

מבט מקרוב על מודלי דיפוזיה



מאת פרופ' מיכאל אלעד

תחילת הסיפור – ניקוי רעש מתמונה

למרבה ההפתעה, הסיפור שלנו על מודלי דיפוזיה מתחיל בתחנה בתחום אחר לחלוטין, במשימה שנחשבת לחלק מהקלאסיקה בעיבוד תמונות – ניקוי תמונה מרעש. אסביר: תמונה מיוצגת כטבלה גדולה למדי של מספרים שמתארים את הגוון שבכל

במאמר זה ברצוני לקחת את הקורא לסיור היכרות עם תחום מרתק שאליו נכנסתי לאחרונה, שנקרא "מודלי דיפוזיה". תחום זה הוא חלק נכבד מזירה רחבה יותר שנקראת "בינה מלאכותית יוצרת" (generative AI), שמדובר עליה רבות בשנים האחרונות בשל תוצאותיה המרתקות ובשל השפעתה הרחבה.



איור 1. משמאל לימין: תמונת מקור נקייה, גרסתה הרועשת ותוצאת ניקוי רעש מתמונה זו



איור 2. תמונות שנוצרו יש מאין בעזרת GAN

קעת נעזוב נושא זה ונעבור, לכאורה, לנושא אחר לחלוטין.

ייצור תמונות יש מאין

מהפכת ה-AI פרצה לחיינו בשנת 2012, עם פרסום המאמר המכונן של הינטון (Hinton) ותלמידיו על סיווג מידע ויזואלי בדיוקים גבוהים בעזרת רשתות נוירונים עמוקות. זמן קצר לאחר מכן, בשנת 2014, העלו קבוצות מחקר מובילות בעולם את השאלה אם ניתן לייצר תמונות יש מאין בעזרת כלי AI. כלומר, אם נוכל לכתוב תוכנת מחשב ש"תהזה" תמונות שתיראינה איכותיות כאילו צולמו. הרעיון הוא להציג לתוכנה זו שפע של תמונות דוגמה, ותהליך הלמידה יטמיע את התנהגותן ויוכל לייצר תמונות חדשות בעלות אופי דומה. כלים רבים ושונים הוצעו לפתרון בעיה זו ברמות שונות של הצלחה, ובהם שיטות כמו VAE, NF, GAN, ו-EBM, וגם מודלי דיפוזיה שבהם מתמקד מאמר זה. באיור 2 מובאות תמונות שיוצרו בעזרת GAN, וניכר שאיכותן קרובה למושלמת.



מיקום בה. תמונה מזוהמת ברעש פירושה שהערכים בטבלה זו משתנים כשנוספים עליהם מספרים אקראיים. נניח שהמספרים האקראיים הללו נדלו מחוק פילוג גאوسی. משימתנו היא בניית תהליך חישובי (אלגוריתם) שינקה רעש זה תוך כדי מאמץ לשמר את התמונה המקורית על כל פרטיה, במידת האפשר. אלגוריתם ניקוי כזה נקרא denoiser – מונח שאשתמש בו בהמשך מאמר זה. באיור 1 מוצגות תמונה מקורית, גרסתה הרועשת ותוצאת הניקוי של denoiser כלשהו (אחד מני רבים המבוסס על למידה עמוקה). ניכר שהאלגוריתם המוצג הוא בעל יעילות מפתיעה בניקוי הרעש.

ניקוי רעש מתמונה היא משימה יסודית בעיבוד תמונות, וככזו היא נלמדה באינטנסיביות במהלך ארבעת העשורים האחרונים. מטבע הדברים, עם הגעת מהפכת ה-AI בעשור האחרון הצטרפו למשפחת הפתרונות לבעיה זו גם שיטות המבוססות על למידת מכונה ועל למידה עמוקה. בראייה רחבה ניתן לומר שעד כה הצטברו יותר מ-30,000 מאמרים שעוסקים בבעיה זו ובדרכים לפתרונה, וככלל ניתן לומר שמשימה זו באה על פתרונה.

מופעלת סדרה מחזורית של צעדים, כשבכל אחת מהן מבוצע ניקוי רעש מחד, והזרקת רעש חדש מאידך, ועוד פעולות פשוטות למדי. כעבור כאלף מחזורים כאלה שיתוכננו בקפידה תיווצר תמונה איכותית. ועוד דבר שחשוב להבין – זהו תהליך סטוכסטי, כלומר אם תהליך זה יופעל שוב ושוב, הוא יניב תמונות שונות זו מזו, כשכל אחת מהן היא דגימה איכותית מחוק הפילוג $P(x)$.

