניברסיטה הפתוחה
המחלקה למתמטיקה ולמדעי המחשב

סקירה השוואתית של שיטות לסגמנטציה תלת מימדית של גליומה

עבודה מסכמת זו הוגשה כחלק מהדרישות לקבלת תואר
"מוסמך למדעים" M.Sc. במדעי המחשב
באוניברסיטה הפתוחה
המחלקה למתמטיקה ולמדעי המחשב

על-ידי
אריאל הלחמי

העבודה הוכנה בהדרכתה של ד"ר מיה הרמן

יוני 2023

תוכן עניינים

4.....	רשימת טבלאות ואיורים.....	1
6.....	תקציר.....	2
7.....	מבוא.....	3
10.....	בסיס הנתונים (BraTS).....	4
11.....	הגדרת הבעיה.....	5
11.....	5.1 בהיבט הרפואי.....	
11.....	5.2 בהיבט מדעי המחשב.....	
13.....	6 למידה עמוקה על קצה המזלג.....	
13.....	6.1 רקע.....	
13.....	6.2 רשתות נוירונים.....	
15.....	6.2.1 תהליך הלמידה של רשת נוירונים [12].....	
15.....	6.2.2 Forward Pass.....	
15.....	6.2.3 Activation Function – פונקציית אקטיבציה.....	
16.....	6.2.4 Loss Function (פונקציית הפסד).....	
16.....	6.2.5 Backward Pass.....	
16.....	6.3 רשת נוירונים עמוקה מסוג CNN.....	
16.....	6.3.1 מהי קונבולוציה.....	
17.....	6.3.2 מבנה הרשת.....	
20.....	6.3.3 אימון המודל.....	
21.....	6.4 רשתות נוירונים עמוקות מסוג מקודד – מפענח.....	
23.....	7 שיטות לסגמנטציה גידולי מוח מסוג גליומה בתמונות מוח-MRI.....	
23.....	7.1 3D U-Net קלאסי [11].....	
29.....	7.2 [17] Scale Attention Network.....	
36.....	7.3 [10] Hybrid High-resolution and Non-local Feature Network.....	
43.....	7.4 [23] Modality-Pairing Network.....	
49.....	7.5 [24] Vox2Vox – 3D Generative Adversarial Network (GAN).....	
54.....	8 ניתוח השוואתי בין השיטות השונות.....	
54.....	8.1 מדדי האיכות.....	
54.....	8.1.1 Dice Coefficient.....	

54.....	Hausdorff Distance p95	8.1.2
55.....	השוואת השיטות השונות על בסיס נתוני BraTS ע"פ המדדים שתוארו	8.2
57.....	סיכום ומסקנות	9
58.....	רשימת מקורות	10

1 רשימת טבלאות ואיורים

8.....	איור 1 הדגמת סיגמנטציה של גידול מוח מסוג גלימה בפרוסות MRI במבטים שונים [1]
10.....	איור 2 – הדגמת מחלקות הסיווג של הגידול, באדום NCR, בירוק ED, בצהוב ET. [10]
13.....	איור 3 – רשת נוירונים בקליפת הראייה במוח. [12]
14.....	איור 4 – דיאגרמה של רשת נוירונים מלאכותית. [12]
15.....	איור 5 – רשת נוירונים מלאכותית רבת שכבות. [12]
16.....	איור 6 – פונקציית האקטיבציה ReLU [13]
17.....	איור 7 – המחשה של מעבר פילטר (Kernel) [12]
18.....	איור 8 – הדגמה של ביצוע Stride Convolution [14]
19.....	איור 9 – המחשת ביצוע של Max pooling בגודל 2×2 [12]
20.....	איור 10 – הדגמת מבנה של שכבות Fully connected [14]
22.....	איור 11 – דוגמה לארכיטקטורת רשת מסוג מקודד-מפענח [15]
23.....	איור 12 – מבנה רשת ה- nnU-Net המבוססת ארכיטקטורת 3D U-Net [11]
27.....	איור 13 – תוצאות ויזואליות של nnU-Net על גבי סט הולידציה [11]
29.....	איור 14 – ארכיטקטורת הרשת SA-Net [17]
31.....	איור 15 – מבנה בלוק ResNet סטנדרטי (משמאל) ומבנה בלוק ResSE (מימין). [18]
32.....	איור 16 – מבנה Scale attention block [17]
33.....	איור 17 – הגדרת פונקציית ההפסד Jaccard loss [20]
33.....	איור 18 – הגדרת פונקציית ההפסד Focal Loss [21]
37.....	איור 19 – מבנה רשת HNF-Net [10]
38.....	איור 20 – מבנה רשת Cascaded HNF-Net [10]
41.....	איור 21 – תוצאות סגמנטציה ויזואליות של מודל H2NF-Net [10]
44.....	איור 22 – מבנה רשת Modality-Pairing Network [23]
44.....	איור 23 – פונקציית ההפסד Modality Pairing loss [23]
44.....	איור 24 – פונקציית ההפסד ששומשה לצורך אימון מודל Modality Pairing Network [23]
47.....	איור 25 – תוצאות ויזואליות של Modality Pairing Network על דוגמאות מסט האימון [23]
50.....	איור 26 – מבנה ה-Generator. [24]
50.....	איור 27 – מבנה ה-Discriminator. [24]
54.....	איור 28 – מדד Dice Coefficient [25]
54.....	איור 29 – איור המתאר את מרחק האוסדורף [26]
25.....	טבלה 1 – סיכומי הביצועים של nnU-Net על סט האימון של BraTS 2020. [11]
26.....	טבלה 2 – סיכומי הביצועים של nnU-Net על גבי סט הולידציה של BraTS 2020. [11]
28.....	טבלה 3 – סיכומי הניסויים של nnU-Net על גבי סט המבחן של BraTS 2020. [11]
34.....	טבלה 4 – תוצאות מודל SA-Net על סט האימון של BraTS 2020 [17]
34.....	טבלה 5 – תוצאות מודל SA-Net על סט הולידציה של BraTS 2020 [17]
34.....	טבלה 6 – תוצאות מודל SA-Net על סט המבחן של BraTS 2020 [17]
39.....	טבלה 7 – תוצאות מודל H2NF-Net על סט הולידציה של BraTS 2020. [10]

טבלה 8 – תוצאות מודל H2NF-Net על גבי סט המבחן של תחרות BraTS 2020 [10]	40
טבלה 9 – תוצאות מודל Modality Pairing Network על סט האימון של BraTS 2020 [23]	45
טבלה 10 – תוצאות מודל Modality Pairing Network על סט הולידציה של BraTS 2020 עבור מודל יחיד [23]	45
טבלה 11 – תוצאות מודל ה-Ensemble של Modality Pairing Network על גבי סט הולידציה של BraTS 2020 [23]	46
טבלה 12 – תוצאות מודל ה-Ensemble של Modality Pairing Network על גבי סט המבחן של BraTS 2020 [23]	46
טבלה 13 – תוצאות מודל Vox2Vox על גבי סט האימון כפונקציה של a [24]	52
טבלה 14 – סיכום תוצאות מודל Ensembled Vox2Vox על גבי סטי הנתונים של BraTS 2020 [24]	53
טבלה 15 – סיכום תוצאות המודלים שהוצגו על סט המבחן של BraTS 2020	56

2 תקציר

גליומה הינו גידול מוח ראשי (הנוצר בתאי המוח עצמם). לגליומה יש שני תתי סוגים: Low Grade Glioma ו-High Grade Glioma, כאשר הסוג השני הינו אגרסיבי יותר, גדל במהירות, דורש התערבות כירורגית וסיכויי ההחלמה ממנו אינם גבוהים.

כיום, הכלי המרכזי לאבחון גידולי מוח מסוג גליומה הינו MRI (דימות תהודה מגנטית), זהו מכשיר שבאמצעותו ניתן ליצור מידול תלת מימדי של רקמת המוח. בעזרת מידול זה ניתן לבצע הערכת גודל וזיהוי מיקום הגידול והשפעותיו על הרקמות הסמוכות. סריקת MRI מורכבת ממספר שכלולים אשר מכל אחד מופקת תת סריקה שונה T1, T1c, T1 (עם חומר ניגוד), T2 ו-Flair כאשר השכלולים משלימים אחד את השני ונועדו בכדי להדגיש מאפיינים שונים של רקמות ואיזורים של התפשטות הגידול. כיום, אבחנת הגידול מתבצעת בצורה ידנית ע"י הרופא המנתח או הרדיולוג. אבחנה זו אינה בהכרח מדויקת, אינה מבצעת הפרדה ברורה בין סוגי הרקמות השונות המרכיבות את הגידול ואינה מספקת הערכת נפח כמותית של הגידול. לפיכך, נדרשת מערכת אוטומטית לביצוע סגמנטציה תלת מימדית לגידולי מוח מסוג גליומה על סמך סריקות מוח MRI. מערכת כזו תהווה כלי עזר נוסף שיסייע לרופא מומחה/רדיולוג בפענוח תוצאות בדיקת MRI. בעזרתה, ניתן יהיה לספק מידע אודות מיקום הגידול, הרכבו ומידת התפשטותו בתלת מימד וכן לנטר ולעקוב אחר גידול קיים. כך, נוכל לחסוך במשאב זמן אנוש יקר.

במהלך השנים האחרונות, לאור התפתחות היכולת החישובית וצמיחת תחום הלמידה העמוקה, החלו להשתמש בלמידה עמוקה לצרכי סגמנטציה תלת מימדית של פתולוגיות שונות בסריקות רפואיות ובפרט גידולים מסוג גליומה בסריקות מוח MRI.

בעבודה מסכמת זו בוצעה סקירה השוואתית של גישות מבוססות למידה עמוקה ובפרט רשתות מבוססות ארכיטקטורת מקודד-מפענח לצורך ביצוע סגמנטציה של גידולי מוח מסוג גליומה בסריקות מוח MRI. כלל השיטות שהוצגו נבחנו בתחרות BraTS 2020 על בסיס נתונים זהה המכיל סריקות מוח MRI בארבעת השכלולים עבור נבדקים שונים הסובלים מגידול מוח מסוג גליומה. בעבודה בוצע ניתוח השוואתי בין השיטות השונות המבוסס על מדדי איכות Dice ו-Hausdorff Distance אשר מהווים מדדים מרכזיים לצורך ביצוע סגמנטציה בתמונות רפואיות. לאור הניתוח ההשוואתי נמצא כי שיטות מבוססות ארכיטקטורת מקודד-מפענח הינן יעילות ביותר ומניבות תוצאות טובות בביצוע סגמנטציה תלת מימדית של גידולי מוח מסוג גליומה.

גידולי מוח נחלקים לשני סוגים [1], גידולים ראשים וגידולים משניים. גידולים ראשיים נוצרים בתאי המוח עצמם וגידולים משניים הינם גרורות אשר הגיעו מאיברים אחרים בגוף. הסוג הנפוץ ביותר של גידולי מוח ראשיים נקרא "גליומה" (Glioma). לגליומה יש שני תתי סוגים: Low Grade Glioma ו-High Grade Glioma, כאשר הסוג השני הינו אגרסיבי יותר, גדל במהירות, דורש התערבות כירורגית וסיכויי ההחלמה ממנו אינם גבוהים.

Magnetic Resonance Imaging - MRI (בעברית: דימות תהודה מגנטית) הינו כלי אבחנתי מרכזי עבור גידולי מוח המשמש להערכת גודל, מיקום הגידול והשפעותיו על הרקמות הסמוכות. סריקת MRI מורכבת ממספר שכלולים אשר מכל אחד מופקת תת סריקה שונה: T1, T1c, T1 עם חומר ניגוד, T2 ו-FLAIR (Fluid Attenuation Inversion Recover) כאשר השכלולים משלימים אחד את השני ונועדו בכדי להדגיש מאפיינים שונים של רקמות ואיזורים של התפשטות הגידול. כיום, אבחנת הגידול מתבצעת בצורה ידנית ע"י הרופא המנתח או הרדיולוג. אבחנה זו אינה בהכרח מדויקת, אינה מבצעת הפרדה ברורה בין סוגי הרקמות השונות המרכיבות את הגידול ואינה מספקת הערכת נפח כמותית של הגידול.

סגמנטציה תלת מימדית אוטומטית לגידולי מוח מסוג גליומה עשויה להוות כלי עזר נוסף שיסייע לרופא מומחה/רדיולוג בפענוח תוצאות בדיקת MRI. כך, ניתן יהיה לחסוך במשאב זמן אנוש יקר, יספק פתרון קל וזמין שיסייע לרופא המטפל בניטור ומעקב אחר גידול קיים וישמש ככלי עזר ע"י הקניית מידע אודות מיקום הגידול, הרכבו ומידת התפשטותו בתלת מימד.

בשנים האחרונות, תחום הלמידה העמוקה (Deep Learning) תפס תאוצה בתחומים מגוונים כגון: סיווג תמונות, זיהוי אובייקטים, עיבוד שפה טבעית ועוד [2]. ובפרט בתחום הדימות הרפואי לדוגמה: זיהוי וסיווג סרטן שד בממוגרפיות [3], סגמנטציה של גידולי כבד בצילומי CT [4] וסגמנטציה של גידולי מוח בצילומי MRI [5].

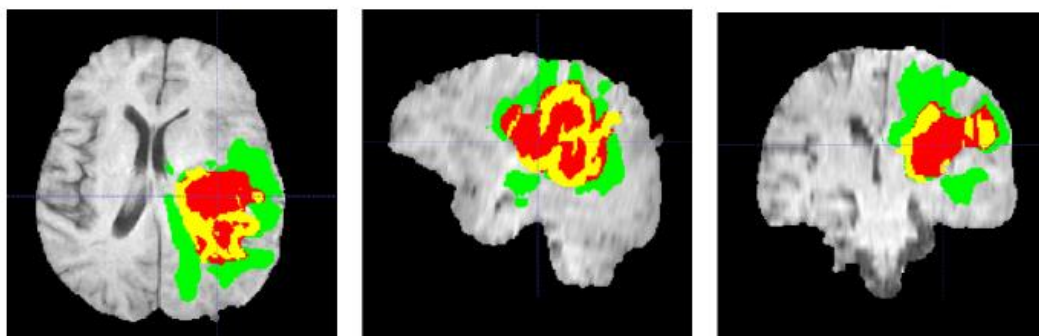
בשנת 2017, הוצע מודל למידה עמוקה מבוסס רשת נוירונים עמוקה (Deep Convolutional Neural Network) לצורך ביצוע סגמנטציה של גידולי מוח מסוג גליומות [6]. במאמר זה הוצגו מספר ארכיטקטורות ספציפיות של רשתות מסוג CNN לצורך ביצוע סגמנטציה זו מימדית של פרוסות MRI מוח. שיטות אלו הניבו תוצאות טובות ביותר בהשוואה לשיטות state-of-the-art נכון לאותה התקופה, הארכיטקטורות השונות שהוצעו הגיעו לרמת דיוק שבין 0.85-0.88 במדד Dice, על סט המבחן (Test set) של התחרות הפופולרית BraTS.

BraTS Multimodal Brain Tumor Segmentation Challenge הינה תחרות המתקיימת אחת לשנה שמטרתה להעריך שיטות שונות לביצוע סגמנטציה של גידולי מוח מסוג גליומה על ידי סיפוק של Dataset המכיל תמונות MRI תלת מימדיות עם תיוגים אשר מהווים את מקור האמת מולו משווים את תוצאות המודלים השונים שמוצעים. כלל התמונות התלת מימדיות סומנו על ידי רופאים מומחים בתחום. סימון הגידול ב-Dataset של BraTS (בעזרת אנוטציות אשר בוצעו ע"י מומחים) מתחלק ל-3 תתי איזורים [1]:

Enhancing Tumor

•

- Peritumoral Edema
- Necrotic/Non-enhancing tumor core
- שלושת תתי האיזורים הנ"ל אוחדו ל-3 תתי איזורים מקוננים :
- (WT) Whole Tumor – אשר מכיל את כלל הגידול
- (TC) Tumor Core – אשר מוכל ב-WT
- (ET) Enhancing Tumor אשר מוכל ב-TC



איור 1 הדגמת סגמנטציה של גידול מוח מסוג גליומה בפרוסות MRI במבטים שונים : (משמאל לימין) אקסיאלי – Axial, סגטלי – Sagittal וקורונלי – Coronal. באיור מודגם WT – המהווה איחוד של שלושת התיוגים השונים (ירוק, אדום וצהוב). TC – הינו האיחוד של התיוג האדום [1]

בתחרות BraTS, ההשוואה ל-Ground Truth מתבצעת בעזרת שני מדדים עיקריים : Dice ו-Hausdorff Distance עבור כל אחד מ-3 תתי האיזורים המקוננים שצוינו לעיל. מדד Dice הינו מדד הערכה הנמצא בשימוש נרחב לצרכי סגמנטציה (וסגמנטציה תלת מימדית בפרט) אשר מחשב את הדמיון בין שתי תמונות (ובפרט סריקות רפואיות). מדד Hausdorff Distance גם הוא מדד פופולרי לצרכי סגמנטציה של תמונות רפואיות המתבסס על חישוב מרחק בין שתי קבוצות ערכים. בספרות קיימים מדדים רבים נוספים המשמשים להערכת ביצועי סגמנטציה כגון : Sensitivity/Specificity, Binary cross entropy ועוד [7]. במסגרת ההשוואה נדון רק ב-Dice ו-Hausdorff Distance Loss.

בשנים האחרונות במסגרת תחרות BraTS הוצעו שיטות רבות מבוססות U-Net (ובפרט 3D U-Net) לסגמנטציה תלת מימדית) לביצוע סגמנטציה של גידולי מוח. U-Net [8] מהווה רשת מסוג מקודד-מפענח. המקודד (Encoder) מוגדר משלב כניסת הנתונים ועד לצוואר הבקבוק. השכבה מצוואר הבקבוק ועד לשכבת יציאת הנתונים מהווה את המפענח (Decoder). באמצעות רשת זו, המקודד יפחית את מימדי הקלט והמפענח (Decoder) יפתח את המידע המקודד למימדים גבוהים יותר תוך כדי שילובו (Skip Connections) עם הקלט הרלוונטי מכל שכבה מצמצמת. כך, ניתן לשחזר את התמונה עם מפת הסגמנטציה המתאימה. היתרון של רשת זו לצרכי סגמנטציה בתמונות רפואיות הינו שתהליך הקידוד מאפשר חילוץ מאפיינים מורכבים מתוך התמונה, ותהליך הפענוח מאפשר בניה של מפות סגמנטציה מתאימות (במימד התמונה המקורי) על סמך המאפיינים שחולצו במימדים הנמוכים.

בעבודה מסכמת זו אבצע סקירה השוואתית לשיטות מודרניות שפורסמו בשנים האחרונות ונבחנו על נתוני BraTS לביצוע סגמנטציה תלת מימדית של גליומות בתמונות מוח-MRI במטרה לזהות ולסווג את שלושת האיזורים הבאים בגידול גליומה :

Enhancing Tumor •

Peritumoral Edema •

Necrotic/Non-enhancing tumor core •

ההשוואה תבוצע על בסיס שני מדדי הערכה מרכזיים אשר הוצגו קודם לכן : Dice ו- Hausdorff Distance.

4 בסיס הנתונים (BraTS)

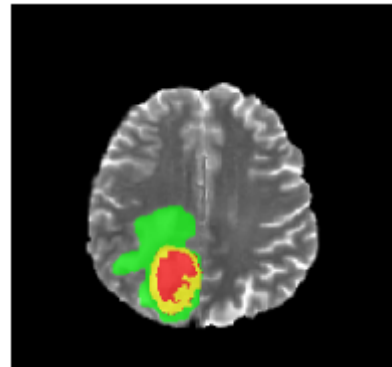
Brain Tumor Segmentation (BraTS) [9] הינה תחרות אשר מתקיימת כבר מעל עשור ומאורגנת ע"י American Society of Radiological Society of North America (RSNA), Image Computing and Computer Assisted Interventions (ASNR) ו-Neuroradiology (MICCAI).

במסגרת תחרות זו מתפרסם כל שנה בסיס נתונים אשר פתוח למשתתפי התחרות וקהל החוקרים אשר מכיל סריקות ראש (mpMRI) multi-parametric magnetic resonance imaging של חולים הסובלים מגידול סרטני מוחי מסוג גליומה. בסיס הנתונים בכל שנה מתבסס על הנתונים מהשנה הקודמת ובדרך כלל מעדכן בנתונים חדשים. כל סריקה מורכבת מ-4 שכלולים שונים:

- (1) T1
- (2) T2
- (3) T1-Gd (T1c) (עם הזרקת חומר ניגוד)
- (4) Flair

בסט האימון מתפרסמת בנוסף מפת סגמנטציה אשר מכילה סיווג של כל ווקסל בסריקה לאחת ממחלקות הסיווג הבאות:

- (1) NCR (מיוצג ע"י הספרה 1) – חלקי הנמק של הגידול
- (2) ED (מיוצג ע"י הספרה 2) – הבצקת שהתפשטה סביב לב הגידול
- (3) ET (מיוצג ע"י הספרה 4) – האיזור שנצבע בהזרקת חומר ניגוד ומראה צפיפות יותר גדולה בהשוואה לרקמות מוח בריאות בשקלול T1-Gd.
- (4) רקמת מוח בריאה ללא גידול/רקע (מיוצג ע"י הספרה 0).



איור 2 – הדגמת מחלקות הסיווג של הגידול, באדום NCR, בירוק ED, בצהוב ET. [10]

מטרת התחרות הינה להשתמש בסט האימון שניתן ובעזרתו לבנות מודל אשר מסוגל לבצע סגמנטציה איכותית (סיווג של כל ווקסל לאחת ממחלקת הסיווג הנ"ל) על גבי סריקות לא מתויגות מסט מבחן.

5 הגדרת הבעיה

5.1 בהיבט הרפואי

ניתן לחלק את הבעיה הרפואית לשתי תתי-בעיות. ראשית, **תיחום הגידול והערכת הנפח שלו**. כיום תיחום זה מתבצע ע"י מומחה אשר נדרש לתחום באופן ידני את הגידול בסריקת המוח-MRI של הנבדק. תיחום ידני של הגידול מצריך מיומנות גבוהה ומשאבי זמן רבים. לפיכך, ישנה חשיבות רבה לפיתוח כלי אוטומטי אשר יהיה מסוגל לבצע תיחום זה בדיוק רב. שנית, **סיווג חלקי הגידול השונים** לרקמות המרכיבות את הגידול (סגמנטציה):

(1) Necrosis – רקמת נמק

(2) Edema – הבצקת אשר התפשטה סביב הגידול

(3) Enhancing Tumor – החלק הפעיל של הגידול.

רקמת מוח בריאה או רקע לא יסווגו לאחת משלוש המחלקות שתוארו לעיל.

סיווג זה חשוב שכן הוא משפיע באופן ישיר על התאמת הטיפול לנבדק. בעבודה זו נעסוק בשתי הבעיות שתוארו, תיחום וסיווג של גידולי מוח מסוג גליומה בסריקות מוח-MRI.

5.2 בהיבט מדעי המחשב

נגדיר את הבעיה באופן הבא: בהינתן בסיס נתונים (Dataset) המורכב מסריקות מוח-MRI של נבדקים ותיוגן על פי סיווג כל ווקסל (פיקסל תלת מימדי) בסריקה לאחד מסוגי הרקמות שהוגדרו לעיל, נדרש לסווג בסריקה חדשה (שלא הופיעה ב-Dataset) כל ווקסל לאחת ממחלקות הסיווג שתוארו לעיל המייצגות רקמות שונות.

באופן פורמלי, ניתן להגדיר סריקת מוח-MRI בודדת כמטריצה ארבע מימדית ממשית X בגודל $(4 \times 155 \times 240 \times 240)$. כאשר 4 מייצג את כמות השכלולים בסריקה, 155 הינו מספר הפרוסות בסריקה התלת מימדית ו-240 מייצג את האורך והרוחב של כל אחת מהפרוסות המרכיבות את הסריקה הכוללת.

מפת הסגמנטציה מכילה סיווג עבור כל אחד מהווקסלים בסריקה המקורית, סיווג כל ווקסל הינו לאחת ממחלקות הסיווג בקבוצה $C = \{0, 1, 2, 4\}$ (מחלקת הסיווג 3 אשר נעשה בה שימוש בעבר אינה בשימוש בבסיסי הנתונים האחרונים של BraTS). מפת סגמנטציה עבור סריקת מוח-MRI

בודדת תוגדר כמערך y בגודל $(155 \times 240 \times 240)$ כאשר $y_{i,j,k} \in C$

לכל $1 \leq i \leq 155, 1 \leq j, k \leq 255$.

על כן, ה-Dataset מורכב מזוגות (X, y) , כאשר X מציין את הסריקה ו- y מציין את מפת הסגמנטציה המתאימה עבור אותה הסריקה.

הגדר מודל f שיתאמן על ה-Dataset ויחזיר את מפת הסגמנטציה הנכונה עבור דגימה חדשה שלא

הופיעה ב-Dataset. באופן פורמלי, המודל נדרש להחזיר את התיוג y^k לקלט X^k

בעזרת הפלט של המסווג $f(X^k, W): \mathbb{R}^{4 \times 155 \times 240 \times 240} \rightarrow \mathbb{C}^{155 \times 240 \times 240}$ עם הפרמטרים W . באופן מתמטי, נדרש למצוא את המודל המסווג f עם הפרמטרים W אשר מביאים למינימום את המרחק בין הסיווג של f לבין הסיווג המהווה Ground Truth.

על מנת להעריך את טיב המודל, נדרש למדוד את ביצועיו בעזרת שני מדדי איכות [11]:

- (1) Dice coefficient.
- (2) Hausdorff distance p95.

מדדים אלו יתוארו בפירוט בהמשך העבודה וכן הסקירה ההשוואתית בין השיטות השונות תתבצע על סמך מדדים אלו.

6 למידה עמוקה על קצה המזלג

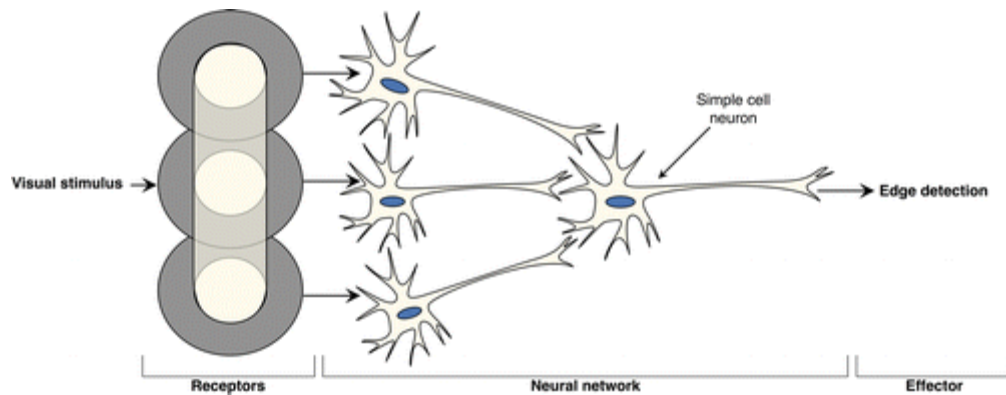
6.1 רקע

למידה עמוקה הינה תת תחום של למידת מכונה אשר תופס תאוצה בשנים האחרונות בתחומים רבים כגון: ראייה ממוחשבת, זיהוי קולי, עיבוד שפה טבעית וכו'. תאוצה זו נבעה מכך שכוח העיבוד של המחשבים התעצם בצורה משמעותית בשנים האחרונות וכמות גדולה של מידע (data) הפכה ליותר זמינה.

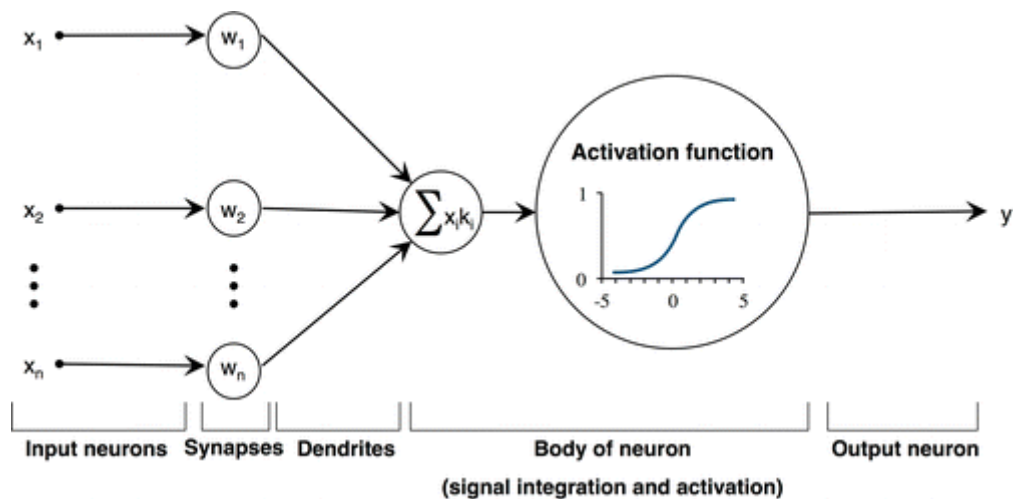
שיטות למידה עמוקה יוצרות מיפוי בין קלט (בתצורתו הגולמית) לפלט (כגון סיווג למחלקה מסוימת). בניגוד לשיטות למידת מכונה קלאסיות, בלמידה עמוקה לא מתבצע שלב מקדים של חילוף התכונות. כמו כן, בלמידה עמוקה המודלים מורכבים משכבות מרובות של רשתות נוירונים מלאכותיות. מודלים כאלה מקבלים השראה מעולם הביולוגיה וממערכת הנוירונים של יצור חי בפרט.

6.2 רשתות נוירונים

במוח של יצור חי, הנוירונים (תאי עצב) מעבירים מידע ביניהם דרך סינפסות (איזור המפגש בין תאי עצב שונים) כימיות או חשמליות. האותות אשר מגיעים לתא עצב, מופצים דרך האיזור הסינפסי שלו אל הדנדריטים (אשר הינם שלוחות של תאי העצב אשר מקבלים מידע מתאים אחרים ומעבירים אותו אל גוף התא) ומשם אל ה-soma (גוף התא). במידה והגירוי עובר סף מסוים (Threshold) התא משחרר אות (Signal) אל האקסון (הסיב העצבי של התא) דרך סינפסות אל תאי עצב שכנים כמודגם באיור 3. באיור 4 מתוארת רשת נוירונים מלאכותית.

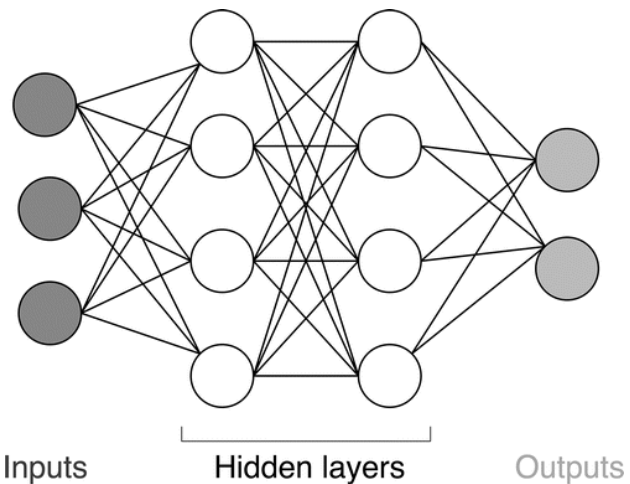


איור 3 - רשת נוירונים בקליפת הראייה במוח אשר מסוגלת לזהות קצוות מתוך קלט המגיע מרשתית העין (העיגולים האפורים באיור הינם קולטנים של הרשתית). כאשר העיגולים הפנימיים של הקולטנים מופעלים בו זמנית ה-simple cell neuron משלב את שלוש האותות שקיבל ומשדר סיגנל "Edge detection" (כלומר, הרשת זיהתה קצה). [12]



איור 4 – דיאגרמה של רשת נוירונים מלאכותית. [12]

ברשת נוירונים מלאכותית כל נוירון מלאכותי מממש מודל סיווג פשוט אשר פולט סיגנל כלשהו על סמך סכום ממושקל של קלטים שקיבל. סכום זה עובר דרך פונקציית אקטיבציה (אשר אינה בהכרח ליניארית ובכך מאפשרת לרשת לחשב בעיות מורכבות יותר) אשר התוצאה שלה מהווה את הפלט של הנוירון. שילוב של כמות רבה של נוירונים מלאכותיים כאלו נקראת רשת נוירונים מלאכותית. המשקלים של הרשת הינם פרמטרים המתעדכנים ע"י שימוש באלגוריתמי למידה אשר מתבססים על זוגות של אותות קלט והפלט הרצוי עבורם. שיטת למידה זו מזכירה את התנהגות המוח אשר מתבסס על גירוי מחיישנים חיצוניים (עיניים, אוזניים וכו') על מנת ללמוד לבצע פעולות מסוימות. למרות שנוירון מלאכותי יחיד הינו פשוט למדי, ארכיטקטורת רשת נוירונים מלאכותית המורכבת משכבות רבות של נוירונים (נקראת גם multilayer perceptron) יכולה לייצג פונקציות לא ליניאריות מורכבות מאוד. Multilayer perceptron בנוי משכבות רבות של נוירונים, כאשר פלט של כל שכבה הוא הקלט לשכבה הבאה. השכבה הראשונה נקראת שכבת הקלט אשר מייצגת קלט כלשהו (כגון עוצמת פיקסלים). השכבה האחרונה נקראת שכבת הפלט אשר התוצאה שלה מייצגת את הסיווג למחלקה. שכבות הביניים בין שכבת הקלט לשכבת הפלט נקראות שכבות נסתרות (Hidden Layers). שכבות אלו לא מניבות פלטים סופיים אלא נועדו לבצע חישובים אשר מהווים את תהליך עיבוד המודל.



איור 5 – רשת נוירונים מלאכותית רבת שכבות. ברשת זו ישנה שכבת נוירוני קלט, מספר רב של שכבות נסתרות ושכבת פלט אחת. [12]

6.2.1 תהליך הלמידה של רשת נוירונים [12]

אלגוריתם הלמידה של רשת נוירונים הוא להלן :

- (1) מתבצעת הרצה של הרשת על קבוצת קלטים (batch), הקלט מוכפל בכל שכבה במשקלים מתאימים (פרמטרים של הרשת). כל נוירון מקבל כקלט את סכום הפלטים מהשכבה שלפניו, מבצע פונקציית אקטיבציה מסוימת על הסכום ומעביר אותו לשכבה הבאה. החישוב מסתיים כאשר מקבלים תוצאה משכבת הפלט (פרדיקציה)
- (2) מתבצע חישוב של פונקציית הפסד (loss function) על גבי הפלט המצוי ביחס לפלט הרצוי.
- (3) מתבצע חישוב של גרדיאנט פונקציית ההפסד ביחס לכל משקל (פרמטר) של הרשת, כלומר גזירה של פונקציית ההפסד לפי כל פרמטר של הרשת.
- (4) לאחר חישוב הגרדיאנט, מתבצע תיקון של משקלי המודל בעזרת אלגוריתם הנקרא gradient descent (ירידה בכיוון הגרדיאנט). המטרה היא לבצע עדכון של פרמטרי המודל בצורה כזו שתקטין את ערך פונקציית ההפסד.
- (5) מתבצעת חזרה על שלבים 1-4 עד להתכנסות המודל.

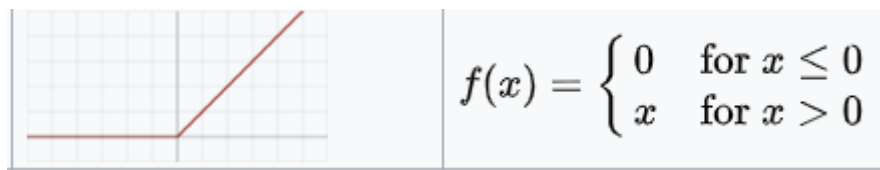
6.2.2 Forward Pass

זהו השלב הראשון באלגוריתם הלמידה בו מתבצעת הרצה של הרשת על דוגמת קלט מסוימת. כל נוירון במודל מקבל כקלט סכום משוקלל של פלטים מנוירונים בשכבה קודמת (או את הקלט המקורי ישירות בשכבה הראשונה).
נוירון המקבל n קלטים $x_1, x_2, \dots, x_n$ ובעל משקלים $w_1, w_2, \dots, w_n$ בהתאמה. הנוירון יחשב את הסכום המשוקלל $\sum_{i=1}^n x_i w_i$ ויפעיל על הסכום פונקציית אקטיבציה $f(\sum_{i=1}^n x_i w_i)$. תוצאה זו תהווה קלט לנוירונים בשכבה הבאה. הפלט של השכבה האחרונה הינה הפרדיקציה של המודל עבור הקלט ההתחלתי.

6.2.3 Activation Function – פונקציית אקטיבציה

פונקציית האקטיבציה הינה פונקציה מוגדרת מראש המופעלת על סכום הקלטים הממושקלים של נוירון מסוים. הפלט של פונקציה זו הוא למעשה הפלט של הנוירון לנוירונים בשכבה הבאה (או פלט סופי במידה ומדובר בשכבה האחרונה).

דוגמה לפונקציית אקטיבציה אשר נעשה בה שימוש רחב בהמשך העבודה הינה פונקציית ReLU (Rectified Linear Unit) – פונקציה שערכה הקבוע 0 משמאל לציר ה-y ופונקציית הזהות מימין לו.



איור 6 - פונקציית האקטיבציה ReLU [13]

6.2.4 Loss Function (פונקציית הפסד)

פונקציית ההפסד היא פונקציה שבאמצעותה ניתן למדוד בצורה כמותית את המרחק בין הפרדיקציה של הרשת לבין התוצאה הצפויה האמיתית. בדרך כלל לוקחים את הפלט של השכבה האחרונה של הרשת, ועליו מפעילים פונקציית softmax. פונקציה זו ממירה את פלט הרשת להתפלגות על מחלקות הסיווג השונות. כאשר ההסתברות הגבוהה ביותר מייצגת את הפרדיקציה של הרשת. לאחר חישוב softmax על פלט הרשת, מפעילים את פונקציית ההפסד על ההסתברויות שהתקבלו על מנת לחשב שגיאה כוללת של המודל על הקלט.

6.2.5 Backward Pass

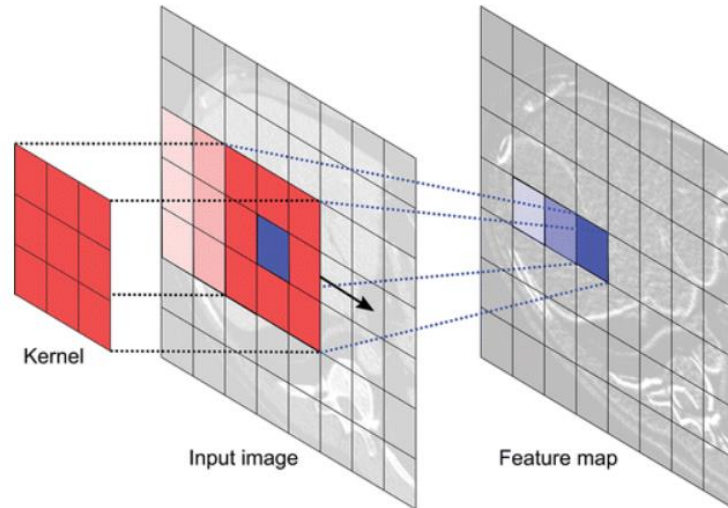
לאחר חישוב ערך פונקציית ההפסד, עושים מעבר בכיוון ההפוך (משכבת הפלט אחורה) על מנת לתקן את כלל פרמטרי הרשת (המשקלים) על סמך ערך השגיאה שהתקבל. על מנת לבצע תיקון זה, מחשבים את הגרדיאנט של פונקציית ההפסד (מחשבים את הנגזרת של פונקציית ההפסד לפי כל פרמטר ברשת). המטרה של חישוב זה הינה לגלות מהו הכיוון (חיובי או שלילי) שבו צריך לתקן את המשקלים על מנת למזער את ערך פונקציית ההפסד. אלגוריתם זה הינו אלגוריתם אופטימיזציה הנקרא Gradient Descent (ירידה בכיוון הגרדיאנט). חישוב הנגזרות החלקיות של פונקציית ההפסד הינו מורכב מפני שהמודל מורכב מהרכבה של פונקציות רבות (אשר חלקן לא ליניאריות, למשל פונקציית האקטיבציה ReLU), לכן על מנת לפשט את חישוב הנגזרות החלקיות משתמשים בכלל השרשרת (Chain Rule) שלפיו הנגזרת הכללית הינה שרשרת של מכפלות של נגזרות: $f(g(x)) = f'(g(x)) * g'(x)$. לאחר חישוב נגזרת פונקציית ההפסד ע"פ משקל מסוים, מעדכנים את ערך המשקל הישן בהתאם ומתקבל ערך חדש המקטין את ערך פונקציית ההפסד. התהליך מסתיים לאחר שכל המשקלים ברשת עודכנו. [12]

6.3 רשת נוירונים עמוקה מסוג CNN

6.3.1 מהי קונבולוציה

קונבולוציה הינה פעולה ליניארית שבה מבצעים מכפלה סקלארית בין פילטר (נקרא גם גרעין – Kernel) לבין סביבה כלשהי של נקודה בקלט (סביבת השכנים מוגדרת לפי מימדי הפילטר). תוצאת המכפלה הסקלארית הינה תוצאת הקונבולוציה של הפילטר על נקודה ספציפית בקלט. בדרך כלל, מימדי הקלט (לדוגמה, תמונה דו מימדית) בעלי רוחב וגובה זהים ועומק משתנה $(N * N)$ כאשר D מייצג את עומק הקלט (נקרא גם ערוצים, עבור תמונת RGB יהיו 3 ערוצים), כל תא בקלט הינו מספר המייצג עוצמה של פיקסל.

על מנת להפעיל קונבולוציה בין פילטר לתמונת קלט כזו, הפילטר חייב להיות בעל אותה כמות ערוצים D כמו הקלט, אך הרוחב והגובה ניתנים לבחירה חופשית (בדרך כלל מימדי פילטר נעים בין $3 * 3 * D$ לבין $5 * 5 * D$), כל תא בפילטר הינו משקל (פרמטר נלמד) של רשת ה-CNN שתפורט בהמשך.



איור 7 – המחשה של מעבר פילטר (Kernel) בגודל $3 * 3 * 1$ על תמונת קלט בגודל $N * N * 1$ [12]

בדוגמה הספציפית הנ"ל נקודת העניין בקלט היא הנקודה המסומנת בכחול, הקונבולוציה של הפילטר סביב נקודה זו מוגדרת כסכום המכפלות של כל תא בפילטר עם התא המתאים בתמונת הקלט. התוצאה הינה סקלאר אשר מיוצג ע"י המשבצת הכחולה ב-feature map. חוזרים על פעולת הקונבולוציה עבור כל תא קלט ע"י הזזת הפילטר בכל שלב וביצוע החישוב הנ"ל. התוצאה של כל קונבולוציה נרשמת לתא המתאים ב-feature map. פעולת הקונבולוציה של פילטר על תמונת קלט מחקה הוצאת תכונות ויזואליות מעניינות מתוך התמונה כגון קצוות ופינות, ולכן ע"י למידת הפרמטרים (משקלי הפילטרים) ניתן לייצר פילטרים המסוגלים לזהות תבניות יותר ויותר מורכבות. הקונבולוציות הינן מרכיב מרכזי ברשתות CNN (מכאן נגזר שמו של המודל).

6.3.2 מבנה הרשת

6.3.2.1 Convolutional layers - שכבות קונבולוציה [14]

שכבות קונבולוציה הינן השכבות העיקריות ב-CNN. שכבה כזו מורכבת ממספר פילטרים מסוים Y , בגודל $H_F * W_F * D_F$ (כאשר הראשון הוא גובה הפילטר, W הוא רוחב הפילטר ו- D הוא עומק/כמות הערוצים בפילטר). הקלט לשכבת הקונבולוציה הינו feature map בגודל $H_I * W_I * D_I$ כאשר בהכרח $D_I = D_F$, כלומר עומק הקלט הוא כעומק הפילטרים (אחרת לא ניתן לבצע קונבולוציה בין הפילטר לקלט).