בואו נצלול לרגע לתוך תהליך זה כדי להבין את מקורותיו ואת דרך פעולתו. נניח שפונקציית ההסתברות $P(x)$ ידועה לנו ושבוצענו לדגום ממנה. כיצד נעשה זאת? הינה הצעה ראשונית: נתחיל בתמונה שרירותית כלשהי (כנראה תמונת רעש) x_0 ונעדכן אותה באמצעות המשוואה האיטרטיבית הזאת:

$$x_{k+1} = x_k + a \cdot \nabla \log P(x_k) + b \cdot z_k$$

כשמ- x_0 ניצור את x_1 וממנו את x_2 וכך הלאה. משוואה זו נשענת על תוצאה משנת 1908 שאליה הגיע פיזיקאי ידוע בשם לנג'ווין (Langevin) שעסק בחקר הדינמיקה של מולקולות בנוזלים.

בביטוי שלמעלה הוספת $a \cdot \nabla \log P(x_k)$ ל- x_k מעדכנת את התמונה באמצעות טיפוס במשטח הפונקצייה $\log P(x)$, ובכך הופכת התמונה לסבירה יותר ויותר. התוספת $b \cdot z_k$ מזריקה רעש חדש לתמונה ומרעידה את כל התהליך כדי לקבל בסופו של דבר דגימה קבילה. בתכנון קפדני של המקדמים a ו- b תהליך זה יכול להוביל לדגימה ראויה מחוק הפילוג המיוחל, $P(x)$. אלא שתהליך זה נשען על הצורך בידעית $\nabla \log P(x_k)$, וכעת עולה השאלה מניין יושג ביטוי זה, המוכר בסטטיסטיקה בשם score function. וכאן אנו נקראים לתוצאה מבריקה משנת 1961 שאליה הגיע סטטיסטיקאי יפני בשם מיאסאווה (Miyasawa), שהראה שניתן לקרב את ה- score בדיוק רב

חשוב להבהיר שמבחינה מתמטית משימת ייצור תמונות יש מאין היא בעיית דגימה מחוק פילוג. כלומר, אנו מניחים שתמונה x היא ישות אקראית שניתן לייחס לה הסתברות $P(x)$, והזיית תמונות אינה אלא הגרלה מחוק פילוג זה, ממש כשם שבזריקת קובייה מוגרל ערך בין 1 ל-6 מחוק פילוג אחיד ודיסקרטי.

כאמור, רבים וטובים עסקו בשאלת לימוד חוק הפילוג $P(x)$ של תמונות והגרלה ממנו, והוצעו פתרונות איכותיים ביותר הכוללים ביצועים מרשימים למדי. תחום זה מרתק, ולו רק משום שמעניין לצפות בתמונות מלאכותיות אלה על מלוא מורכבויותיהן. אך מלבד זאת תחום זה חשוב ברמה המעשית, כיוון שהיכולת הזאת פירושה היכרות בלתי אמצעית עם חוק הפילוג של תמונות טבעיות, ואפשר לרתום היכרות זו לפתירת בעיות מעשיות רבות בעיבוד תמונות, כגון שחזור תמונות מקלקולים, דחיסת תמונות ועוד, וכן יש בה כדי להוביל לשיטות חדשות שתעלינה באיכות תוצריהן על כל שיטה קודמת.

מודלי דיפוזיה להפקת תמונות יש מאין

השאלה שנתמקד בה כעת היא זו: האם ניתן למנף אלגוריתם לניקוי תמונה מרעש - denoiser - ממש כפי שתואר קודם כדי לפתור את אותה בעיה של הזיית תמונות? התשובה לשאלה זו חיובית, והיא שמביאה אותנו למודלי דיפוזיה שהם הטובים ביותר כיום למשימה זו. שאלה זו הועלתה לפני כחמש שנים, ובשתי קבוצות מחקר, האחת בסטנפורד בידי ארמון (Ermon) ותלמידיו, והשנייה באוניברסיטת ניו יורק בידי סימונצ'לי (Simoncelli) וקבוצת המחקר שלו, הוצעו תשובות דומות וחיוביות - אכן, כן, ניתן להסב שיטת ניקוי רעש לתהליך חישובי שיוליד תמונות יש מאין. התהליך המוצע מתחיל בתמונת רעש אקראית לחלוטין ללא כל ערך ויזואלי, ואז

תמונות רועשות אינו אלא חוק הפילוג המקורי $P(x)$ כשהוא מטושטש בגאוסיאן שרוחבו נקבע לפי עוצמת הרעש בתמונות. זה אומר שחוק פילוג זה כולל זנבות רחבים שנפרסים לכל עבר במרחב, ואלה משרים כוחות כיווניים שיכוונו בייעילות את התמונה המתעדכנת. וכך, בתהליך מתוקן שנקרא נוצרת כשהיא רועשת מאוד ובהדרגה מתחדדת ומתנקה. הדרגה זו של עוצמת הרעש יוצרת שרשרת של חוקי פילוג של תמונות עם תהליך דיפוזיבי ביניהן, ומכאן השם שנבחר - מודלי דיפוזיה.