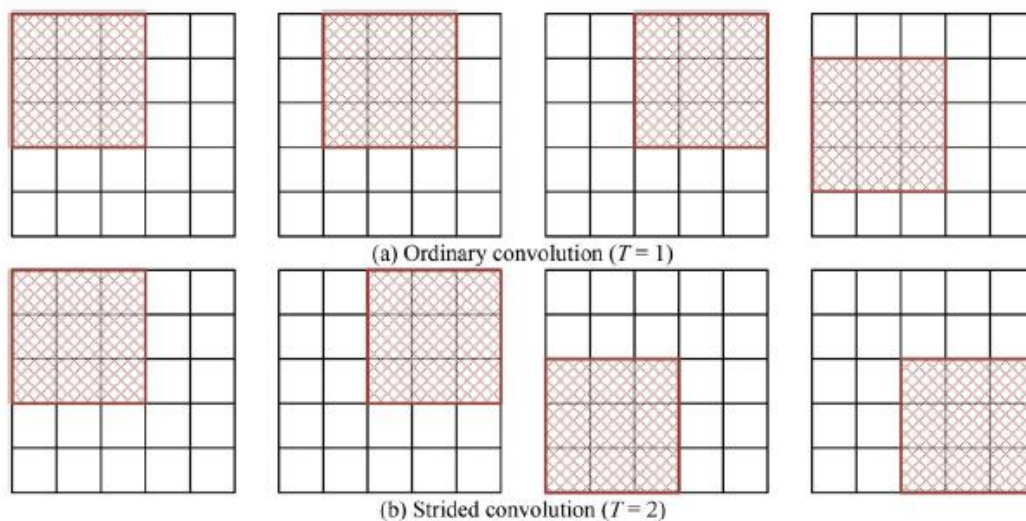
מבצעים קונבולוציה בין כל פילטר לקלט בצורה שתוארה לעיל, תוצאת הקונבולוציה עבור כל פילטר הינה סקלאר בשכבה המתאימה ב-feature map הפלט של שכבת הקונבולוציה. תוצאת הקונבולוציות של הפילטר הראשון עם הקלט יצרו את השכבה הראשונה ב-feature map הפלט, תוצאת הקונבולוציות של הפילטר השני עם הקלט יצרו את השכבה השנייה ב-feature map הפלט

וכן הלאה. לכן תחת ההנחה כי בשכבת הקונבולוציה היו Y פילטרים, הפלט של שכבה זו יהיה feature map בגודל $H_O * W_O * Y$ [14].

קונבולוציה אינה חייבת להתבצע עבור כל תא בקלט, ניתן להגדיר דילוג (Stride) אשר את גודלו נסמן ב- T .

$T = 1$ – מקרה זה מייצג קונבולוציה רגילה שבה גודל הדילוג הוא 1. כלומר, עוברים על כל תא בקלט.

$T = 2$ – מקרה זה מייצג קונבולוציה עבור כל תא שני (מדלגים על כל תא אי זוגי).



איור 8 – הדגמה של ביצוע Stride Convolution [14]

בנוסף, על מנת שהקונבולוציה תפעל גם על קצוות הקלט, ניתן לרפד את הקלט (Padding) בשכבות אפסים מסביב. עובי הריפוד יסומן ב- N .

לאחר הגדרת N ו- T ניתן לחשב במדויק את מימדי feature map הפלט של שכבת הקונבולוציה:

$$H_O = 1 + \frac{H_I - H_F + 2 * N}{T}$$

$$W_O = 1 + \frac{W_I - W_F + 2 * N}{T}$$

כל סקלאר המייצג תוצאת קונבולוציה של פילטר על הקלט עובר דרך פונקציית אקטיבציה אשר הפלט שלה יהיה ב-feature map הפלט לשכבה הבאה.

כלל המשקלים בפילטרים הינם פרמטרים אשר ערכם מתעדכן בתהליך הלמידה של הרשת. כל פילטר לומד לזהות תבניות מסוימות בקלט.

6.3.2.2 שכבות Down sampling – Pooling

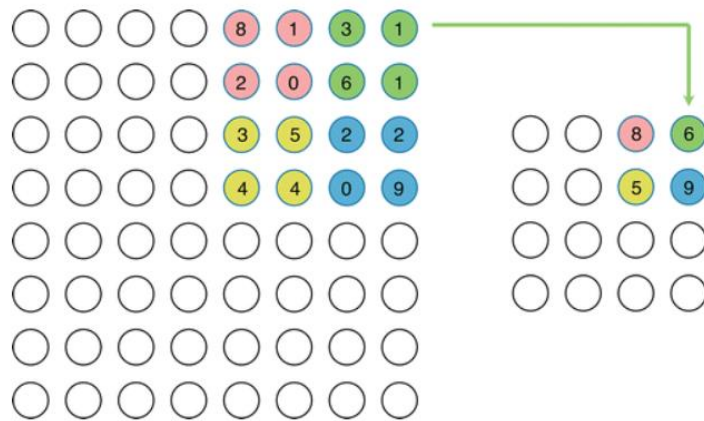
בדרך כלל, השכבה הבאה אחרי שכבת קונבולוציה הינה שכבת pooling. מטרת שכבה זו הינה לצמצם את מימדי ה-feature map שהתקבל בקלט משכבת הקונבולוציה וזאת משתי סיבות עיקריות:

- צמצום המימדים מקטין את רגישות המודל לתזוזות קטנות של אובייקטים שונים.

- הקטנת כמות הזיכרון שהמודל מצריך וצמצום החישובים בשכבות הבאות.

תהליך ה-pooling מתבצע בצורה הבאה :

- מגדירים את מימדי ה-pooling להיות $n * n$ עבור n כלשהו.
 - עבור כל איזור בגודל $n * n$ ב-feature map הקלט, בוחרים מספר אחד אשר ייצג איזור זה ב-feature map הפלט.
- השיטות הפופולריות ביותר לבחירת נציג כזה הן :
- Max pooling – בוחרים את הערך הגבוה ביותר באזור, ערך זה יהיה הנציג ב-feature map הפלט עבור איזור זה. ניתן לראות דוגמה ל-Max pooling באיור 9
 - Average pooling – מבצעים מיוצע על כל הערכים באזור, הערך הממוצע יהיה הנציג ב-feature map הפלט עבור איזור זה.



איור 9 – המחשת ביצוע של Max pooling בגודל $2 * 2$. ניתן לראות כי עבור כל איזור בוחרים את הערך המקסימלי אשר עובר לתא המתאים ב-feature map הפלט. ניתן לראות כי פעולה זו הקטינה את ה-feature map פי 2 בכל מימד [12]

Dropout 6.3.2.3

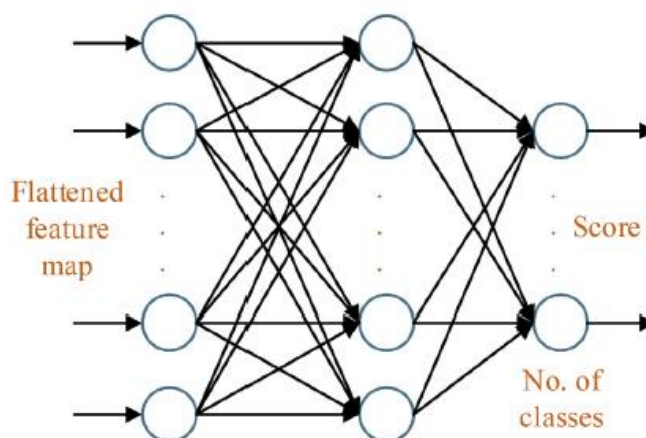
שכבת Dropout הינה שכבה אופציונלית שניתן לשים לפני convolutional layer או fully-connected layer (שנפרט עליה בהמשך). עבור שכבת dropout מגדירים היפר-פרמטר p אשר קובע את ההסתברות "לזרוק" נוירון בשכבה העוקבת. הכוונה היא שבשלב האימון, עבור כל איטרציית forward pass של הרשת, קובעים בהסתברות p עבור כל נוירון בשכבה הרלוונטית האם לנתק את קשרי הקלט שלו מהשכבה הקודמת ואת קשרי הפלט לשכבה העוקבת. בצורה זו כל נוירון "שנזרק" לא משפיע על תוצאת הרשת באותה איטרציה.

כאשר מתבצעת הרצה של הרשת על סט מבחן, כל הנוירונים חוזרים להיות פעילים אך מתבצע downscale לכלל המשקלים בשכבה הרלוונטית (ע"י הכפלה ב- p) בכדי לפצות על ה-dropout בשלב האימון.

הסיבה העיקרית לשימוש בשכבות dropout הינה למנוע התאמת יתר (overfitting) על סט האימון. התאמת יתר הינה מצב שבו הרשת למדה את סט האימון בצורה מיטבית אך לא למדה לבצע סיווג כראוי למקרה הכללי. רשת אשר עברה התאמת יתר, תתקשה לבצע סיווג מדויק של דוגמאות חדשות אשר היא לא התאמנה עליהן.

6.3.2.4 Fully Connected Layer

לאחר סיום המעבר של הקלט בכל שכבות הקונבולוציה (וה-pooling) של הרשת, מתבצע שיטוח של ה-feature map שהתקבל כפלט לווקטור חד מימדי. וקטור זה נכנס כקלט לשכבה הנקראת fully connected layer, זוהי שכבת נוירונים חד מימדית אשר כל נוירון קלט שלה מחובר לכל ערכי הפלט מהשכבה הקודמת (ועל כן שמה fully connected). ניתן לבנות כמה שכבות fully connected אחת אחרי השנייה (אשר כל נוירון בשכבה מחובר לכל הפלטים מהשכבה הקודמת). מספר הנוירונים בשכבת ה-fully connected האחרונה הינו בדרך כלל כמספר מחלקות הסיווג הסופיות של הרשת. הפלטים של שכבת ה-fully connected האחרונה מנורמלים ע"י פונקציית softmax וכך מתקבל וקטור התפלגות ההסתברויות של שיוך הקלט המקורי למחלקות הסיווג השונות.



איור 10 – הדגמת מבנה של שכבות Fully connected. השכבה השמאלית מייצגת את הווקטור המשוטח (חד מימדי) של ה-feature map שהתקבל משכבת הקונבולוציה/pooling האחרונה, השכבה האמצעית הינה שכבת fully connected והשכבה האחרונה הינה שכבת fully connected בעלת מספר נוירונים הזהה לכמות מחלקות הסיווג הסופי. [14]

6.3.3 אימון המודל

6.3.3.1 Data

אימון מודל כדוגמת CNN מתבסס על מידע מתויג (למשל תמונות וסיווג שלהן למחלקה כלשהי). כמות המידע הנדרשת על מנת לאמן מודל בצורה טובה משתנה בהתאם למורכבות המשימה אותה המודל צריך לבצע (כדוגמת זיהוי גידולים מסוג גליומה בסריקות מוח MRI). בדרך כלל, על מנת להשיג תוצאות מדויקות נדרשת בדרך כלל כמות גדולה של מידע מתויג. לעיתים, השגת מידע מתויג באיכות גבוהה הינה פעולה האורכת זמן רב וצורכת הרבה כסף ומשאבים מפני שקודם כל נדרש להשיג/לייצר את המידע ולאחר מכן לבצע תיוג בצורה ידנית.

במידה וישנה רק כמות קטנה יחסית של מידע, עדיין ניתן להגיע לתוצאות טובות בעזרת פעולת אוגמנטציה (העשרת מידע). פעולה זו נועדה בכדי להגדיל בצורה מלאכותית את דוגמאות האימון מתוך ה-Data הקיים. אוגמנטציה מתבצעת על ידי ביצוע טרנספורמציות רנדומליות על המידע המקורי (כגון: היפוך, סיבוב בזווית קטנה מסוימת, הזזה, Scaling וכו').

באופן כללי, נהוג לחלק את המידע לשלושה סטים: סט אימון (training set), סט ולידציה (validation set) וסט מבחן (test set). סט האימון נועד לצורך אימון הרשת וקביעת הפרמטרים (משקלים). החלוקה מתבצעת בצורה כזו שבה סט האימון הינו הגדול ביותר ואילו סט הולידציה וסט המבחן הינם קטנים משמעותית (לדוגמה, 80% לסט האימון, 10% לסט הולידציה ו-10% לסט המבחן). ניתן גם לבצע אימון בצורת K-fold cross validation שבו מחלקים את המידע ל-K

קבוצות שונות ומבצעים אימון K פעמים כאשר בכל פעם קבוצה אחרת מתוך K הקבוצות משמשת כסט המבחן. בשלב האימון מתבצעת הרצה של הרשת על דוגמאות שונות מסט האימון ולאחר מכן תיקון המשקלים על סמך המרחק בין פלט המודל לבין התוצאה הרצויה – Ground Truth. סט הוולידציה נועד בכדי לנטר את ביצועי המודל בזמן האימון (ובעזרתו ניתן לבצע תיקונים כלליים במודל כגון שינוי היפר-פרמטרים כדוגמת מספר שכבות נסתרות, מספר משקלים בפילטר, הסתברות Dropout וכו'). לאחר סיום תהליך האימון וקביעת המשקלים באופן סופי ניתן להעריך את ביצועי המודל על סט המבחן. בד"כ מתבצעת מדידה של ביצועי המודל על סט המבחן בסוף המחקר על מנת לדווח את הביצועים הסופיים של המודל על ה-dataset.

6.3.3.2 דרישות טכניות

אימון של מודל Deep Learning מודרני כגון CNN בדרך כלל דורש כוח עיבוד רב. ההצלחה של מודלים אלו התרחשה בעיקר בזכות פיתוח רכיבי חומרה המאפשרים כמות גדולה של חישובים מקביליים. רכיבים אלו נקראים GPU (Graphics Processing Unit) או בעברית מאיץ גרפי. רכיבי חומרה אלו נוצרו במקור בעיקר לצורך הרצת משחקי מחשב הדורשים חישובים גרפיים כבדים ורבים. בשנים האחרונות רכיב ה-GPU הפך לרכיב נחוץ ומרכזי בתהליך האימון של מודלי Deep Learning רבים בכך שמאפשר אימון בזמן קצר משמעותית ביחס לאימון המתבסס על CPU בלבד. שימוש ב-GPU על פני CPU (Central Processing Unit – המעבד המרכזי של המחשב) מקצר את זמני האימון ובכך מאפשר לאמן מודלים המורכבים ממיליוני פרמטרים בשעות/ימים בודדים לעומת שבועות/חודשים ארוכים.

בנוסף לחומרה, יש צורך גם בסביבת הרצה תוכניתית שעל גביה ירוץ המודל. ישנן ספריות מוכרות לבניית רשתות נוירונים רבות שכבות כגון: TensorFlow, PyTorch ועוד.

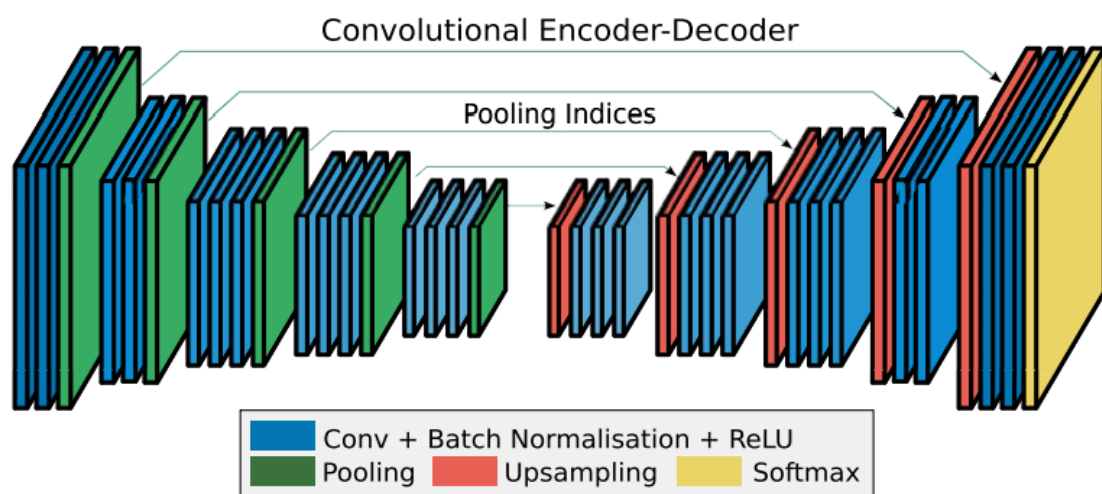
6.4 רשתות נוירונים עמוקות מסוג מקודד – מפענח

רשתות נוירונים עמוקות מסוג מקודד מפענח [15] כמודגם באיור 11 הינן רשתות המורכבות משני חלקים:

(1) מקודד (Encoder) – החלק הראשון ברשת, המורכב משכבות קונבולוציה שמטרתן ליצור סט של feature maps על פי כמות הפילטרים, שכבות Pooling שנועדו על מנת לצמצם את מימדי הקלט ופונקציית אקטיבציה. מסלול ה-Encoder דומה למסלול המצמצם ברשת CNN סטנדרטית, בכך שהוא מורכב משכבות קונבולוציה ושכבות Pooling המשמשות להורדת מימדי הקלט, וכן נעשה שימוש בפונקציית אקטיבציה.

(2) מפענח (Decoder) – החלק השני ברשת (לאחר ציור הבקבוק בחלק המקודד), בכל רמה במסלול זה מתבצעת פעולת Upsampling על גבי ה-Feature map שהתקבל מהשכבה הקודמת (העלאת הקלט למימדים גבוהים יותר) ולאחר מכן מתבצעות פעולות קונבולוציה בדומה לפעולות הקונבולוציה אשר מתבצעות במסלול המקודד. השכבה האחרונה בנתיב המפענח הינה שכבת Softmax אשר הפלט שלה הינו בגודל הקלט המקורי והוא מכיל הסתברות שייכות למחלקת סיווג עבור כל פיקסל בנפרד (pixelwise classification).

רשתות מסוג מקודד-מפענח מניבות פלט אשר זהה במימדיו לממדי הקלט ולכן בעזרת רשת מסוג זה ניתן לבצע סיווג של כל פיקסל בנפרד ולכן רשתות מסוג זה הינן בשימוש נרחב לצורכי סגמנטציה הכוללת חלוקה של איזורים בתמונה למחלקות סיווג שונות (בניגוד לרשת CNN סטנדרטית אשר מסווגת את התמונה למחלקת סיווג כמקשה אחת). בנוסף, רשתות מסוג מקודד מפענח הינן בעלות מבנה גנרי המאפשר התאמה לבסיסי נתונים שונים במימדים שונים על ידי שינוי קל של פרמטרים. רשת מסוג מקודד-מפענח מאפשרת חילוץ של מאפיינים מורכבים מהתמונה במימדים שונים לאורך הקידוד ועל סמך מאפיינים אלה לבנות מפת סגמנטציה איכותית בשלב הפענוח.



איור 11 - דוגמה לארכיטקטורת רשת מסוג מקודד-מפענח [15]

7 שיטות לסגמנטציית גידולי מוח מסוג גליומה בתמונות מוח-MRI

בפרק הקודם הצגנו מבנה כללי של רשת נוירונים עמוקה מסוג מקודד מפענח לצורך ביצוע סגמנטציה. בפרק זה נסקור מודלים שונים אשר מבוססים על ארכיטקטורת מקודד/מפענח לצורך ביצוע סגמנטציה של גידולים סרטניים מסוג גליומה בסריקות מוח – MRI.

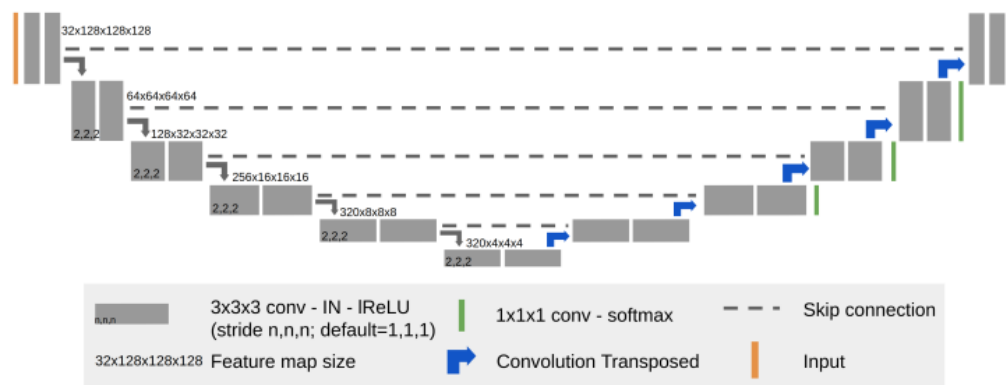
7.1 3D U-Net קלאסי [11]

nnU-Net [11] הינה ארכיטקטורת רשת נוירונים עמוקה מסוג 3D U-Net אשר מורכבת מנתיב מצמצם (מקודד – Encoder) ומנתיב מרחיב (מפענח – Decoder). ארכיטקטורת nnU-Net הוצעה בתחילה לביצוע משימות סגמנטציה על גבי בסיסי מידע שונים מהעולם הרפואי [16] ולאחר מכן הותאמה על מנת לבצע סגמנטציה על בסיס המידע של BraTS בפרט.

7.1.1 עיבוד מקדים

nnU-Net מבצעת נירמול של ווקסלי (פיקסל תלת מימדי) הסריקה המייצגים רקמת מוח ע"י מעבר על כל ווקסל כזה בנפרד והחסרת ממוצע העוצמה של כלל הווקסלים ממנו. ולאחר מכן מתבצעת חלוקה של כל ווקסל כזה בסטיית התקן הכללית של כלל הווקסלים. כל הווקסלים שאינם מייצגים רקמת מוח (רקע) יהיו בעלי ערך קבוע 0.

7.1.2 מבנה הרשת



איור 12 – מבנה רשת ה- nnU-Net המבוססת ארכיטקטורת 3D U-Net [11].

מבנה רשת nnU-Net הינו מבנה המבוסס על ארכיטקטורת 3D U-Net אשר מורכב ממקודד (נתיב מצמצם) ומפענח (נתיב מרחיב) אשר מחוברים ביניהם בעזרת Skip connections (הקווים המקווקים). גודל הקלט הינו קבוע ובגודל 128x128x128 (האימון מתבצע על Patches בגודל קבוע אשר חולצו מהקלט המקורי), גודל ה-Batch הינו 2.

ברשת ישנן 6 רמות, כאשר בכל רמה ישנן שתי שכבות קונבולוציה בגודל 3x3x3, שלאחריהן יש Instance normalization (שמטרתו לנרמל את הפלט של כל שכבת קונבולוציה עבור כל דגימה בנפרד, וזאת בניגוד ל-Batch normalization שבו הנירמול מתבצע ביחס לכל הדגימות ב-Batch) ושימוש בפונקציית אקטיבציה מסוג Leaky ReLU. ובסוף כל רמה ישנה פעולת Down sampling שמטרתה להקטין את מימדי הקלט (פי 2 בכל מימד).

ברמה התחתונה בנתיב המצמצם גודל ה-Feature map הינו 4x4x4.

מספר הגרעינים (Kernels) אשר נבחר ברמה הראשונה הינו 32 ובכל רמה מתחת כמות הגרעינים מוכפלת פי 2 (עד למקסימום של 320 ברמה התחתונה).
 בנתיב המרחיב, כמות הגרעינים בכל רמה מחקה את הכמות ברמה המקבילה בנתיב המצמצם. Feature map הפלט בכל רמה בנתיב המצמצם משורשר ל-Feature map הקלט של הרמה המקבילה בנתיב המרחיב (Skip connection) ועל שרשר זה מבוצעות שתי פעולות קונבולוציה, Instance normalization והפעלת פונקציית אקטיבציה מסוג Leaky ReLU בדומה לכל שכבה בנתיב המצמצם. העלאת מימד מתבצעת בעזרת פעולת קונבולוציה הפוכה.
 בסוף הנתיב המרחיב ישנה שכבת softmax אשר גודל הפלט שלה זהה לגודל הקלט. הפלט מכיל את ההסתברויות של כל ווקסל להשתייך לאחת ממחלקות הסיווג שהוגדרו בפרק הדן בבסיס הנתונים של BraTS.

7.1.3 אימון הרשת

מטרת האימון הינה לצמצם את ערך פונקציית ההפסד. פונקציית ההפסד שנבחרה לצורך אימון המודל הינה סכום פונקציות ההפסד Dice ו-Cross-entropy. כאשר פונקציית ההפסד מתייחסת ל-3 מחלקות הסיווג המייצגות את הגידול (בצקת, נמק, וגידול פעיל). אלגוריתם האימון שנבחר הינו Stochastic gradient decent עם מקדם למידה התחלתי של 0.01. תהליך האימון רץ בסך הכל כ-1,000 epochs כאשר בכל epoch ישנן 250 איטרציות.
 בתהליך האימון נדגמים באופן אקראי Patches בגודל קבוע של 128x128x128 מתוך הסריקה המקורית בסט האימון ועל כל קלט מתבצעת אוגמנטציה לצורך העשרת המידע.

7.1.4 שיטות אוגמנטציה להעשרת המידע

בתהליך האימון מתבצעת אוגמנטציה על הקלט המקורי על מנת להעשיר את המידע הקיים בצורה מלאכותית. בין השיטות שבהן נעשה שימוש לאימון המודל:
 ביצוע סיבוב (רוטציה) סביב ציר מסויים.
 ביצוע scale כלפי מעלה או כלפי מטה (בפקטור אקראי בטווח 0.65-1.6)
 ביצוע Elastic deformation בהסתברות של 0.3.
 שימוש בהגברת בהירות.
 שימוש ב-Gamma correction.

7.1.5 שיפורים שונים שבוצעו במודל במסגרת הניסויים השונים שהורצו

במהלך הניסויים השונים, הוצעו מספר שיפורים למודל על מנת להבין את השפעתם על ביצועי המודל.

הגדלת גודל ה-Batch – גודל ה-Batch המקורי שנבחר לרוב הניסויים הוא 2. הגודל הקטן של Batch מצד אחד מקטין את הסיכוי ל-Overfitting אך מצד שני מגביל את ביצועי האימון על סט האימון. לכן, בוצעו ניסויים שבהם גודל ה-Batch הינו גדול יותר ושווה ל-5. נסמן שיפור זה כ- **BL***.

אוגמנטציה אגרסיבית יותר – הגדלת ההסתברויות לביצוע אוגמנטציה והגדלת טווח הפרמטרים המוגרלים לביצוע האוגמנטציה וזאת על מנת לכפות אוגמנטציה עשירה ומגוונת יותר. נסמן שיפור זה כ- **DA/DA*** (כוכבית מסמנת שאוגמנטציית הבהירות מופעלת בהסתברות של 0.5).

ביצוע Batch Normalization – במקום Instance Normalization. נסמן שיפור זה כ- **BN**.
ביצוע Batch Dice – במימוש הסטנדרטי של Dice loss מחשבים את פונקציית ההפסד על כל דגימה ב-Batch בנפרד ואז מבצעים מיצוע של פונקציית ההפסד הכוללת על ה-Batch כולו. שגיאות קטנות בדגימות עם מעט ווקסלים מסומנים (כאשר הגידול מאוד קטן) עלולות להשפיע לרעה על עדכוני המשקלים בתהליך האימון. לכן, מומש חישוב Dice loss על גבי כל הדגימות ב-Batch יחדיו (ע"י התייחסות לכל הדגימות ב-Batch כדגימה אחת גדולה). נסמן שיפור זה כ- **BD**.

Region-based training – מחלקות הסיווג השונות הנתונות בסט האימון הינן בצקת, ליבת הגידול + נמק והגידול הפעיל. לעומת זאת, הערכת ביצועי המודל בתחרות BraTS היא על גבי 3 תתי אזורים חופפים:

- (1) Whole tumor – כלל הגידול המורכב מ-3 המחלקות שתוארו לעיל
- (2) Tumor core – המורכב מליבת הגידול + נמק והגידול הפעיל
- (3) Enhancing tumor – המורכב רק מהגידול הפעיל.

לכן, ביצוע אופטימיזציה למודל לפי 3 המחלקות המבצעות להערכה (במקום 4 המחלקות המקוריות שסומנו בסט האימון) עלולות לשפר את ביצועי המודל על גבי סט המבחן של BraTS. לשם כך, מבצעים את אופטימיזציית המודל ביחס למחלקות שמשמשות להערכת המודל בתחרות BraTS. נסמן שיפור זה כ- **R**.

7.1.6 סיכום הניסויים השונים

האימון בוצע מספר פעמים עם קומבינציות שונות של השיפורים שהוצגו לעיל וזאת על מנת להעריך את השפעת השיפורים השונים על ביצועי המודל. הניסויים מסומנים בצורה הבאה לדוגמה: BL^*+R+BD (שמשמעותו ניסוי שבו גודל ה-Batch היה גדול יותר, התבצע Region-based training ו-Batch dice).

סיכומי הניסויים על סט האימון:

טבלה 1 – סיכומי הביצועים של nnU-Net על סט האימון של BraTS 2020 [11]

Model	Training Set (Dice, 5-fold CV)			Mean
	Whole	Core	Enh.	
BL	91.6	87.23	80.83	86.55
BL*	91.85	86.24	80.18	86.09
BL*+R	91.75	87.24	82.21	86.73
BL*+R+DA	91.87	87.97	81.37	87.07
BL*+R+DA+BN	91.57	87.59	81.29	86.82
BL*+R+DA+BD	91.76	87.67	80.94	86.79
BL*+R+DA+BN+BD	91.7	87.21	81.7	86.87
BL*+R+DA*+BN	91.6	87.51	80.94	86.68
BL*+R+DA*+BN+BD	91.47	87.13	81.33	86.64

גודל סט האימון הינו 369, בכל הניסויים הורץ 5-fold cross validation (חלוקת סט האימון ל-5 קבוצות שוות וביצוע 5 תהליכי אימון שבכל אחד מהם בוחרים את אחת הקבוצות להיות סט המבחן וביצוע מיצוע התוצאות על גבי 5 הניסויים).

בטבלה מוצג מדד ה-Dice עבור כל אחת ממחלקות ההערכה בנפרד וכן הממוצע של שלוש המחלקות. ניתן לראות כי הניסוי אשר הניב את התוצאות הטובות ביותר על גבי סט האימון הוא הניסוי הרביעי (שבו התבצע אימון על Batch בגודל 5, Region based training ואוגמנטציה אגרסיבית יותר). עם זאת, ההבדלים בין הניסויים השונים אינם מאוד משמעותיים וכי הפער בין הניסוי המוצלח ביותר לניסוי הגרוע ביותר הוא בערך 1%.

סיכומי הניסויים על סט הולידציה:

הערכת ביצועי המודלים אל מול סט הולידציה של BraTS מתבצע אל מול שרתי BraTS. החיזוי התבצע בעזרת ensemble של 5 המודלים שחושבו בתהליך ה-5-Fold cross validation.

טבלה 2 – סיכומי הביצועים של nnU-Net על גבי סט הולידציה של BraTS 2020 [11]

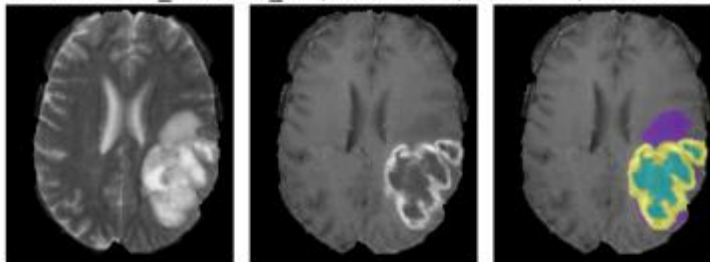
Model	Dice				HD95			
	Whole	Core	Enh.	Mean	Whole	Core	Enh.	Mean
BL	90.6	84.26	77.67	84.18	4.89	5.91	35.10	15.30
BL*	90.93	83.7	76.64	83.76	4.23	6.01	41.06	17.10
BL*+R	90.96	83.76	77.65	84.13	4.41	8.80	29.82	14.34
BL*+R+DA	90.9	84.61	78.67	84.73	4.70	5.62	29.50	13.28
BL*+R+DA+BN	91.24	85.04	79.32	85.2	3.97	5.17	29.25	12.80
BL*+R+DA+BD	90.97	83.91	77.48	84.12	4.11	8.60	38.06	16.93
BL*+R+DA+BN+BD	91.15	84.19	79.99	85.11	3.72	7.97	26.28	12.66
BL*+R+DA*+BN	91.18	85.71	79.85	85.58	3.73	5.64	26.41	11.93
BL*+R+DA*+BN+BD	91.19	85.24	79.45	85.29	3.79	7.77	29.23	13.60

מדדי האיכות שעל פיהם מתבצעת ההערכה בתחרות BraTS הינם Dice ו-Hausdorff distance. p95. ההערכה מתבצעת עבור 3 האיזורים שתוארו לעיל עבור כל אחד ממדדי האיכות. ניתן לראות כי הניסוי שהניב את התוצאות הטובות ביותר במדד Dice והן במדד HD95 הינו הניסוי השמיני שבו בוצע אימון על Batch size בגודל 5, Region-based training, אוגמנטציה אגרסיבית ו-Batch normalization.

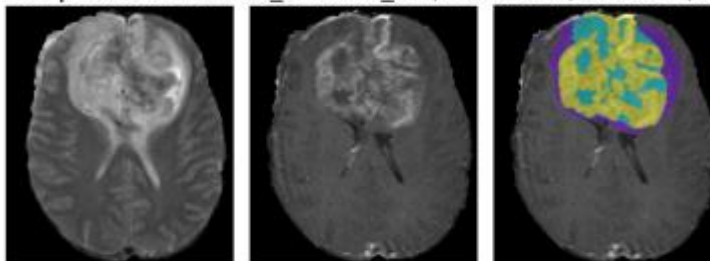
תוצאות ויזואליות:

באיור 13 מוצגות תוצאות ויזואליות על גבי סט הולידציה. כל שורה מייצגת סריקה שונה, הראשונה מייצגת את הסריקה שהניבה את התוצאות הטובות ביותר, השנייה את האחוזון ה-75, השלישית את הסריקה שהניבה את התוצאות החציוניות, הרביעית את האחוזון ה-25 והחמישית הינה הסריקה שהניבה את התוצאות הגרועות ביותר. בכל שורה, התמונה השמאלית מייצגת תמונת T2 ללא סימונים, באמצע תמונת T1c (עם הזרקת חומר ניגוד) ובימין תמונת T1c עם סימוני הסגמנטציה. הבצקת מסומנת בסגול, הגידול הפעיל בצהוב, וליבת הגידול והנמק מסומנים בטורקיז.

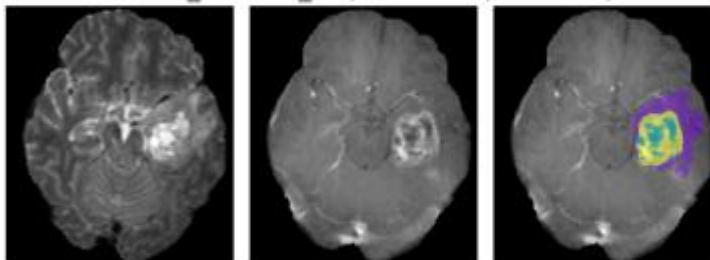
Best: BraTS20_Validation_040, whole 96.68, core 97.94, enh 95.87



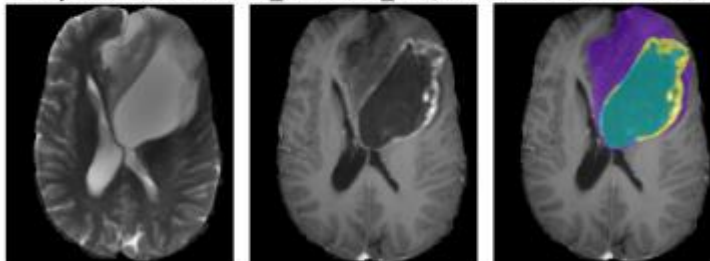
75th percentile: BraTS20_Validation_071, whole 96.75, core 90.11, enh 91.50



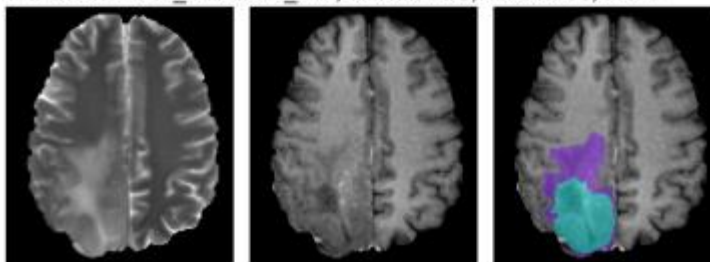
Median: BraTS20_Validation_059, whole 90.80, core 87.91, enh 91.40



25th percentile: BraTS20_Validation_100, whole 94.63, core 73.04, enh 76.48



Worst: BraTS20_Validation_084, whole 88.53, core 60.38, enh 0



איור 13 – תוצאות ויזואליות של nnU-Net על גבי סט הולידציה [11]

סיכומי הניסויים על סט המבחן:

טבלה 3 – סיכומי הניסויים של nnU-Net על גבי סט המבחן של BraTS 2020 [11].

	Dice			HD95		
	Enh	Whole	Core	Enh	Whole	Core
Mean	82.03	88.95	85.06	17.805	8.498	17.337
StdDev	19.71	13.23	24.02	74.886	40.650	69.513
Median	86.27	92.73	92.98	1.414	2.639	2.000
25th percentile	79.30	87.76	88.33	1.000	1.414	1.104
75th percentile	92.25	95.17	96.19	2.236	4.663	3.606

בטבלה הנ"ל מוצגות תוצאות ביצועי המודל על גבי סט המבחן הרשמי של תחרות BraTS 2020. המדדים שלפיהם הוערך המודל הינם Dice ו- Hausdorff Distance p95, וזאת עבור כל אחד מהאזורים שתוארו קודם. עבור כל אחד מהמדדים מציגים את המדד הממוצע, סטיית התקן, המדד החציוני, האחוזון ה-25 והאחוזון ה-75.

חשוב לציין כי מודל nnU-Net זכה במקום הראשון בתחרות הרשמית של BraTS 2020.

7.1.7 סיכום ומסקנות

nnU-Net הינו מודל גנרי אשר הניב תוצאות טובות מאוד במשימת סגמנטציה של גידולי מוח מסוג גליומה בסריקות מוח-MRI. מודל זה זכה במקום הראשון בתחרות BraTS 2020. מהתוצאות על סט המבחן נובע כי המודל הגיע לתוצאות הטובות ביותר על גבי מחלקת הערכה Whole tumor אשר מורכבת מכלל איזורי הגידול, גם במדד Dice וגם במדד Hausdorff Distance 95. והתוצאות הגרועות ביותר בשני המדדים הנ"ל היו על גבי מחלקת ההערכה של Enhancing Tumor אשר מורכבת מהחלק הפעיל של הגידול בלבד. זוהי מחלקת הסיווג הקשה ביותר לסגמנטציה שכן היא אינה קיימת בכלל הגידולים.

באיור 13 ניתן לראות כי עבור תוצאות הסגמנטציה הגרועות ביותר מדד ה-Dice הינו 0 עבור מחלקת Enhancing Tumor וזאת משום שהמודל לא הצליח לזהות איזור זה כלל בסריקה. עם זאת אלו מקרים יחסית נדירים ולמרות זאת מודל זה זכה בתחרות.

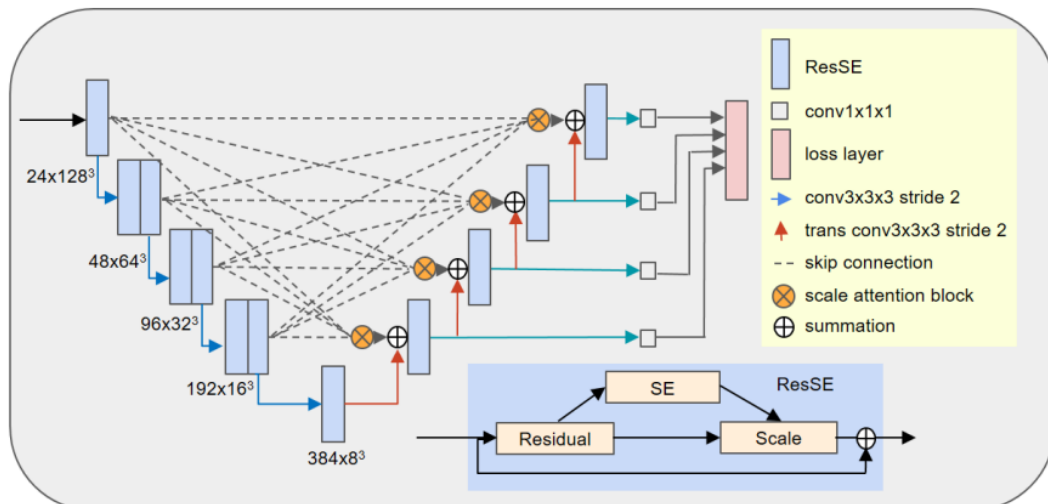
7.2 Scale Attention Network [17]

Scale Attention Network או בקצרה SA-Net [17] הינה רשת המבוססת על ארכיטקטורת מקודד-מפענח בדומה ל-nnU-Net. אך בשונה מ-nnU-Net, החיבורים בין השכבות השונות בין הנתבי המצמצם לנתבי המרחיב הינם חיבורים מלאים (כלומר כל שכבה בנתבי המצמצם מחוברת לכל אחת מהשכבות בנתבי המרחיב) וזו בניגוד ל-nnU-Net שבה החיבורים מתבצעים באותה הרמה בלבד. החיבורים המלאים נועדו על מנת ללכוד את המאפיינים שחולצו בכל הרמות בנתבי המצמצם ובכל רמה בנתבי המרחיב בו זמנית ובכך מאפשרים למקודד לבנות את מפת הסגמנטציה בצורה איכותית יותר בשלבי העלאת המימדים.

7.2.1 עיבוד מקדים

SA-Net מבצעת נירמול של ווקסלי הסריקה ע"י החסרת ממוצע העוצמה של כלל הווקסלים בסריקה. ולאחר מכן מתבצעת חלוקה של כל ווקסל כזה בסטיית התקן הכללית של הסריקה. זאת בשונה מ-nnU-Net ששם הנרמול מתבצע על ווקסלים המייצגים רקמת מוח בלבד.

7.2.2 מבנה הרשת



איור 14 – ארכיטקטורת הרשת SA-Net [17]

גודל הקלט של הרשת הינו $4 \times 128 \times 128 \times 128$ מכיוון שישנן 4 סריקות תלת מימדיות (עבור כל אחד משקלולי ה-MRI), וגם במודל זה בדומה ל-nnU-Net מתבצעת דגימה אקראית של Patch בגודל קבוע מתוך כל סריקה ($128 \times 128 \times 128$) וזו מפאת אילוצי זכרון GPU. ברמה הראשונה של הנתבי המצמצם ישנה שכבת ResSE [18] (מבנה שכבה זו יתואר בהמשך), לאחריה יש פעולת קונבולוציה עם דילוג בגודל 2 כך שגודל ה-Feature map מצטמצם פי 2 בכל אחד מהמימדים בעוד שכמות הערוצים גדלה פי 2 בכל שלב. בשאר הרמות בנתבי המצמצם ישנן שתי שכבות ResSE בניגוד לרמה הראשונה שבה יש רק שכבה אחת. ברמה התחתונה ביותר בנתבי המצמצם מימדי ה-Feature map הינם $384 \times 8 \times 8 \times 8$. בסוף הנתבי המצמצם ישנו נתבי מרחיב בעל אותה כמות רמות כמו הנתבי המצמצם. בכל רמה בנתבי המצמצם מבצעים פעולת קונבולוציה הפוכה עם דילוג בגודל 2 כך שמימדי ה-Feature map יגדלו פי 2 וכמות הערוצים תקטן פי 2. לאחר מכן, מחברים ל-Feature map שגדל במימדיו את הפלט של שכבת Scale attention block (אשר תתואר בהמשך) שמבצעת עיבוד על כל פלטי ה-

Feature map ברמות השונות בנתיב המצמצם. ניתן לראות כי קיים Skip connection בין כל שכבות הנתיב המצמצם לכל אחד מה-Scale attention blocks. לאחר ביצוע החיבור, ישנה שכבת ResSE בדומה לנתיב המצמצם.

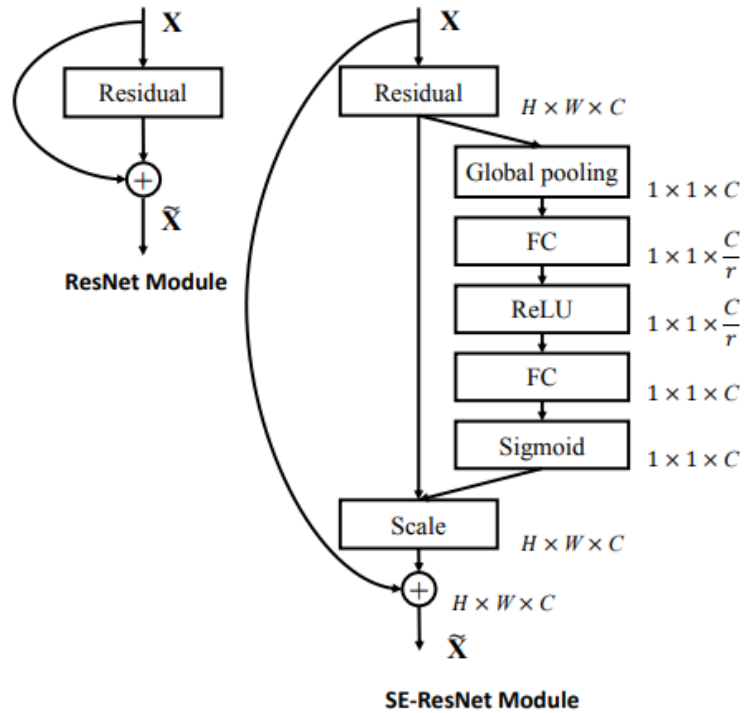
בכל אחת מרמות הנתיב המרחיב, מבצעים פעולת Upsampling על מנת להעלות את מימדי ה-Feature map ל- $128 \times 128 \times 128$ (פרט לשכבה העליונה שבה מימדי ה-Feature map כבר כאלה) ומבצעים פעולת קונבולוציה עם גרעין בגודל $1 \times 1 \times 1$ וזאת על מנת להקטין את מספר הערוצים ב-Feature map ל-3, ובנוסף מפעילים פונקציית אקטיבציה מסוג סיגמואיד. ארבעת הפלטים של כל אחת מהרמות בנתיב המרחיב מחוברות לפונקציית ה-Loss, אשר הפלט שלה הינו בגודל $3 \times 128 \times 128 \times 128$ (זהה למימדי הקלט אך בעלת 3 ערוצים), כאשר כל ערוץ מציין את ההסתברויות של הווקסלים להשתייך לכל אחת מ-3 מחלקות הסיווג.