זוה עובד! ALD, ואלגוריתמי דיפוזיה אחרים הדומים לו מאוד, מסוגלים לייצר תמונות יש מאין באיכות מרהיבה שעולה על כל השיטות האחרות שהוזכרו קודם. כיום זו הטכניקה המובילה והעיקרית למשימה זו בתחום הרחב של בינה מלאכותית יוצרת, ומכאן הפופולריות הרבה שלה. במאמר מוסגר נזכיר כי בשינוי עדין לתהליך שתואר ניתן לייצר תמונות שתתאמנה לתיאור מילולי הנתון כמשפט (prompt). תהליך זה, הנקרא "טקסט מנחה" (text-to-image), מרתק בתוצריו, אך אנו לא נתעמק בו כאן. איור 3 ממחיש אפשרות זו באמצעות שלוש דוגמאות שנוצרו בשיטה שנקראת imagen.

בשימוש ב-denoiser, לפי הנוסחה הזאת:

$$\nabla \log P(x_k) = x_k - D(x_k)$$

כאשר $D(x_k)$ הוא ה-denoiser הפועל על x_k . וכך, בסיכומו של דבר התקבל תהליך חישובי שכל מרכיביו פשוטים ומוכרים, כשאחד מהם הוא הפעלת אלגוריתם לניקוי תמונות, ומובטח ששרשרת פעולות זו תספק דגימה ראויה מחוק הפילוג $P(x)$.

התהליך שהצגנו, שנקרא "דינמיקת לנג'ווין" (Langevin dynamic), יכול להצליח, אך הוא סובל מפגם מהותי שבעטיו יידרשו מאות אלפי צעדים כדי להפיק ממנו דגימות ראויות. כאשר תמונה היא בגודל של 1,000 על 1,000 פיקסלים והיא צבעונית, שלושה מיליון מספרים קובעים את תוכנה. כשאנו פועלים במרחב הרב-ממדי הזה, ואנו רחוקים מהיריעה שעליה נמצאות תמונות איכותיות, אין כל כוח שימשוך את עדכוני התמונה המבוצעים בתהליך הזה לעבר הכיוון הנכון, וכך נימצא משוטטים במדבריות צחיחים ללא התנסות לתמונה איכותית. הפתרון לזה הוא טכניקה מוכרת היטב בפיזיקה ובכימיה - annealing - שמשמעותה היא שבמקום לדגום מחוק הפילוג של תמונות נקיות נדגום תמונות רועשות. למרבה הפלא (והמזל!), חוק הפילוג של



A cute sloth holding a small treasure chest. A bright golden glow is coming from the chest



A cute corgi lives in a house made out of sushi



Teddy bears in a swimming pool

איור 3. ייצור תמונות בעזרת מודלי דיפוזיה באמצעות הזנת טקסט מנחה (text-to-image)

זוהי פונקציית ההסתברות של כלל התמונות האפשריות בהסבירן את המדידה y . לכן תמונות x שתהיינה בעלות הסתברות מותנית גבוהה הן שחזורים ראויים וטובים למשימת ההיפוך. וכך דגימה מחוק הפילוג המותנה $P(x|y)$ תוכל להפיק שחזורים שונים, כשכל אחד מהם הוא פתרון ראוי ותחרותי למשימת תיקון התמונה שלפנינו. כך נפתחו אפשרויות חדשות בזירה זו של בעיות היפוך, שמעולם לא נבחנו בעבר, ונוצר פרק חדש ועדכני בזירה זו.

כיצד דוגמים מחוק הפילוג המותנה? על נושא זה לא נרחיב במאמר זה, אך נאמר שהדבר מורכב ומחייב קירובים שונים. עשרות מאמרים שתקפו בעיה זו נכתבו בשלוש השנים האחרונות בקבוצות מחקר מובילות בתעשייה ובאקדמיה. קבוצת המחקר שלנו בטכניון – מכון טכנולוגי לישראל סיפקה תרומות נכבדות למגרש משחקים זה.