7.2.2.1 ResSE

כל בלוק ResSE מבוסס על בלוק ResNet [19] כאשר כל בלוק כזה מורכב משתי שכבות קונבולוציה עם Instance Normalization ו-ReLU כפונקציית אקטיבציה שלאחריהן מתבצע חיבור של ה-Feature Map שהתקבל כפלט ל-Feature Map שהתקבל כקלט (Skip connection פנימי בתוך הבלוק). בנוסף, על מנת לשפר את יכולות המודל, נוסף מודול מסוג Squeeze and Excitation (SE) [18] לכל אחד מבלוקי ה-ResNet.

היתרונות בשימוש בבלוק ResSE הינם :

- השימוש ב-Residual block מונע פגיעה בביצועים במידה וערך פונקציית האקטיבציה מתבדר (שכן יש Skip connection בין קלט הבלוק לבין פלט הבלוק שעליו הופעלה הפונקציה)
- מאפשר לכידת תלויות בין הערוצים השונים בסריקה וביצוע הבחנה בין מאפיינים שונים בין הערוצים.
- יעילות חישובית שמושגת על ידי ביצוע קידוד של כל ערוץ בנפרד ע"י נציג בודד.



איור 15 – מבנה בלוק ResNet סטנדרטי (משמאל) ומבנה בלוק ResSE (מימין). [18]

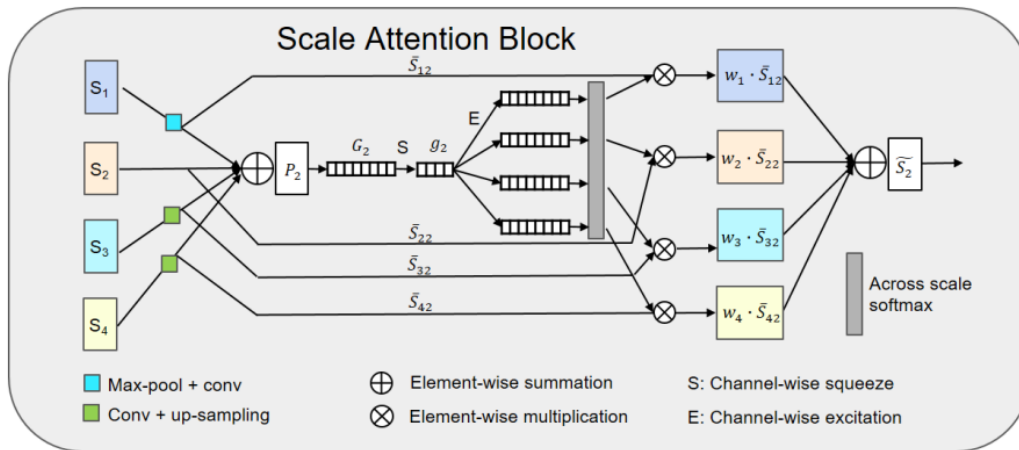
פעולת ה-Squeeze נועדה על מנת לקודד מידע על כל ערוץ בקלט בעזרת נציג יחיד. על מנת לחשב נציג כזה, משתמשים ב-Global average pooling בכדי לחשב את הערך הממוצע של ה-Feature map בכל ערוץ בנפרד, כל ערך כזה הינו הנציג של הערוץ. פלט פעולה זו הינה בגודל $1 \times 1 \times C$. כאשר C מייצג את מספר הערוצים.

לאחר פעולת ה-Squeeze מבצעים פעולת Excitation על מנת לזהות את התלויות בין הערוצים השונים. פעולה זו נעשית ע"י הורדת מימד בפקטור r (שנקבע להיות 4 במודל SA-Net). פעולה זו מתבצעת ע"י מעבר בשכבת Fully Connected שלאחריה פונקציית אקטיבציה לא ליניארית מסוג ReLU. לאחר מכן מעבר בשכבת Fully Connected נוספת על מנת להעלות חזרה את מימד הערוצים ל-C המקורי. לאחריה פונקציית אקטיבציה מסוג סיגמואיד.

מבצעים פעולת Scale ע"י ביצוע מכפלה סקלארית של פלט שכבת ה-SE בפלט שכבת ה-Residual (אשר מורכבת משתי שכבות קונבולוציה) ולבסוף מחברים את פלט שכבת ה-Scale עם ה-Feature Map שהתקבל כקלט לבלוק ה-ResSE ע"י Skip connection.

7.2.2.2 Scale attention block

Scale attention block (בקצרה SA-block) הינו בלוק המורכב מ-Skip connections בין כל הרמות (Scales) בנתיב המצמצם לכל הרמות בנתיב המרחיב.



איור 16 – מבנה Scale attention block. כאן S1-S4 הינם הקלטות אשר הגיעו מרמות שונות בנתיב המצמצם. בדוגמה זו $d=2$. [17]

השלב הראשון של SA-block הינו ביצוע טרנספורמציה של כלל ה-Feature maps שהתקבלו כקלט לאותם המימדים. כאן ישנן 4 רמות (פרט לרמה התחתונה) אשר נכנסות כקלט לכל SA-block ($S_1 - S_4$). נסמן ב- d ו- e את הרמה בנתיב המצמצם ובנתיב המרחיב בהתאמה.

פונקציית הטרנספורמציה $f_{ed}(S_e)$ מוגדרת כך:

במידה ו $e < d$, הפונקציה מבצעת הורדת מימד ל- S_e בפקטור $2^{(d-e)}$ ע"י ביצוע max pooling ושכבת קונבולוציה עם נורמליזציה ו-ReLU כפונקציית אקטיבציה. במידה ו $e = d$, אזי הפונקציה מתנהגת כפונקציית הזהות (שכן אין צורך להוריד או להעלות מימד) במידה ו $e > d$, הפונקציה מבצעת העלאת מימד על ידי ביצוע קונבולוציה נורמליזציה ו-ReLU כפונקציית אקטיבציה שלאחריה ישנה פעולת Upsampling.

לאחר מכן מבצעים סכימה על כלל הפלטות של הפונקציה $f_{ed}(S_e)$, אשר נסמנה כ- P_d ועל Feature map זה מבצעים פעולת Pooling על כל ערוץ בנפרד ובכך מתקבל וקטור אשר גודלו כמספר הערוצים המקורי וכל כניסה מכילה נציג עבור הערוץ המקביל. לאחר מכן מבצעים פעולת Squeeze-excitation על הווקטור שהתקבל בדומה לבלוק ResSE וזאת ע"י ביצוע מעבר בשכבת Fully connected עם מקדם צמצום r .

לאחר מכן, מבוצע מעבר ב-4 שכבות Fully connected נוספות עם פונקציית אקטיבציה מסוג סיגמואיד אשר מעלות את מימד אורך הווקטורים למספר הערוצים המקורי. ארבעת הווקטורים שהתקבלו מועברים בשכבת Softmax על מנת לקבל משקל (המייצג את התרומה של כל וקטור המייצג Scale אחר). כל משקל כזה מוכפל ב-Feature map המקביל (לאחר השוואת ה-Scale שנעשתה על ידי פונקציית $f_{ed}(S_e)$ הפלט הסופי של SA-block הינו סכום מכפלת ארבעת ה-Feature maps במשקלים שהתקבלו.

7.2.3 אימון הרשת

לצורך ביצוע האימון נדגמו אקראית מהסריקות Patches בגודל קבוע של 128x128x128 ווקסלים. גודל ה-Batch נקבע ל-1. אימון הרשת ארך כ-300 Epochs ובוצע שימוש ב-Adam optimizer עם מקדם למידה התחלתי של כ-0.003.

פונקציית ההפסד שבוצע בה שימוש לצורך האימון הינה סכום של שתי פונקציות הפסד שונות

(1) Jaccard Loss [20] - הינה פונקציית הפסד המוגדרת כמפורט באיור הבא :

$$L_{d,J} = 1 - \frac{\sum_{i,j} (t_{ij} p_{ij})}{\sum_{i,j} t_{ij}^2 + \sum_{i,j} p_{ij}^2 - \sum_{i,j} (t_{ij} p_{ij})}$$

איור 17 - הגדרת פונקציית ההפסד Jaccard loss [20]

$t_{i,j}$ ו- $p_{i,j}$ מסמנים את ערך ה-Ground truth וערך הפרדיקציה עבור אותו הווקסל

בהתאמה. מטרת תהליך האימון הינו למזער ביטוי זה.

(2) Focal Loss [21] - הינה פונקציית הפסד המוגדרת כמפורט באיור הבא :

$$FL(p_t) = -(1 - p_t)^\gamma \log(p_t).$$

איור 18 - הגדרת פונקציית ההפסד Focal Loss [21]

p_t נקבע להיות ערך הפרדיקציה p של המודל כאשר הווקסל משתייך למחלקת הסיווג הרצויה, אחרת (כאשר הווקסל משתייך לרקע) ערכו של p_t נקבע להיות $1-p$.

$\gamma \geq 0$ הינו פרמטר המשמש ל-*tuning* ונקרא *focusing parameter*.

במודל זה פונקציית ההפסד חושבה עבור כל אחד מ-3 מחלקות הסיווג המרכיבות את הגידול (לא כולל הרקע) נפרד, ולאחר מכן בוצע סכימה של התוצאות.

תהליך אימון המודל התבצע בעזרת 5-fold cross validation על סט האימון בדומה ל-nnU-Net.

7.2.4 שיטות אוגמנטציה להעשרת המידע

על מנת למזער את הסיכוי ל-Overfitting, בוצעה אוגמנטציה מכמה סוגים על מנת להעשיר בצורה מלאכותית את ה-Dataset :

(1) היפוך מראה של הסריקה סביב אחד הצירים בהסתברות של 0.5.

(2) שינוי הניגודיות של כל ערוץ בסריקה בפקטור רנדומלי שנדגם בטווח של 0.9 – 1.1.

תוצאות המודל על סט האימון:

טבלה 4 – תוצאות מודל SA-Net על סט האימון של BraTS 2020 [17]

		fold-0	fold-1	fold-2	fold-3	fold-4	ALL
SA-Net	WT	0.9256	0.9002	0.9217	0.9085	0.9195	0.9151
	TC	0.8902	0.8713	0.8758	0.8538	0.8953	0.8773
	ET	0.8220	0.7832	0.8198	0.8107	0.8268	0.8125
	AVG	0.8793	0.8516	0.8724	0.8577	0.8805	0.8683

גודל סט האימון הינו 369, בכל הניסויים הורץ 5-fold cross validation (חלוקת סט האימון ל-5 קבוצות שוות וביצוע 5 תהליכי אימון שבכל אחד מהם בוחרים את אחת הקבוצות להיות סט המבחן וביצוע מיצוע התוצאות על גבי 5 הניסויים).

בטבלה מוצג מדד ה-Dice עבור כל אחד ממחלקות ההערכה השונות עבור כל אחד מהרצות ה-Fold ובנוסף מוצג ממוצע המדד הכולל של כל Fold.

תוצאות המודל על סט הולידציה:

טבלה 5 – תוצאות מודל SA-Net על סט הולידציה של BraTS 2020 [17]

	DSC			HD95		
	WT	TC	ET	WT	TC	ET
The best single model	0.9044	0.8422	0.7853	5.4912	8.3442	20.3507
Ensemble of 11 models	0.9108	0.8529	0.7927	4.0975	5.8879	18.1957

הערכת ביצועי המודלים אל מול סט הולידציה של BraTS מתבצע אל מול שרתי BraTS. החיזוי התבצע בעזרת ensemble של 11 מודלים שונים שחושבו בעזרת 5 Fold cross validation והתאמות נוספות.

בטבלה מוצגים מדדי ה-Dice ו- Hausdorff Distance p95 עבור כל אחד ממחלקות הערכה השונות. השורה הראשונה בטבלה מייצגת את תוצאות המודל הבודד הטוב ביותר. והשורה השנייה מייצגת את תוצאות ה-Ensemble של 11 מודלים שונים.

תוצאות המודל על סט המבחן:

טבלה 6 – תוצאות מודל SA-Net על סט המבחן של BraTS 2020 [17]

	DSC			HD95		
	WT	TC	ET	WT	TC	ET
Ensemble of 11 models	0.8828	0.8433	0.8177	5.2176	17.9697	13.4298

לצורך, הערכת המודל מול סט המבחן של BraTS 2020, בוצעה הערכה בודדת של Ensemble 11 המודלים ששומשו לצורך הערכת המודל מול סט הולידציה של BraTS 2020. עבור סט המבחן חושבו מדדי האיכות הממוצעים של Dice ו-Hausdorff distance p95 עבור כל אחת משלוש מחלקות ההערכה.

7.2.6 סיכום ומסקנות

SA-Net הניב תוצאות טובות ביותר בתחרות BraTS וזכה במקום ה-3, זאת למרות אילוצי המשאבים (ובפרט GPU) ושימוש ב-Batch size של 1 בלבד מפאת אילוצי זכרון. דבר זה התאפשר בזכות השימוש ב-Scale Attention Blocks. ניתן לראות כי המודל הניב את התוצאות הטובות ביותר במדד ה-Dice עבור מחלקת ההערכה הכללית Whole Tumor המייצגת את כלל הגידול ואילו התוצאות הגרועות ביותר התקבלו עבור מחלקת Enhancing Tumor במדד ה-Dice בדומה ל-nnU-Net שכן זוהי מחלקת ההערכה הקשה ביותר לסיווג הכוללת רק את מחלקת הסיווג Enhancing Tumor המתאימה.

7.3 [10] Hybrid High-resolution and Non-local Feature Network

Hybrid High-resolution and Non local Feature Network או בקצרה H2NF-Net [10] הינה רשת מסוג מקודד מפענח הבנויה בצורה מדורגת. בדומה ל-Scale Attention Network שהוצגה קודם לכן, ישנם חיבורים בין שכבות מרמות שונות.

7.3.1 עיבוד מקדים

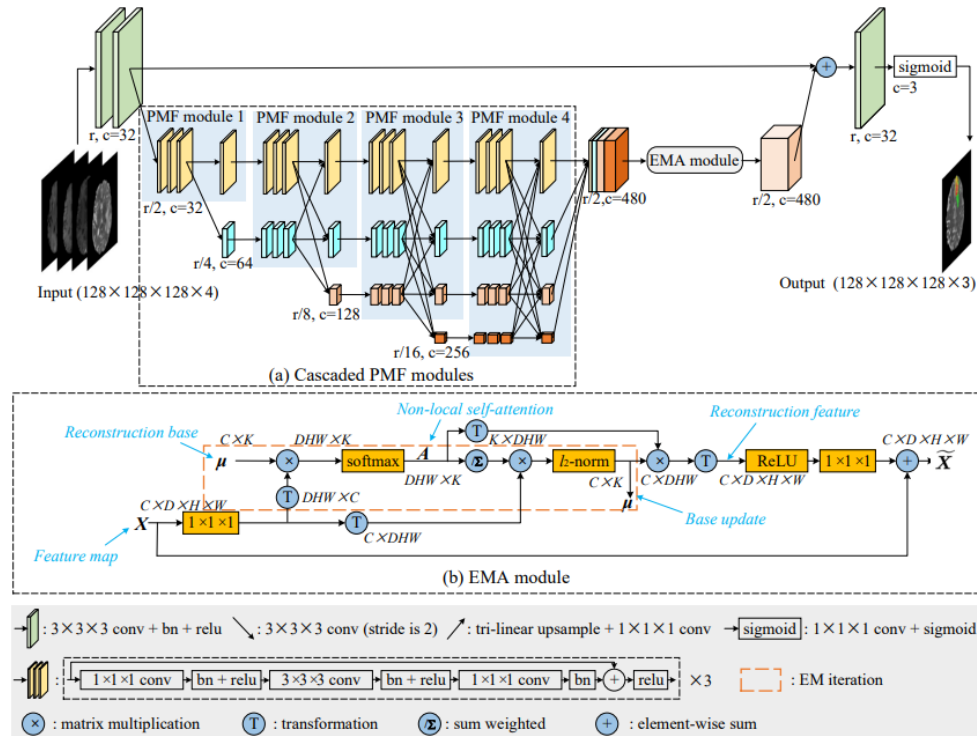
לצורך העיבוד המקדים, חולצו כל הווקסלים אשר מציינים רקמת מוח, כל ווקסלים אלו נורמלו ע"י החסרת ממוצע הווקסלים מכל ווקסל וחלוקה בסטיית התקן של כלל הווקסלים המייצגים רקמת מוח.

7.3.2 מבנה הרשת

לצורך הבנת מבנה רשת H2NF-Net המדורגת, נציג קודם לכן את מבנה רשת HNF-Net [22]

7.3.2.1 Single HNF-Net

HNF-Net הינה ארכיטקטורת רשת מסוג מקודד מפענח. קלט הרשת הינו שרשרת של 4 השכלולים של סריקת MRI אשר מתוכו נדגם Patch בגודל $128 \times 128 \times 128$. הסריקה מעובדת ב-5 רמות שונות, כאשר מימדי ה-Feature map בכל רמה קטנים פי 2 (באיור הרמות מסומנות בירוק, צהוב, כחול, ורוד וכתום). ברמה העליונה (אשר מימדיה זהים למימדי הקלט) ישנם ארבעה בלוקי קונבולוציה, שניים בנתיב המצמצם ושניים בנתיב המרחיב, ברמה זו ישנו חיבור ישיר (Skip connection) בין המקודד למפענח וזאת על מנת לשמר את המידע הקיים ברזולוציה הגבוהה. בארבע הרמות התחתונות, ישנם ארבע מודולים של (parallel multi-scale fusion) PMF, מודול זה (יתואר בהמשך) המחברים יחדיו בכדי לאגור מאפיינים אשר חולצו מהרמות השונות. בסוף מודולי ה-PMF, ארבעת ה-Feature maps הפלט שהתקבלו בארבע הרמות השונות, מועלים למימד זהה הקטן פי 2 ממימד הקלט המקורי, ולאחר מכן משורשים ל-Feature Map יחיד. ה-Feature map שהתקבל עובר ב-EMA module (expectation-maximization attention), גם מודול זה יתואר בהמשך) שמטרתו ללכוד בצורה יעילה את התלויות בין המידע שנאגר ברמות השונות ולהפחיתן. הפלט של מודול זה מועלה למימדי הקלט המקוריים (וזאת ע"י העלאת המימד וביצוע פעולת קונבולוציה בגודל $1 \times 1 \times 1$ עם גרעין בעל 32 ערוצים) ולאחר מכן מתבצע חיבור של פלט זה עם ה-Feature map אשר התקבל כקלט מהרמה העליונה בנתיב המצמצם, לאחר מכן מופעלת פונקציית אקטיבציה מסוג סיגמואיד וזהו הפלט הסופי של הרשת. לאורך הרשת, פעולות ה-down sample מבוצעות בעזרת Convolutions עם דילוג (Stride) בגודל 2, וכל פעולות העלאת המימד לאורך הרשת מבוצעות ע"י איחוד של פעולות קונבולוציה עם גרעין בגודל $1 \times 1 \times 1$ ואינטרפולציה ליניארית כמודגם באיור 19.



איור 19 – מבנה רשת HNF-Net [10]

7.3.2.1.1 Parallel Multi-Scale Fusion (PMF) Module

מודול ה-PMF הינו מודול המורכב משני חלקים:

- 1) Parallel multiscale convolutional block – אשר מכיל ענפים מקבילים ברמות שונות המורכבים מבלוקי קונבולוציה עם Skip Connection (מסומן בצהוב במקרא באיור).
- 2) Fully connected fusion block – אשר משלב את כלל הפלטים שהתקבלו ברמות השונות בבלוקים המקבילים ע"י הבאתם למימד אחיד וחיבורם (Summation).

החלק הראשון במודול ה-PMF נועד על מנת להשתמש במאפיינים אשר חולצו ברמות (רזולוציות) שונות תוך כדי שמירה על ייצוג המאפיינים ברזולוציה הגבוהה. החלק השני במודול נועד על מנת לבצע אגרגציה על מידע מרמות שונות.

בנוסף, משתמשים בארבעה מודולי PMF מדורגים אשר בהם עומק הענפים גדל ככל שהעומק ברשת גדל. כתוצאה מכך, המאפיינים ברמות הגבוהות מוגברים ע"י היתוך חוזר עם מאפיינים ברמות נמוכות יותר.

7.3.2.1.2 Expectation-Maximization Attention (EMA) Module

המטרה של מנגנון Non-Local self attention הינה ביצוע אגרגציה של מידע אשר התקבל לאורך התהליך ברמות השונות וללכוד את התלויות שנוצרו בין המאפיינים לאורך התהליך, דבר המאפשר למודל לזהות את התלויות בין מחלקות הסיווג השונות המרכיבות את הגידול ולהבחין ביניהן בצורה טובה יותר. עם זאת, העלות החישובית של תהליך זה אינה מאפשרת להשתמש בטכניקה זו לביצוע סגמנטציה תלת מימדית.

לצורך כך שולב ברשת מודול ה-EMA אשר נועד להכניס מנגנון אשר פועל כמנגנון Non-Local attention אך קל יותר מבחינה חישובית. מנגנון זה פועל על סט של בסיסים הבונים את המאפיינים

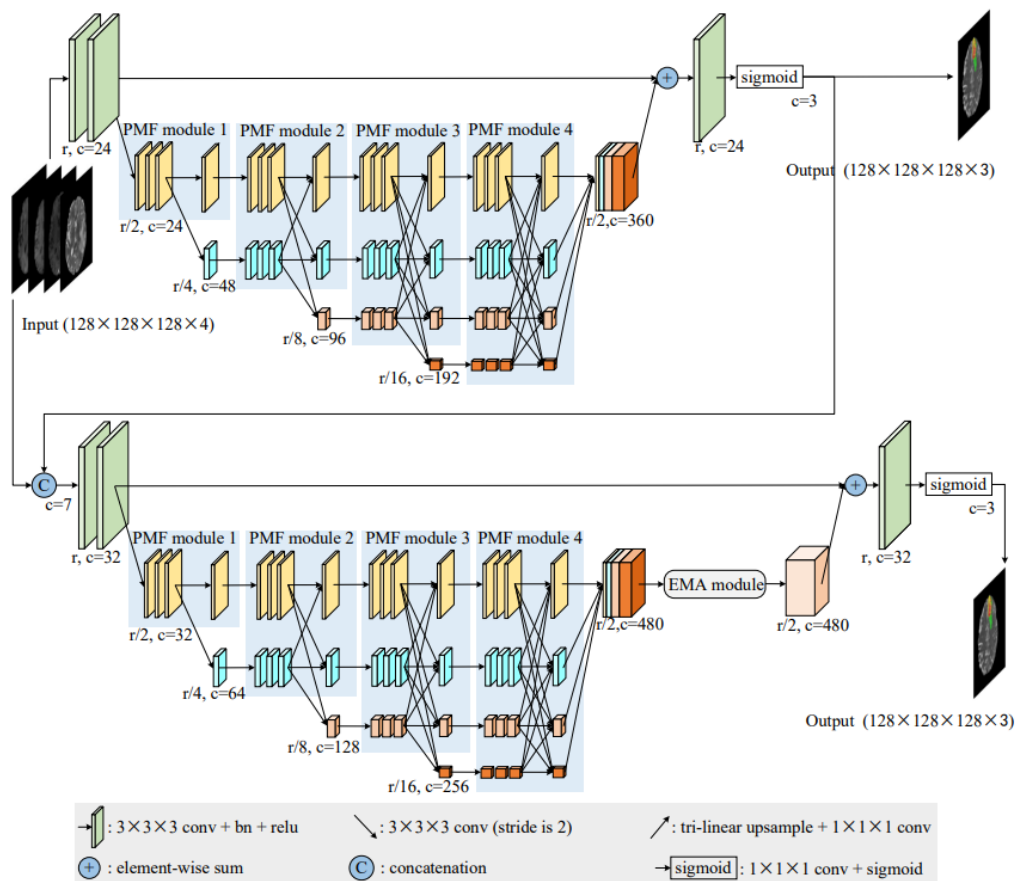
במקום לפעול באופן ישיר על Feature maps מלאים במימד גבוה. מכיוון שבסיסים אלו מכילים כמות פחותה משמעותית של אלמנטים ביחס ל-Feature maps המקוריים, העלות החישובית קטנה משמעותית.

כפי שניתן לראות באיור התחתון, הגודל של קלט ה-Feature map אשר מיוצג כ- X הינו $C \times D \times H \times W$ כאשר C מייצג את מספר הערוצים, המרובע המקווקו האדום מייצג את פעולת ה-EM (Expectation – Maximization).

A מייצג את ה-Non-local attention map, ו- μ מייצג את הבסיס. A ו- μ מתעדכנים במהלך איטרציית ה-EM. ולאחר התכנסות ערכים אלה (החלק המקווקו הינו תהליך איטרטיבי) מתבצעת מכפלה של A ב- μ על מנת לייצר את ה-Feature map שנבנה מחדש על סמך הבסיס שנוצר, נסמנו ב- $\tilde{X}$. בסוף מודול ה-EMA מתבצעת פעולת קונבולוציה עם גרעין בגודל $1 \times 1 \times 1$ על מנת להחזיר את מימד ה-Feature map שהתקבל למימד הקלט המקורי ועל מנת למנוע Overfitting גם מתבצעת סכימה של X (הקלט המקורי) עם $\tilde{X}$ (Skip connection).

7.3.2.2 Cascaded HNF-Net

רשת זו הינה רשת מדורגת אשר בנויה משני שלבים דמויי HNF-Net כפי שניתן לראות באיור.



איור 20 – מבנה רשת Cascaded HNF-Net [10]

בהשוואה ל-Single HNF-Net, השלב הראשון ב-Cascaded HNF-Net הינו צר יותר ואינו מכיל EMA module ואילו השלב השני בעל מבנה זהה ל-Single HNF-Net פרט לכך שהקלט של שלב זה הינו השרשור של הקלט המקורי והפלט שהתקבל מהשלב הקודם. Single HNF-Net הציג ביצועים טובים יותר על מחלקות הסיווג ET (החלק הפעיל בגידול) ו-WT (כלל הגידול), ואילו Cascaded HNF-Net הציג ביצועים טובים יותר על מחלקת ה-TC (ליבת הגידול המכילה את רקמות הנמק והגידול הפעיל) על סט הולידציה של BraTS, ולכן בוצע שימוש ב-Single HNF-Net לסיווג הבצקת והגידול הפעיל וב-Cascaded HNF-Net לצורך סיווג רקמות הנמק וחלק הגידול הלא פעיל. מודל משולב זה נקרא H2NF-Net.

7.3.3 אימון הרשת

לצורך תהליך האימון, חולצו Patches תלת מימדים מתוך הסריקות בצורה רנדומלית בגודל קבוע של כ-128x128x128, כאשר קלט הרשת הינו שרשור של ארבעה Patches כאלה (עבור כל אחד משקלולי סריקת ה-MRI).

המודל אומן לאורך 450 epochs, ובוצע שימוש ב-Adam optimizer עם מקדם למידה התחלתי של כ-0.0085.

פונקציית ההפסד בה נעשה שימוש לאורך תהליך האימון הינה שילוב של שתי פונקציות הפסד, הראשונה הינה Dice loss והשנייה הינה Binary cross-entropy. תהליך האימון ווידוא המודל התבצע בעזרת 5-fold cross validation על סט האימון.

7.3.4 שיטות אוגמנטציה להעשרת המידע

על מנת להקטין את הסיכוי ל-Overfitting, בוצעה אוגמנטציה על גבי סט האימון על מנת להעשיר את המידע בצורה מלאכותית. בין שיטות האוגמנטציה שבוצעו:

- (1) ביצוע היפוך מראה רנדומלי סביב כל אחד מ-3 הצירים (בצורה בלתי תלויה)
- (2) ביצוע סיבוב רנדומלי של עד 10 מעלות, סביב כל אחד מ-3 הצירים (בצורה בלתי תלויה)
- (3) ביצוע Scaling בפקטור שבין 0.9-1.1.

7.3.5 תוצאות המודל

תוצאות המודל על סט הולידציה:

טבלה 7 – תוצאות מודל H2NF-Net על סט הולידציה של BraTS 2020 [10]

Method	Dice(%)				95%HD(mm)			
	ET	WT	TC	Mean	ET	WT	TC	Mean
Single*	0.78492	0.91261	0.83532	0.84428	26.60476	4.17888	5.41503	12.06622
Single ⁺	0.78908	0.91218	0.84887	0.85004	26.50355	4.10500	5.14468	11.91774
Cascaded*	0.77647	0.91084	0.85631	0.84787	26.68954	4.38397	4.93158	12.00169
Cascaded ⁺	0.77338	0.91022	0.85701	0.84687	29.71248	4.30247	4.93369	12.98288
H ² NF-Net	0.78751	0.91290	0.85461	0.85167	26.57525	4.18426	4.97162	11.91038

בטבלה ניתן לראות את תוצאות הערכת המודל על גבי סט הולידציה של BraTS 2020 ע"פ מדדי האיכות Dice ו- Hausdorff Distance P95 עבור כל אחת מ-3 מחלקות ההערכה של התחרות.

Single* ו- Cascaded* מסמנים את ה-Ensemble של 5 מודלי Single HNF-Net ו-5 מודלי Cascaded HNF-Net בהתאמה, אשר התקבלו כפלט מתהליך האימון עם 5-fold cross validation.

Single+ ו- Cascaded+ מסמנים Ensemble של 7 מודלי Single HNF-Net על גבי כלל סט האימון (ללא Cross validation) ו-Ensemble של 10 מודלי Cascaded HNF-Net על גבי כלל סט האימון בהתאמה.

מהטבלה ניתן לראות כי מודלי ה-Single HNF-Net הניבו תוצאות טובות יותר על גבי מחלקות ההערכה ET ו- WT ואילו מודלי ה-Cascaded HNF-Net הניבו תוצאות טובות יותר על גבי מחלקת ההערכה TC.

H2NF-Net, אשר מהווה את ה-Ensemble של כל 27 המודלים הני"ל הניבה את התוצאות הטובות ביותר על גבי סט הולידציה עם ציון Dice ממוצע של כ-0.85167 ו ציון HD95 ממוצע של כ 11.91038.

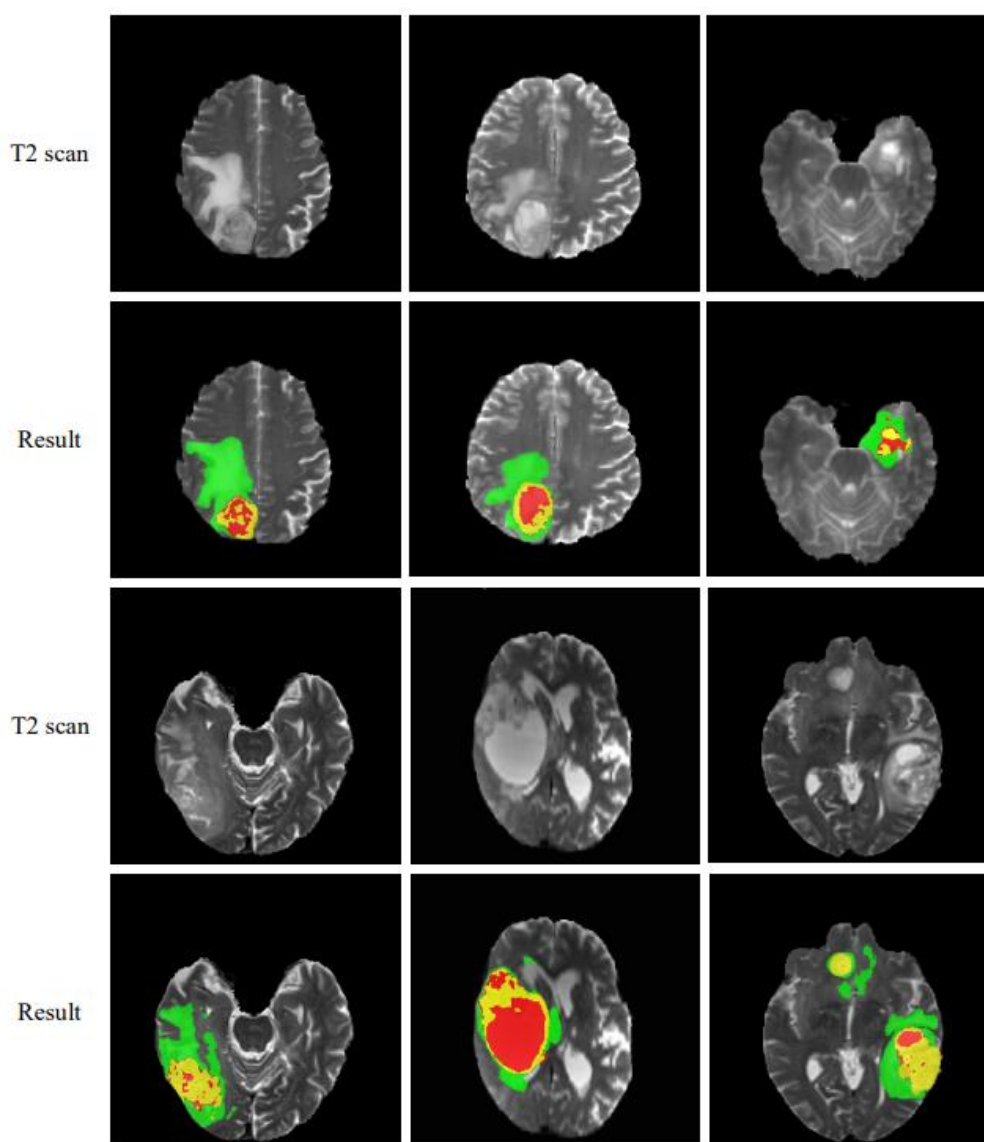
תוצאות המודל על סט המבחן:

טבלה 8 – תוצאות מודל H2NF-Net על גבי סט המבחן של תחרות BraTS 2020 [10]

Method	Dice(%)			95%HD(mm)		
	ET	WT	TC	ET	WT	TC
H ² NF-Net	0.82775	0.88790	0.85375	13.04490	4.53440	16.92065

לצורך, הערכת המודל מול סט המבחן של BraTS 2020, בוצעה הערכה בודדת של H2NF-Net (המודל שהתקבל מביצוע Ensemble של 27 המודלים השונים ששומשו להערכה על גבי סט הולידציה) על גבי הנתונים בסט המבחן אל מול שרתי BraTS.

עבור סט המבחן חושבו מדדי האכיות הממוצעים של Dice ו- p95 Hausdorff distance עבור כל אחת משלוש מחלקות ההערכה.



איור 21 – תוצאות סגמנטציה ויזואליות של מודל H2NF-Net [10]

באיור 21 מוצגות תוצאות סגמנטציה ויזואליות של מודל H2NF-Net על גבי סריקות MRI אמיתיות מתוך סט המבחן של BraTS 2020. בשורה הראשונה ובשורה השלישית מוצגות 3 פרוסות מתוך סריקה בשקלול T2 ללא סימוני סגמנטציה, ובשורה השנייה ובשורה הרביעית מוצגת הסגמנטציה (סימון הגידול) עבור כל אחת מהפרוסות בהתאמה. כאשר ירוק מייצג את הבצקת, אדום מייצג את רקמת הנמק והגידול הלא פעיל, וצהוב מייצג את רקמת הגידול הפעילה.

7.3.6 סיכום ומסקנות

מודל H2NF-Net הניב תוצאות טובות ביותר בביצוע סגמנטציה של גידולי מוח מסוג גליומה בסריקות מוח-MRI. מודל זה זכה במקום השני בתחרות BraTS 2020 (לאחר nnU-Net). גם מודל זה בדומה לשאר המודלים הניב את התוצאות הטובות ביותר על גבי מחלקת ההערכה Whole Tumor המורכבת משלוש מחלקות הסיווג המרכיבות את הגידול (הן במדד Dice והן במדד Hausdorff Distance). התוצאות הגרועות ביותר התקבלו על מחלקת ההערכה Enhancing Tumor בדומה למודלים הקודמים שהוצגו. למרות התוצאות המרשימות של מודל H2NF-Net, לא ניכר כי היה יתרון בשימוש בארכיטקטורה מורכבת הכוללת בלוקי EMA ו-PMF שכן מודל זה לא הציג עליונות על גבי מודל nnU-Net שמבנהו פשוט הרבה יותר.

7.4 [23] Modality-Pairing Network

Modality Pairing Network [23] הינה רשת מסוג מקודד-מפענח, אשר בניגוד למודלים הקודמים שהוצגו, Modality Pairing Network אינה מבצעת שרשרת של ארבעת השכלולים השונים של סריקת ה-MRI בקלט המודל אלא מבצעת צימוד בין זוגות שכלולים וריצה בשני ענפים מקבילים.

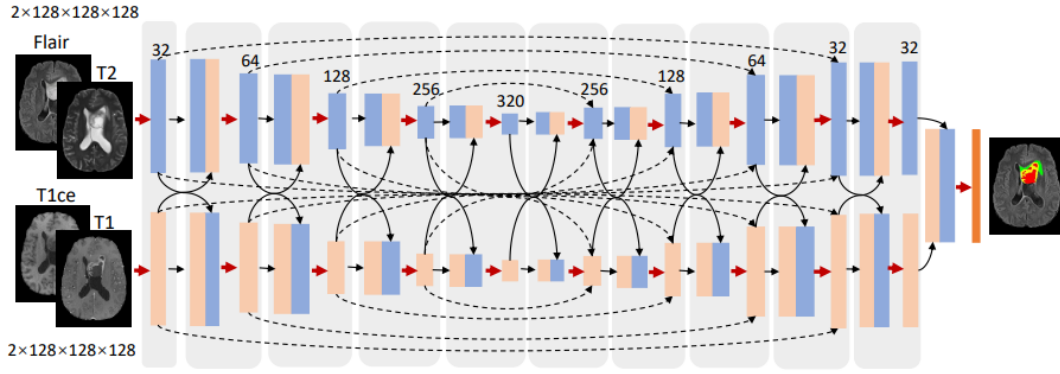
7.4.1 עיבוד מקדים

בדומה ל- $nnU-Net$, מתבצע נימול של ווקסלי (פיקסל תלת מימדי) הסריקה ע"י החסרת ממוצע העוצמה של כלל הווקסלים המייצגים רקמת מוח מכל ווקסל המייצג מוח באופן אינדיבידואלי. ולאחר מכן מתבצעת חלוקה של כל ווקסל כזה בסטיית התקן הכללית. כל הווקסלים שאינם מייצגים רקמת מוח (רקע) יהיו בעלי ערך 0.

7.4.2 מבנה הרשת

Modality Pairing Network מורכבת מענפים מקבילים, כאשר כל ענף מתמקד בזוג שכלולים (במקום להתמקד בארבעת השכלולים). הענפים מחוברים ביניהם לאורך הרשת. המטרה של הנתבים המזווגים הינה לחלץ את המאפיינים הרלוונטיים ביותר לכל Modality, וכן העשרת המידע בעזרת Modalities אחרים (נתייחס ל-Modality בתור שקלול בודד). השכלולים (Modalities) השונים חולקו לשתי קבוצות, כאשר בכל קבוצה ישנם שני שכלולים. מניסויים שונים נבע כי התוצאות המיטביות הושגו ע"י שילוב של T1 עם T1ce וכן T2 עם Flair. שתי קבוצות אלה מוזנות כקלט אל שני הענפים בנפרד, לכן גודל הקלט בכל ענף הינו $2 \times 128 \times 128 \times 128$ (כאשר $128 \times 128 \times 128$ הינו גודל ה-Patch שנדגם לצורך ביצוע האימון). שני הענפים מבוססים על ארכיטקטורת 3D U-Net וזה בעלת נתיב מצמצם (מקודד) ונתיב מרחיב (מפענח) אשר מחוברים ביניהם בעזרת Skip connections. כל בלוק (מסומן באפור באיור) מורכב משתי שכבות קונבולוציה בעלות גרעין בגודל $3 \times 3 \times 3$ אשר לאחר מכן מבוצעת Instance normalization וכן Leaky ReLU כפונקציית אקטיבציה (פונקציה זו דומה ל-ReLU מלבד העובדה שעבור קלטים שליליים מתבצעת הכפלה בקבוע במקום איפוס הפלט). על מנת לבצע מיזוג מיטבי בין מאפיינים מענפים שונים ישנם חיבורים בין שכבות מענפים שונים. הפלטים מענף Flair/T2 ומענף T1/T1ce משורשרים זה לזה על מנת ליצור שני Feature Maps שיהיו קלטים לשכבה הבאה בכל אחד מהענפים (באיור פלט מענף Flair/T2 מסומן בכחול ואילו פלט משכבת T1/T1ce מסומן בוורוד). בסוף הנתביב המרחיב, הפלטים משני הענפים משורשרים יחדיו ועוברים בשכבת Softmax לצורך סיווג לארבע מחלקות סיווג:

- (1) רקע – רקמת מוח בריאה או הרקע סביב המוח
- (2) בצקת
- (3) החלק הלא פעיל של הגידול ונמק
- (4) החלק הפעיל של הגידול.



איור 22 – מבנה רשת Modality-Pairing Network [23]

7.4.3 אימון הרשת

לצורך תהליך האימון, חולצו Patches תלת מימדים מתוך הסריקות בצורה רנדומלית בגודל קבוע של $128 \times 128 \times 128$, כאשר לרשת ישנם שני קלטים (עבור כל ענף). הקלט לענף הראשון הינו שרשרת של Patch משקלול T2 ו-Patch משקלול Flair, ואילו הקלט לענף השני הינו שרשרת של Patch משקלול T1 ו-Patch משקלול T1ce.

גודל ה-Batch שנקבע לאורך האימון הינו קבוע ושווה ל-2.

המודל אומן לאורך 1,000 epochs, ובוצע שימוש ב-Stochastic gradient descent (SGD) optimizer עם מקדם למידה התחלתי של כ-0.01.

פונקציית ה-Loss שבה התבצע שימוש לאורך תהליך האימון הינה שילוב של כמה פונקציות הפסד שונות. פונקציות ההפסד הקלאסיות לביצוע סגמנטציה בתמונות רפואיות הינן Dice loss ו-Cross entropy, אך במקרה זה פונקציות אלה אינן נותנות מענה מספק מפני שהשכלולים השונים בסריקה מפוצלים בין ענפים שונים וזהו דבר העלול לגרום לשונות בין הפרדיקציות בענפים השונים. לצורך כך הוצעה פונקציית הפסד ייעודית Modality Pairing Loss המוגדרת כך:

$$\mathcal{L}_{MP}(X_i^A, X_i^B) = - \frac{\sum_{j=1}^N (X_{ij}^A - \bar{X}_i^A) (X_{ij}^B - \bar{X}_i^B)}{\sqrt{\sum_{j=1}^N (X_{ij}^A - \bar{X}_i^A)^2 \sum_{j=1}^N (X_{ij}^B - \bar{X}_i^B)^2}},$$

איור 23 – פונקציית ההפסד Modality Pairing loss [23]

A, B מייצגים את שני הענפים, N הוא מספר הווקסלים הכללי, $X_{i,j}$ מייצג את הווקסל ה-j בדגימה i. והמקביל עם סימון הגג מייצג את ערך ה-Ground truth עבור הווקסל המתאים. פונקציית ההפסד הכללית ששומשה לצורך תהליך האימון הינה:

$$L = \lambda_1 \cdot L_{Dice} + \lambda_2 \cdot L_{CE} + \lambda_3 \cdot L_{MP};$$

איור 24 – פונקציית ההפסד הכללית ששומשה לצורך אימון מודל Modality Pairing Network [23]

כאשר L_{Dice} הינה פונקציית ההפסד Dice loss, L_{CE} הינה פונקציית ההפסד Cross entropy. ו- L_{MP} הינה פונקציית ההפסד הייעודית Modality Pairing Loss. $\lambda_1, \lambda_2, \lambda_3$ הינם קבועים בעלי הערכים 1, 1, 0.5 בהתאמה.

תהליך האימון ווידוא המודל התבצע בעזרת 5-fold cross validation על סט האימון. בניגוד לשיטות הקודמות שהוצגו, בשיטה זו לא צוין כי בוצע שימוש בשיטות אוגמנטציה לצורך העשרת המידע.

7.4.4 תוצאות המודל

תוצאות המודל על סט האימון :

טבלה 9 – תוצאות מודל Modality Pairing Network על גבי סט האימון של BraTS 2020 [23]

Method	Tumor	Dice	Sensitivity	Specificity	Hausdorff95
Vanilla U-Net	Enhancing Tumor	0.848	0.863	1.000	12.145
	Whole Tumor	0.923	0.909	0.999	4.508
	Tumor Core	0.900	0.880	1.000	3.368
Modality-Pairing learning	Enhancing Tumor	0.863	0.875	1.000	7.179
	Whole Tumor	0.924	0.910	0.999	4.131
	Tumor Core	0.898	0.877	1.000	3.448

בטבלה הנ"ל מוצגות תוצאות האימון של Modality Pairing Network על גבי סט האימון של BraTS 2020. התוצאות מוצגות עבור כל אחת ממחלקות ההערכה של תחרות BraTS עבור כל אחד ממדדי האכיות Dice, Hausdorff p95, sensitivity, specificity (כאשר השניים האחרונים לא משמשים לצורך הערכת המודלים בתחרות). התוצאות של מודל Modality Pairing Network מוצגים אל מול התוצאות של Vanilla U-Net (אשר לא תוצג במסגרת עבודה זו).