באיור 4 מומחשת היכולת לפתור בעיות היפוך בעזרת מודלי דיפוזיה עבור בעיית סופר-רזולוציה. תוצאות אלו לקוחות מעבודתם של בהג'ת קאוואר וגרישה וקסמן, דוקטורנטים בקבוצת המחקר שלנו בטכניון.

בדוגמה זו אנו מתחילים בארבע תמונות נתונות (הימניות ביותר באיור) אשר נפגמו בגלל הורדת רזולוציה והוספת רעש – אלה נתונות בתמונות השניות משמאל. המשימה היא לשחזר את תמונות המקור בהינתן המידע המקולקל, ובאיור זה מוצעות שמונה דגימות מחוק הפילוג המותנה לכל אחת מארבע תמונות מבחן אלה. כפי שניתן לראות, התוצאות הן בעלות איכות ויזואלית גבוהה, וכל אחת מספקת הסבר ראוי למידע שבו התחלנו. עוד דבר שניכר הוא שבתוך כל מקבץ של שמונה דגימות יש הבדלים עדינים ביניהן, דבר שמעיד על אי-הוודאות בפתרון הבעיה שמטופלת. בשתי העמודות הימניות של האיור רואים את

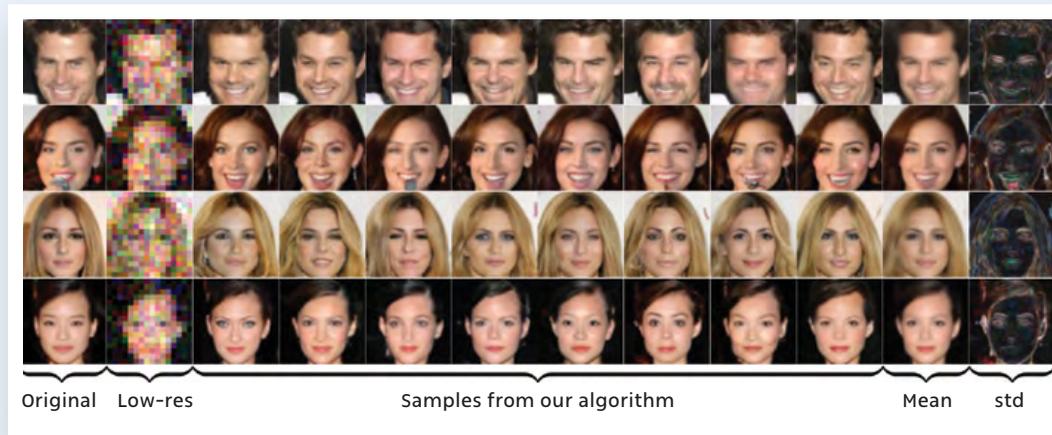
בהמשכו של מאמר זה נראה כיצד ניתן לרתום מודלי דיפוזיה לשם הצעת פתרונות חדשניים ואיכותיים לבעיות מעשיות בעיבוד תמונה. נתחיל בטיפול בבעיות היפוך ומשם נמשיך לנושא של דחיסת תמונות.

מודלי דיפוזיה לבעיות היפוך

פרק מרכזי בתחום של עיבוד תמונות עוסק בשחזור תמונות שעברו קלקולים שונים. במסגרת זו תמונה מקורית x עוברת קלקול באמצעות אופרטור H ולאחר מכן גם מזהמת בהוספת רעש n . המדידה המקולקלת היא $y = Hx + n$, ומשימתנו היא שחזור x מתוך y , דהיינו היפוך הקלקול. מכאן שמן הכולל של בעיות אלו – "בעיות היפוך". לשם דוגמה, בעיית ניקוי רעש שבה עסקנו למעלה היא מקרה פרטי של משפחה זו, ובה האופרטור H הוא מערכת יחידה שאינה משנה דבר. בעיות אחרות במשפחה זו הן שחזור תמונה מטשטוש או מאובדן רזולוציה, השלמת חורים בתמונות, שחזור צבעים לתמונת שחור-לבן, שחזור תמונה מהיטליה כבטומוגרפיה חישובית ועוד.

ממש כפי שהוזכר בהקשר של ניקוי תמונה מרעש, גם באשר לבעיות היפוך מהסוג הזה ולדרכים לפתרון קיים שפע עצום של עבודות שדנות בהן, ואלו פורסמו בעשרות אלפי מאמרים לכל אורך ארבעת העשורים האחרונים, וגם בזירה זו שיטות שחזור מבוססות AI שפותחו בעשור האחרון לקחו את הבכורה.