תוצאות המודל על סט הולידציה :

טבלה 10 – תוצאות מודל Modality Pairing Network על גבי סט הולידציה של BraTS 2020 עבור מודל יחיד [23]

	Dice			Sensitivity			Specificity			Hausdorff95		
	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC
Mean	0.785	0.907	0.837	0.783	0.901	0.804	1.000	0.999	1.000	32.25	4.39	8.34
StdDev	0.272	0.072	0.178	0.286	0.103	0.211	0.000	0.001	0.000	101.03	5.97	33.64
Median	0.880	0.931	0.902	0.886	0.931	0.891	1.000	1.000	1.000	1.73	2.83	3.16
25quantile	0.790	0.890	0.808	0.795	0.885	0.733	1.000	1.000	1.000	1.00	1.73	1.41
75quantile	0.923	0.949	0.944	0.951	0.964	0.943	1.000	1.000	1.000	3.00	4.24	5.83

בתחילה בוצעה הערכה של כל אחד מהמודלים של Modality Pairing Network שנוצרו במסגרת 5 fold cross validation בתהליך האימון. התוצאות המוצגות בטבלה הנ"ל הינן עבור המודל הבודד אשר הניב את התוצאות הטובות ביותר על גבי סט הולידציה מכלל המודלים שנוצרו במסגרת ה-5 fold cross validation. בטבלה מוצגים מדדי האכיות Dice, sensitivity, specificity, Hausdorff p95 עבור כל אחת ממחלקות ההערכה של התחרות.

לאחר מכן, בוצעה הערכה נוספת על גבי Ensemble של 3 המודלים הטובים ביותר של Modality Pairing Network אשר נוצרו במסגרת ה-5-fold cross validation עם 3 מודלים נוספים של Vanilla U-Net.

תוצאות מודל ה-Ensemble על גבי סט הולידציה מוצגות בטבלה הבאה:

טבלה 11 – תוצאות מודל ה-Ensemble של Modality Pairing Network על גבי סט הולידציה של BraTS 2020 [23]

	Dice			Sensitivity			Specificity			Hausdorff95		
	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC
Mean	0.787	0.908	0.856	0.786	0.905	0.822	1.000	0.999	1.000	35.01	4.71	5.70
StdDev	0.276	0.078	0.130	0.289	0.103	0.175	0.000	0.001	0.000	105.54	7.62	10.17
Median	0.885	0.931	0.907	0.887	0.934	0.895	1.000	1.000	1.000	1.73	2.83	3.00
25quantile	0.809	0.892	0.827	0.801	0.887	0.768	1.000	0.999	1.000	1.00	1.73	1.41
75quantile	0.925	0.950	0.944	0.960	0.967	0.947	1.000	1.000	1.000	2.83	4.47	5.39

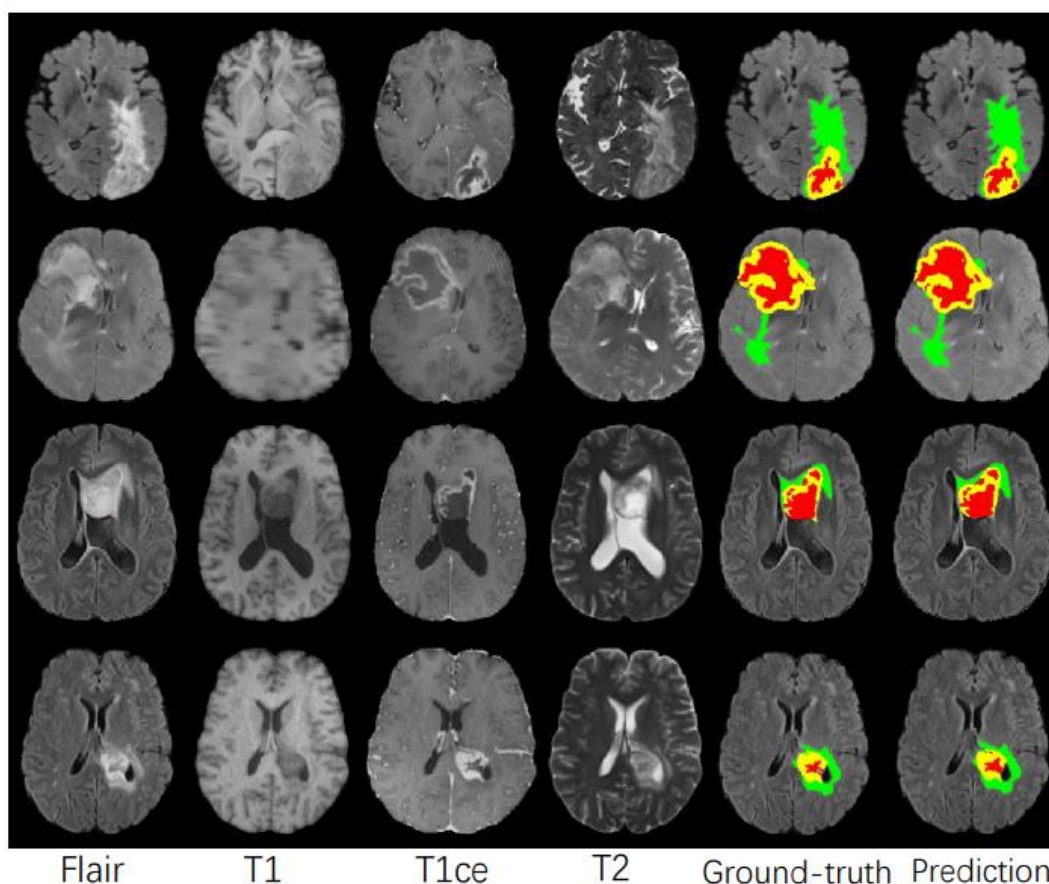
מודל ה-Ensemble הניב תוצאות ממוצעות טובות יותר במדד Dice לעומת המודל היחיד הטוב ביותר, והניב שיפור במדד Hausdorff p95 עבור ליבת הגידול (TC)

תוצאות המודל על סט המבחן:

טבלה 12 – תוצאות מודל ה-Ensemble של Modality Pairing Network על גבי סט המבחן של BraTS 2020 [23]

	Dice			Sensitivity			Specificity			Hausdorff95		
	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC
Mean	0.816	0.891	0.842	0.847	0.911	0.853	1.000	0.999	1.000	17.79	6.24	19.54
StdDev	0.197	0.112	0.244	0.211	0.118	0.225	0.000	0.001	0.001	74.87	28.98	74.78
Median	0.857	0.925	0.925	0.916	0.941	0.931	1.000	0.999	1.000	1.41	2.83	2.00
25quantile	0.788	0.884	0.865	0.823	0.905	0.854	1.000	0.999	1.000	1.00	1.41	1.41
75quantile	0.921	0.950	0.959	0.963	0.968	0.966	1.000	1.000	1.000	2.24	4.66	3.74

המודל הסופי שנבחר לביצוע ההערכה אל מול סט המבחן הוא מודל ה-Ensemble שנבחר אל מול סט הולידציה. בטבלה הנ"ל ניתן לראות את תוצאות המודל על גבי סט המבחן של BraTS 2020 (המורכב מ-166 סריקות).



איור 25 – תוצאות ויזואליות של Modality Pairing Network על דוגמאות מסט האימון [23]

באיור מוצגות תוצאות ויזואליות של המודל על סריקות מסט האימון. בסך הכל מוצגות תוצאות ויזואליות של כ-4 סריקות (4 שורות) כאשר ארבע העמודות השמאליות מייצגות פרוסה של סריקה בכל אחד מארבעת השכילים השונים. ובנוסף מוצגות עמודות עם סימוני הסגמנטציה על גבי הפרוסות, אחת עבור ה-Ground truth והשנייה עבור הפרדיקציה של המודל. בצהוב מסומן החלק הפעיל של הגידול, באדום הנמק והחלק הלא פעיל ובירוק האיזור הבצקתי.

7.4.5 סיכום ומסקנות

שיטת ה-Modality Pairing הינה שיטה ייחודית שבה הקלט מחולק לשתי קבוצות שונות (מיזוג של T1 עם T1ce ומיזוג של T2 עם Flair), בניגוד לשאר המודלים שראינו קודם אשר מקבלים כקלט את איחוד ארבעת השכילים. דבר זה מאפשר חילוף מאפיינים מהענפים המקבילים תוך שילובם לאורך המודל. שיטה זו הוכיחה את עצמה כיעילה וזכתה במקום השני בתחרות BraTS 2020 (אשר אותו חולקת עם מודל H2NF-Net שהוצג קודם לכן).

מסיכום התוצאות על סט המבחן, ניתן לראות כי המדדים הטובים ביותר הושגו על גבי מחלקת ההערכה Whole Tumor אשר כוללת את כלל מרכיבי הגידול, ואילו התוצאות הגרועות ביותר הושגו על גבי מחלקת ההערכה Enhancing Tumor בדומה לשיטות הקודמות שהוצגו.

התוצאות הוויזואליות שמוצגות באיור 25 הינן מרשימות ביחס ל-Ground Truth עבור ארבע הסריקות שנבחרו. עם זאת, השימוש ב-Modality Pairing לא הראה עליונות ביחס לשימוש בארכיטקטורת 3D U-Net כדוגמת nnU-Net.

7.5 Vox2Vox – 3D Generative Adversarial Network (GAN) [24]

Vox2Vox [24] הינה רשת מסוג Generative Adversarial Network, מודל זה מורכב משני רכיבים מרכזיים:

- Generator – מחולל, מטרתו לקבל כקלט סריקת MRI ולפלוט עבודה מפת סגמנטציה מתאימה
- Discriminator – מפענח, מטרתו לקבל כקלט סריקת MRI ומפת סגמנטציה מתאימה עבור אותה הסריקה ולפלוט כקלט את איכות הסגמנטציה על גבי הסריקה הנתונה.

7.5.1 עיבוד מקדים

מתבצע נירמול של ווקסלי הסריקה ע"י החסרת ממוצע העוצמה של כלל הווקסלים המייצגים רקמת מוח מכל ווקסל המייצג מוח באופן אינדיבידואלי. ולאחר מכן מתבצעת חלוקה של כל ווקסל כזה בסטיית התקן הכללית. כל הווקסלים שאינם מייצגים רקמת מוח (רקע) יהיו בעלי ערך 0. נרמול זה מתבצע עבור כל אחד מארבעת הערוצים בנפרד (כאשר כל ערוץ מציין את אחד משקלולי הסריקה).

7.5.2 מבנה הרשת

Vox2Vox הינה רשת מסוג Generative Adversarial Network הבנויה משני רכיבים

7.5.2.1 Generator - מחולל

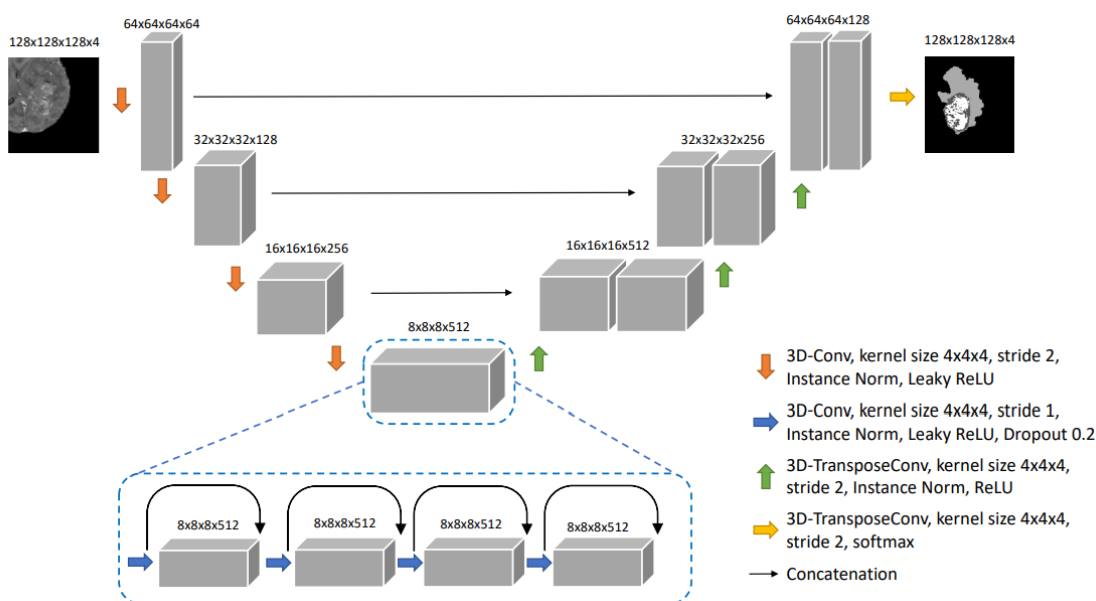
מחולל הינו רכיב מסוג 3D U-Net המקבל כקלט Patch שנדגם רנדומלית מתוך הסריקה בגודל קבוע של $128 \times 128 \times 128 \times 4$ (שרשור של כלל השכלולים המרכיבים את הסריקה).

בנתיב המצמצם של המחולל ישנם ארבעה בלוקים של הורדת מימד כאשר בכל אחד מהם מתבצעת פעולת קונבולוציה עם גרעין בגודל $4 \times 4 \times 4$ ודילוג (Stride) בגודל 2 (על מנת לבצע הורדת מימד). לאחר כל פעולת קונבולוציה מתבצע Instance normalization שלאחריו מופעלת פונקציית אקטיבציה מסוג Leaky ReLU. מספר הפילטרים בפעולת הקונבולוציה הראשונה הינו 64, כאשר לאחר כל פעולת הורדת מימד, מספר זה גדל פי 2.

בתחתית הנתיב המצמצם, ישנם ארבעה בלוקי ResNet המורכבים כל אחד מפעולת קונבולוציה עם גרעין בגודל $4 \times 4 \times 4$ שלאחריו Instance normalization ו-Leaky ReLU כפונקציית אקטיבציה. כל פלט של בלוק קונבולוציה משורשר לקלט של אותו הבלוק ויחד מועברים כקלט לבלוק הבא. לאחר ארבעת בלוקי הקונבולוציה במימד הנמוך ביותר, ישנו נתיב מרחיב שמטרתו להחזיר את ה-Feature map למימדיו המקוריים. נתיב זה מורכב משלושה בלוקים של העלאת מימד, כל אחד מורכב מפעולת קונבולוציה הפוכה עם גרעין בגודל $4 \times 4 \times 4$ ודילוג (Stride) בגודל 2, שלאחריו מתבצעת פעולת Instance normalization ו-Leaky ReLU כפונקציית אקטיבציה. הקלט של כל בלוק בנתיב המרחיב מורכב מהפלט של הבלוק הקודם בנתיב, ומ-Feature map הפלט של הבלוק המקביל בנתיב המצמצם (Skip connection).

הפלט של המחולל מתקבל ע"י ביצוע פעולת קונבולוציה הפוכה (המורכבת מ-4 פילטרים כמספר מחלקות הסיווג) עם גרעין בגודל $4 \times 4 \times 4$ ודילוג בגודל 2. לאחריו, פונקציית אקטיבציה מסוג

softmax שבסופה מתקבלת מפת סגמנטציה עם ההסתברות של כל ווקסל להשתייך לאחת ממחלקות הסיווג השונות.

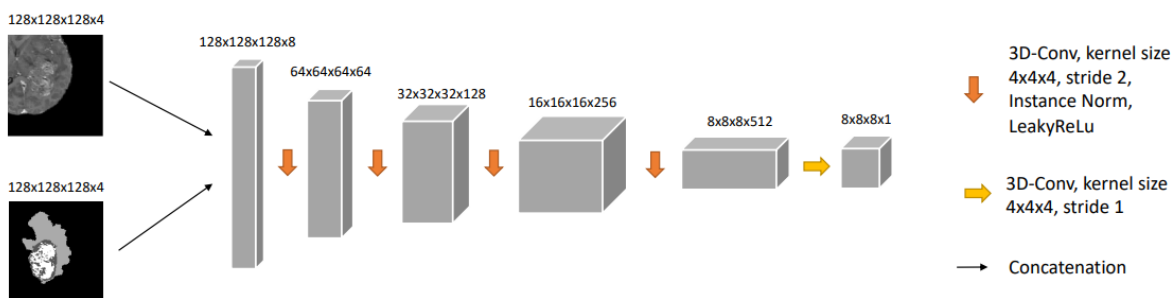


איור 26 – מבנה ה-Generator. [24]

7.5.3 Discriminator - מפענח

המפענח הינו רכיב בעל נתיב מצמצם בלבד (בניגוד למחולל המכיל גם נתיב מצמצם וגם נתיב מרחיב). המחולל מקבל כקלט Patch של סריקה בגודל $128 \times 128 \times 128 \times 4$ וכן מפת סגמנטציה מתאימה (היכולה להיות Ground truth או פלט של המחולל), גם היא בגודל זהה של $128 \times 128 \times 128 \times 4$.

הנתיב המצמצם של המפענח הינו זהה לחלוטין לזה של המחולל. הפלט של המפענח מתקבל ע"י ביצוע קונבולוציה עם פילטר בודד וגרעין בגודל $4 \times 4 \times 4$. פלט המפענח הינו בגודל $8 \times 8 \times 8 \times 1$ אשר מייצג את איכות הסגמנטציה שהתקבלה כקלט על גבי הסריקה הנתונה.



איור 27 – מבנה ה-Discriminator. [24]

7.5.4 אימון הרשת

לצורך תהליך האימון, חולצו Patches תלת מימדים מתוך הסריקות בצורה רנדומלית בגודל קבוע של $128 \times 128 \times 128$, כאשר קלט המחולל הינו שרשרת של ארבעה Patches כאלה (עבור כל אחד

משקלולי סריקת ה-MRI). ואילו קלט המבחין הינו שרשרת של Patches אשר נדגמו מהסריקה משורשר עם ה-Patch המתאים ממפת הסגמנטציה המתאימה.

לצורך האימון בוצע שימוש ב-Adam optimizer עם מקדם למידה התחלתי של כ-0.0002. בנוסף, בוצע Dropout עם הסתברות של 0.2 לאחר כל אחד מבלוקי הקונבולוציה בתחתית המחולל.

מכיוון ש-Vox2Vox מורכב משני רכיבים (המחולל והמבחין) בוצע שימוש בשתי פונקציות הפסד:

(1) פונקציית ההפסד L_D – Discriminator loss

הגדרת הפונקציה:

$$L_D = L_2 [D(x, y), 1] + L_2 [D(x, \hat{y}), 0]$$

פונקציה זו משמשת כפונקציית ההפסד של המבחין, $\hat{y}, y, x$ מסמנים את התמונה המקורית, מפת הסגמנטציה שהתקבלה כ-*Ground truth* ומפת הסגמנטציה שהתקבלה כפרדיקציה של המחולל בהתאמה, D מסמן את פלט המבחין על זוג הקלטים.

פונקציית ההפסד מחשבת את פונקציית ההפסד L_2 (ריבוע ההפרש) בין פלט המבחין על גבי הסריקה המקורית עם מפת הסגמנטציה המהווה *Ground truth* לבין הווקטור הקבוע 1. ובנוסף את פונקציית ההפסד L_2 בין פלט המבחין על גבי הסריקה המקורית עם מפת הסגמנטציה שהתקבלה כפרדיקציה המחולל לבין הווקטור הקבוע 0.

פונקציית ההפסד הינה הסכום של שתי פונקציות L_2 שחושבו.

(2) פונקציית ההפסד L_G – Generator loss

הגדרת הפונקציה:

$$L_G = L_2 [D(x, \hat{y}), 1] + \alpha GDL(y, \hat{y})$$

פונקציה זו משמשת כפונקציית ההפסד של המחולל, גם כאן, $\hat{y}, y, x$ מסמנים את התמונה המקורית, מפת הסגמנטציה שהתקבלה כ-*Ground truth* ומפת הסגמנטציה שהתקבלה כפרדיקציה של המחולל בהתאמה, D מסמן את פלט המבחין על זוג הקלטים.

פונקציית ההפסד מחשבת את פונקציית ההפסד L_2 בין פלט המבחין על גבי הסריקה המקורית עם מפת הסגמנטציה שהתקבלה כפרדיקציה המחולל לבין הווקטור הקבוע 1. ומבצעת חיבור עם פונקציית ההפסד Dice loss (GDL) בין מפת הסגמנטציה המהווה *Ground truth* לבין מפת הסגמנטציה שהתקבלה כפרדיקציה המחולל. פונקציית ה-Dice loss מוכפלת בפרמטר α .

ניתן לראות כי כאשר α שואף לאינסוף, המודל מתנהג כ-3D U-Net קלאסי, שכן מתעלמים לחלוטין מפלט המבחין.

בדומה למודלים הקודמים שהוצגו, גם מודל זה אומן בתצורת Fold cross validation עם 10 Folds. בשונה מהמודלים הקודמים בהם ה-Ensemble בוצע על גבי כל ווקסל בנפרד (ע"י מיצוע), ב-Vox2Vox הוצעה שיטה לביצוע Ensemble המתחשבת בשכני הווקסל ע"י אימון רשת CNN שמקבלת כקלט את שרשרת הפלטים של כל המודלים של ה-Ensemble ומחזירה פלט בגודל

128x128x128x4, המשתמשת בפעולת קונבולוציה יחידה עם גרעין בגודל 3x3x3 ופונקציית אקטיבציה מסוג softmax.

7.5.5 שיטות אוגמנטציה להעשרת המידע

על מנת להקטין את הסיכוי ל-Overfitting, בוצעה אוגמנטציה על גבי סט האימון על מנת להעשיר את המידע בצורה מלאכותית. בין שיטות האוגמנטציה שבוצעו:

- היפוך רנדומלי סביב הצירים
- סיבוב רנדומלי סביב הצירים בזווית של עד כ-30 מעלות
- תיקון גאומטרי – עם קבועים המוגדרים בטווח של 0.8 – 1.2

7.5.6 תוצאות המודל

תוצאות המודל על סט האימון כפונקציה של α

טבלה 13 – תוצאות מודל Vox2Vox על גבי סט האימון כפונקציה של α [24]

α	Dice score [%]			Hausdorff distance 95 [mm]		
	WT	TC	ET	WT	TC	ET
0	58.96	23.57	37.63	77.74	104.35	73.84
1	86.36	72.48	62.50	13.12	17.26	40.50
3	92.98	90.82	79.12	4.49	4.06	31.32
5	93.21	91.70	79.31	3.80	3.18	30.91
10	87.50	80.66	70.55	12.09	9.92	34.40
25	84.35	83.01	71.04	9.62	8.75	32.53
50	87.87	76.79	61.54	9.55	8.33	31.57
100	91.81	89.22	77.31	5.28	5.03	31.63
200	88.72	90.16	77.62	6.33	3.38	31.19
250	86.65	85.67	78.35	9.32	8.90	31.87

בתחילה, הורץ מודל Vox2Vox על סט האימון של BraTS 2020 תוך בחינת ערכי α שונים. בטבלה מוצגים מדדי האכיות Dice ו- Hausdorff distance p95 עבור כל אחת משלוש מחלקות ההערכה השונות. ניתן לראות כי הערכים המיטביים התקבלו עבור ערך של $\alpha=5$.

לאחר קביעת הפרמטר α להיות 5, בוצע תהליך אימון 10-fold cross validation שבמסגרתו נוצרו 10 מודלים מסוג Vox2Vox שהפלט שלהם שורשר יחדיו כקלט לרשת ה-Ensembler.

תוצאות מודל ה-Ensembled על גבי סט האימון, סט הולידציה וסט המבחן:

טבלה 14 – סיכום תוצאות מודל *Ensembled Vox2Vox* על גבי סטי התונים של *BraTS 2020* [24]

dataset		Dice score [%]			Hausdorff distance 95 [mm]		
		WT	TC	ET	WT	TC	ET
training	Mean	91.63	89.25	79.56	3.66	3.52	30.04
	StdDev	6.36	11.49	24.74	3.87	4.57	96.67
	Median	93.39	92.50	87.16	2.44	2.23	1.73
validation	Mean	89.26	79.19	75.04	6.39	14.07	36.00
	StdDev	8.28	22.30	28.68	11.44	47.56	105.28
	Median	91.75	88.13	85.87	3.0	3.74	2.23
test	Mean	87.20	81.14	78.67	6.44	24.36	18.95
	StdDev	14.07	25.47	20.17	11.09	79.78	74.91
	Median	91.42	90.59	83.65	3.16	3.0	2.23

בטבלה הנ"ל מוצג סיכום תוצאות מודל ה- *Ensembled Vox2Vox* על גבי סט האימון, סט הולידציה וסט המבחן של *BraTS 2020*. התוצאות מורכבות מהערכת מדדי האיכות *Dice loss* ו- *Hausdorff distance p95* על גבי כל אחת משלוש מחלקות ההערכה של התחרות.

7.5.7 סיכום ומסקנות

מודל *Vox2Vox* הינו מודל המורכב משני רכיבים נפרדים, מחולל שתפקידו לחשב מפת סגמנטציה עבור סריקה נתונה ומפענח שתפקידו לתת ציון למפת סגמנטציה על גבי סריקה נתונה. מודל זה הניב תוצאות תחרותיות ב-*BraTS 2020* אך פחות טובות ביחס למודלים האחרים שהוצגו בעבודה זו. גם כאן, התוצאות הטובות ביותר שהמודל הניב היו ביחס למחלקת ההערכה *Whole Tumor* ואילו התוצאות הגרועות ביותר התקבלו ביחס למחלקת ההערכה *Enhancing Tumor*. לא ניכר יתרון בשימוש בארכיטקטורת GAN לצורך סגמנטציה של גידולי מוח מסוג גליומה בסריקות מוח-MRI, שכן מודל זה הניב את התוצאות הגרועות ביותר ביחס לשאר המודלים שהוצגו בעבודה.

8 ניתוח השוואתי בין השיטות השונות

בפרק זה יוצגו מדדי האיכות שעל פיהם בוצעה הערכת המודלים אל מול סט המבחן בתחרות BraTS, מדד ה-Dice coefficient ומדד Hausdorff distance p95. לאחר מכן יוצג ניתוח השוואתי של השיטות שהוצגו בפרק הקודם על סמך מדדים אלו.

8.1 מדדי האיכות

8.1.1 Dice Coefficient

Dice coefficient הינו מדד להערכת הדמיון בין שתי קבוצות ערכים (במקרה הספציפי בו העבודה עוסקת, בין מטריצה תלת מימדית המכילה את מפת הסיווג שחושבה עבור כל אחד מהווקסלים לבין מטריצת ה-Ground Truth הזהה במימדיה למטריצת הסיווג שהמודל חישב).

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

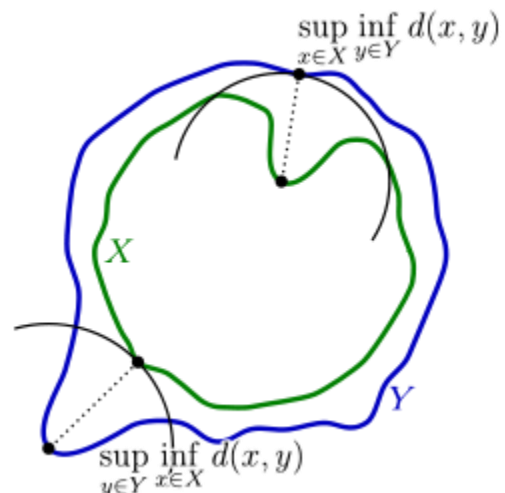
איור 28 - מדד Dice Coefficient [25]

המדד מחושב ע"י חישוב גודל החיתוך בין הקבוצות כפול 2, חלקי סכום הגדלים של כל קבוצה בנפרד. לכן, כאשר יש דמיון מוחלט (שתי הקבוצות זהות) אז ערך המדד יהיה 1. כאשר הקבוצות זרות לחלוטין ערך המדד יהיה 0.

לכן, בתחרות BraTS השאיפה היא למקסם את מדד ה-Dice הממוצע (על גבי כל הסריקות בסט המבחן) עבור כל אחת ממחלקות ההערכה של התחרות.

8.1.2 Hausdorff Distance p95

מרחק האוסדורף, הינו מדד אשר נועד לחשב כמה שתי קבוצות שונות רחוקות זה מזו במרחב מטרי. שתי קבוצות נחשבות קרובות זה לזו אם כל נקודה בכל אחת מהקבוצות קרובה לנקודה כלשהי בקבוצה הנגדית.



איור 29 – איור המתאר את מרחק האוסדורף [26]

מרחק האוסדורף, הינו המקסימום בין כל המרחקים המינימליים בין שתי נקודות בקבוצות שונות. במקרה הספציפי בו העבודה עוסקת, הקבוצה הראשונה הינה קבוצת כל הווקסלים המשתייכים לאחת ממחלקות הסיווג המקוננות שתוארו לעיל על סמך חיזוי המודל, והקבוצה השנייה הינה

קבוצת כל הווקסלים המשתייכים למחלקת הסיווג המקוננת המתאימה על סמך ה-Ground Truth. מרחק האוסדרוף בין שתי קבוצות אלה, יהיה המקסימום בין כלל המרחקים המינימליים במרחב מטרי בין שני ווקסלים המשתייכים לקבוצות מנוגדות. בתחרות BraTS, לא משתמשים בערך המקסימלי לביצוע ההערכה, אלא באחוזון ה-95 של כלל המרחקים, וזאת על מנת לסנן מקרי קצה אשר עלולים לגרום לרעש בתוצאה. לכן, בתחרות BraTS השאיפה היא למזער את מדד ה-Hausdorff Distance p95 הממוצע (על גבי כל הסריקות בסט המבחן) עבור כל אחת ממחלקות ההערכה של התחרות.

8.2 השוואת השיטות השונות על בסיס נתוני BraTS ע"פ המדדים שתוארו

בסעיף זה נבצע ניתוח השוואתי תוך התייחסות לסט המבחן של בסיס הנתונים של תחרות BraTS 2020 ששימש את החוקרים אשר השתתפו בתחרות. במסגרת ההשוואה, תוצג טבלה המסכמת את תוצאות כלל השיטות שהוצגו בעזרת שני מדדי האיכות שתוארו לעיל על גבי כל אחת מ-3 מחלקות ההערכה של תחרות BraTS.

מניתוח תוצרי טבלת ההשוואה עולות מספר נקודות. ראשית המודלים אשר הוצגו בעבודה זו השיגו תוצאות טובות במדדי האיכות שנקבעו לשימוש במסגרת תחרות BraTS 2020 ותוצרי הסגמנטציה של מודלים אלו הינם איכותיים ומניבים תוצאות ויזואליות מרשימות.

שנית, למרות שמודל nnU-Net הוא זה שניצח בתחרות BraTS 2020 ניתן לראות כי אין מודל שהניב תוצאות עיליות בכל מדדי ההערכה ביחס לשאר המודלים שהוצגו. בנוסף, ההבדלים בין המודלים השונים אינם מהותיים ביחס למדדי האיכות השונים על גבי מחלקות הסיווג השונות. ההבדל במדד ה-Dice הממוצע על גבי 3 מחלקות ההערכה בין המודל שהניב את התוצאות הטובות ביותר לבין המודל שהניב את התוצאות הפחות טובות עומד על 3% בלבד. וכן ההבדל במדד Hausdorff Distance p95 הממוצע על גבי 3 מחלקות ההערכה בין המודל שהניב את התוצאות הטובות ביותר לבין המודל שהניב את התוצאות הפחות טובות עומד על 5mm.

ניכר כי כל אחד מ-5 המודלים שהוצגו בעבודה זו הצליחו להשיג תוצאות יחסית דומות עבור אותה משימת סגמנטציה (הכוללת זיהוי וחלוקה של הגידול ל-3 מרכיבים שונים), על גבי אותו בסיס נתונים ובעזרת שימוש באותם מדדי איכות בדיוק.

לפיכך, ניכר כי שיטות מבוססות ארכיטקטורת מקודד-מפענח מניבות תוצאות טובות בביצוע סגמנטציה של גידולי מוח מסוג גליומה ולא ניתן לקבוע כי שיטה אחת עולה בהכרח על השיטות האחרות, אלא ישנן מספר שיטות שונות אשר מסוגלות להביא לתוצאות דומות במשימה זו.

טבלה 15 – סיכום תוצאות המודלים שהוצגו על סט המבחן של BraTS 2020. עבור כל אחד מהמדידים, מסומן הטוב ביותר בהדגשה והרע ביותר באדום.

הערה	Dice Coefficient				Hausdorff Distance p95 (mm)				מודל
	Whole Tumor	Tumor Core	Enhancing Tumor	Mean	Whole Tumor	Tumor Core	Enhancing Tumor	Mean	
זוכה במקום הראשון בתחרות BraTS 2020	0.8895	0.8506	0.8203	0.8535	8.498	17.337	17.805	14.5467	nnU-Net [11]
זוכה במקום השלישי בתחרות BraTS 2020	0.8828	0.8433	0.8177	0.8479	5.2176	17.9697	13.4298	12.2057	SA-Net [17]
זוכה במקום השני בתחרות BraTS 2020	0.88790	0.85375	0.82775	0.8564	4.53440	16.92065	13.04490	11.5	H2NF-Net [10]
חולק את המקום השני בתחרות H2NF-Net עם BraTS 2020	0.891	0.842	0.816	0.8497	6.24	19.54	17.79	14.5233	Modality Pairing Network [23]
	0.8720	0.8114	0.7867	0.8234	6.44	24.36	18.95	16.5833	Vox2Vox [24]

9 סיכום ומסקנות

בעבודה זו עסקנו בבעיית ביצוע סגמנטציה של גידולי מוח מסוג גליומה בסריקות מוח MRI ע"י זיהוי וסיווג כל ווקסל בסריקה לאחד משלוש הרקמות המרכיבות את הגידול. הגדרנו את עולם הבעיה בהיבט הרפואי ובהיבט מדעי המחשב והצגנו את בסיס הנתונים של תחרות BraTS המכיל סריקות מוח MRI של נבדקים הסובלים מגידול מסוג גליומה. לאחר מכן, הצגנו רקע קצר על למידה עמוקה המסביר מהי רשת נוירונים, מהם הרכיבים ממנה היא מורכבת וכיצד מתבצע תהליך הלמידה של רשת מסוג זה. הסברנו מהי רשת מסוג CNN (Convolutional Neural Networks) ובפרט רשתות נוירונים מסוג מקודד-מפענח.

בלב העבודה סקרנו מודלים שונים המבוססים ארכיטקטורת מקודד-מפענח אשר השתתפו בתחרות BraTS לצורך הערכת ביצועי סגמנטציה של גידולי מוח מסוג גליומה בסריקות מוח MRI. מודלים אשר הניבו תוצאות טובות בתחרות וזכו במקומות הראשונים. הסקירה התייחסה לארכיטקטורת המודלים השונים, לשלב העיבוד המקדים של המידע, תהליך אימון המודל בדגש על פונקציית ההפסד, שיטות אוגמנטציה שונות להעשרה מלאכותית של המידע ותוצאות המודלים השונים על גבי סט האימון, סט הולידציה וסט המבחן של בסיס נתוני BraTS.

לבסוף, הצגנו את סיכום תוצאות המודלים השונים ביחס למדדי האכיות Dice ו Hausdorff Distance p95 על גבי סט המבחן של תחרות BraTS בהתחשב בכל אחת ממחלקות ההערכה של התחרות. במהלך העבודה נמצא שעולם הסגמנטציה של גידולי מוח מסוג גליומה בסריקות מוח MRI הינו מורכב ואין שיטה אחת המכריעה בסוגיה זו. עם זאת, ניכר כי שיטות מבוססות ארכיטקטורת מקודד-מפענח מניבות תוצאות טובות והינן הפופולריות ביותר כיום לצורך משימה זו.

עבודה מסכמת זו מתבססת על הפרויקט שלי בנושא "מערכת לסגמנטציה תלת מימדית לגידול גליומה מתמונות מוח MRI" שבמסגרתה פותחה מערכת מבוססת ארכיטקטורת מקודד-מפענח מסוג 3D U-Net לצורך ביצוע סגמנטציה של גידולי מוח מסוג גליומה בסריקות מוח MRI, וכן מהווה המשך ישיר לפרויקט ע"י ביצוע סקירה מקיפה של שיטות שונות לביצוע אותה המשימה. להמשך מחקר, ניתן לבחון את ביצועי הסגמנטציה של המודלים שהוצגו על גבי בסיסי נתונים המכילים תמונות תלת מימדיות מעולמות שונים (ותמונות רפואיות בפרט), וכן לבחון כיצד ניתן להפוך מודלים אלה לגנריים, דבר אשר יאפשר למודלים אלה לבצע סגמנטציה איכותית על גבי תמונות תלת מימדיות בעלות מאפיינים שונים, מימדים שונים וכן מחלקות סיווג שונות, וכל זה תוך שינויים מינימליים בהגדרת המודל.

- [1] A. Myronenko, "3D MRI brain tumor segmentation using autoencoder regularization," in *International MICCAI Brainlesion Workshop*, 2018, pp. 311-320.
- [2] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," *nature*, vol. 521, pp. 436-444, 2015.
- [3] A. S. Becker, M. Marcon, S. Ghafoor, M. C. Wurnig, T. Frauenfelder and A. Boss, "Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer," *Investigative radiology*, vol. 52, pp. 434-440, 2017.
- [4] M. Drozdal, G. Chartrand, E. Vorontsov, M. Shakeri, L. Di Jorio, A. Tang, A. Romero, Y. Bengio, C. Pal and S. Kadoury, "Learning normalized inputs for iterative estimation in medical image segmentation," *Medical image analysis*, vol. 44, pp. 1-13, 2018.
- [5] M. Havaei, N. Guizard, H. Larochelle and P.-M. Jodoin, "Deep learning trends for focal brain pathology segmentation in MRI," in *Machine Learning for Health Informatics*, Springer, 2016, pp. 125-148.
- [6] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal and P.-M. a. L. H. Jodoin, "Brain tumor segmentation with deep neural networks," *Medical image analysis*, vol. 35, pp. 18-31, 2017.
- [7] S. Jadon, "A survey of loss functions for semantic segmentation," in *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2020, pp. 1-7.
- [8] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus and K. H. Maier-Hein, "No new-net," in *International MICCAI Brainlesion Workshop*, 2018, pp. 234-244.
- [9] M. BRATS, "The Multimodal Brain Tumor Segmentation Challenge," [Online]. Available: <http://braintumorsegmentation.org/>.
- [10] H. Jia, W. Cai, H. Huang and Y. Xia, "H2NF-Net for Brain Tumor Segmentation Using Multimodal MR Imaging," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop*,

BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6, 2021.

- [11] F. Isensee, P. F. Jaeger, P. M. Full, P. Vollmuth and K. H. Maier-Hein, "nnU-Net for Brain Tumor Segmentation," *arXiv preprint arXiv:2011.00848*, 2020.
- [12] "Chartrand, G., Cheng, P. M., Vorontsov, E., Drozdal, M., Turcotte, S., Pal, C. J., ... & Tang, A. (2017). Deep learning: a primer for radiologists. *Radiographics*, 37(7), 2113-2131."
- [13] "Rectifier (neural networks)," [Online]. Available: [https://en.wikipedia.org/wiki/Rectifier_\(neural_networks\)](https://en.wikipedia.org/wiki/Rectifier_(neural_networks)).
- [14] "Zhang, Y. D., Pan, C., Sun, J., & Tang, C. (2018). Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. *Journal of computational science*, 28, 1-10."
- [15] V. Badrinarayanan, A. Kendall and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, pp. 2481-2495, 2017.
- [16] F. Isensee, P. F. Jager, S. A. Kohl, J. Petersen and K. H. Maier-Hein, "Automated design of deep learning methods for biomedical image segmentation," *arXiv preprint arXiv:1904.08128*, 2019.
- [17] Y. Yuan, "Automatic Brain Tumor Segmentation with Scale Attention Network," *BrainLes 2020*, pp. 285-294, 2020.
- [18] J. Hu, L. Shen and G. Sun, "Squeeze-and-excitation networks," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018.
- [19] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.
- [20] Y. Yuan, M. Chao and Y.-C. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance," *IEEE transactions on medical imaging*, 2017.
- [21] T.-Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollar, "Focal loss for dense object detection," *Proceedings of the IEEE international conference on computer vision*, 2017.

- [22] H. Jia, Y. Xia, W. Cai and H. Huang, "Learning high-resolution and efficient non-local features for brain glioma segmentation in MR images," *Medical Image Computing and Computer Assisted Intervention--MICCAI 2020: 23rd International Conference, Lima, Peru, October 4--8, 2020, Proceedings, Part IV* 23, 2020.
- [23] Y. Wang, Y. Zhang, F. Hou, Y. Liu, J. Tian, C. Zhong, Y. Zhang and Z. He, "Modality-pairing learning for brain tumor segmentation," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020*, 2021.
- [24] M. D. Cirillo, D. Abramian and A. Eklund, "Vox2Vox: 3D-GAN for brain tumour segmentation," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020*, 2021.
- [25] "Sørensen–Dice coefficient," [Online]. Available: https://en.wikipedia.org/wiki/S%C3%B8rensen%E2%80%93Dice_coefficient.
- [26] "Hausdorff distance," [Online]. Available: https://en.wikipedia.org/wiki/Hausdorff_distance.
- [27] "Activation function," [Online]. Available: https://en.wikipedia.org/wiki/Activation_function.
- [28] S. Ronaghan, "Deep Learning: Which Loss and Activation Functions should I use?," 2018. [Online]. Available: <https://towardsdatascience.com/deep-learning-which-loss-and-activation-functions-should-i-use-ac02f1c56aa8>.

Abstract

Glioma is a primary brain tumor (A tumor which originates from brain cells). Glioma can be classified as Low-Grade Glioma (LGG) or High Grade Glioma, where the second is more aggressive, develops very quickly and requires surgical procedures. Chances of recovery are low.

Nowadays, the primary method for diagnosing brain tumors (and especially gliomas) is MRI (Magnetic Resonance Imaging). MRI is a device that can generate a 3D modeling of the human brain. This modeling helps to identify the tumor and its location, evaluate the volume of the tumor and the impact of the tumor over adjacent tissues. MRI is composed of four modalities: T1, T1c (Contrast enhanced), T2 and Flair, while the four modalities are complimentary. Nowadays, the diagnosis is done manually by the surgeon or the radiologist. This diagnosis is not necessarily accurate, doesn't necessarily classify the different tissues that compose the tumor and doesn't provide an estimation of the volume of the tumor. Hence, an automatic 3D segmentation tool of Gliomas in MRI scans is needed. Such a tool will help the surgeon or the radiologist to interpret the resulting scan, it will save efforts and time, it will help monitor an existing tumor and give an exact estimation of its location and the tissue that composes it.

In the recent years, computational abilities have developed extremely developed, and the field of deep learning gained momentum. Therefore, researchers started to use deep learning widely for 3D segmentation of different pathologies in medical scans and especially gliomas in brain MRI scans.

In this work, a comparison of 3D segmentation methods of gliomas was done. Especially neural networks based on encoder-decoder architecture. All the methods that were introduced in this work competed in BraTS 2020 competition using the same 3D MRI scans dataset of patients that suffer from gliomas. The comparison was based on two evaluation metrics which are very common in evaluating segmentation in medical scans. The metrics are Dice and Hausdorff Distance. The comparison showed that encoder-decoder based methods are efficient and produce satisfactory results in 3D segmentation of gliomas.

Table of Contents

1	Drawing List	4
2	Abstract	6
3	Introduction	7
4	BraTS Dataset	10
5	Defining the Problem	11
5.1	Medical Aspect	11
5.2	Computational Aspect	11
6	Deep Learning – a Short Review	13
6.1	Background	13
6.2	Neural Networks	13
6.2.1	Learning Procedure of Neural Network	15
6.2.2	Forward Pass	15
6.2.3	Activation Function	15
6.2.4	Loss Function	16
6.2.5	Backward Pass	16
6.3	Convolutional Neural Networks	16
6.3.1	What is a Convolution	16
6.3.2	Network Architecture	17
6.3.3	Training the Model	20
6.4	Neural Networks Based on Encoder-Decoder Architecture	21
7	Different Methods for Glioma Segmentation in 3D Brain MRI	23
7.1	Classic 3D U-Net	23
7.2	Scale Attention Network	29
7.3	Hybrid High-resolution and Non-local Feature Network	36

7.4	Modality-Pairing Network	43
7.5	Vox2Vox – 3D Generative Adversarial Network (GAN)	49
8	Comparison of the Different Methods.....	54
8.1	Evaluation Metrics	54
8.1.1	Dice Coefficient.....	54
8.1.2	Hausdorff Distance p95.....	54
8.2	Comparison of the Different Methods Over BraTS Test Set.....	55
9	Summary and Conclusions.....	57
10	Bibliography	58

The Open University of Israel
Department of Mathematics and Computer Science

Comparison of 3D Segmentation Methods of Gliomas

Final Paper submitted as a partial fulfillment of the requirements
Towards an M.Sc degree in Computer Science
The Open University of Israel
Department of Mathematics and Computer Science

By
Ariel Halachmi

Prepared under the supervision of Dr. Maya Herman
June 2023