עם פריצתם של מודלי הדיפוזיה לחיינו לפני שנים ספורות עבר פרק זה של בעיות היפוך מהפכה מטלטלת. משהתברר שמודלי דיפוזיה יכולים להפיק דגימות איכותיות מחוק פילוג מורכב למדי כמו זה של תמונות טבעיות $P(x)$, נפתחה אפשרות חדשה של דגימה מחוק הפילוג המותנה $P(x|y)$. מהו חוק פילוג מותנה זה?



איור 4. פתרון בעיית סופר-רזולוציה (הגדלת תמונה נתונה) באמצעות מודלי דיפוזיה

דחיסת תמונות גם היא משימה קלאסית בעיבוד תמונות, ורבים טובים עסקו בה. בניגוד לסוגיות האחרות שהזכרנו בעיבוד תמונות (ניקוי רעש ובעיות היפוך), לסוגיה זו גם ערך מסחרי עצום, ובשל כך קבעו חברות בעלות עניין תקינה שהסדירה שיטות אחידות לדחיסה. כך למשל נוצרה שיטת ה-JPEG ב-1992 כסטנדרט לדחיסה שפועל עד היום בכל המצלמות הדיגיטליות, במצלמות הטלפונים הסלולריים, באתרי אינטרנט ועוד. בעזרת אלגוריתם JPEG ניתן לדחוס תמונה בשיעור של פי עשרה ואפילו פי עשרים מנפח המקור של התמונה תוך כדי שימור יעיל (אך לא מושלם!) של הפרטים בה. אסטרטגיה זו קרויה "דחיסה מאבדת נתונים" (lossy compression) כיוון שהיא מתירה לעצמה שינויים עדינים בתמונה לטובת משימת הדחיסה. כל שיטת דחיסה כוללת אלגוריתם לדחיסה (נקרא לו גם ה"משדר") שממיר תמונה למידע דחוס, ואלגוריתם תואם לו לפריסה (שייקרא ה"מקלט") שממיר חזרה מידע דחוס לתמונת התוצאה, שכאמור עשויה להיות מעט שונה מתמונת המקור.

יש שפע של שיטות דחיסה lossy שנוצרו לאחר JPEG ושעולות עליה בביצועיהן. ◀

התמונה הממוצעת של שמונה הדגימות, ולימינה את סטיית התקן בכל מקום בתמונה - לשתי אלה תפקיד חשוב באומדן של אי-הוודאות בפתרונות המתקבלים, אך לא נרחיב על כך כאן.

לסיכום, אם נתונות לנו מדידות y שמתייחסות לתמונה מקורית לא נודעת x , כשהקשר ביניהן נתון במשוואה $y = Hx + n$, אזי דגימה מחוק הפילוג המותנה $P(x|y)$ היא משימה בת השגה. כעת נראה כיצד יכולת זו בדיוק הופכת לנכס במשימה אחרת - דחיסת תמונות.

מודלי דיפוזיה לדחיסת תמונה

כפי שהזכרנו קודם, תמונה צבעונית בגודל של 1,000 על 1,000 פיקסלים נשמרת כשלושה מיליון מספרים, כלומר בקובץ שגודלו יהיה כשלושה מגבייט. האם זה מחויב המציאות? האם באמת כמות האינפורמציה בתמונה כזאת היא כה רבה? מסתבר שלא! יש יתירות רבה בתמונות, שנובעת בין היתר מהעובדה שערכים של פיקסלים סמוכים נוטים להיות דומים. משמעות הדבר היא שתמונה ניתנת לשמירה כקובץ הרבה יותר קטן, והדבר פותח פתח לדיון בשיטות לדחיסת תמונות.

המידע על תמונת המקור הוא המועט ביותר. כיוון זה ייבחר כשורה הבאה במטריצה H , כדי לספק מידע מרבי על התמונה המקורית בהינתן המדידות עד כה.

4. אלגוריתם הדחיסה ייקח את השורה החדשה הזאת ויחשב את האיבר ה-11 בווקטור באמצעות הטלת התמונה המקורית על כיוון זה ויעביר מידע זה למקלט לאחר העגלה כלשהי לשם תיאור תמציתי.

5. החישובים שנעשו בצעדים השני והשלישי במשדר יכולים להיעשות באותה מידה בדיוק במקלט, כדי שהשורה החדשה של H תיבנה גם שם ללא כל צורך בהעברתה מהמשדר.

6. וכך הגענו למצב שגם במשדר וגם במקלט אנו מחזיקים כעת מידע עדכני על H ועל y , ואנו ערוכים לשלב הבא שיחזור לצעד שתיים וימשיך בצבירת שורות ומדידות.

הינה כמה מאפיינים חשובים של התהליך שתואר:

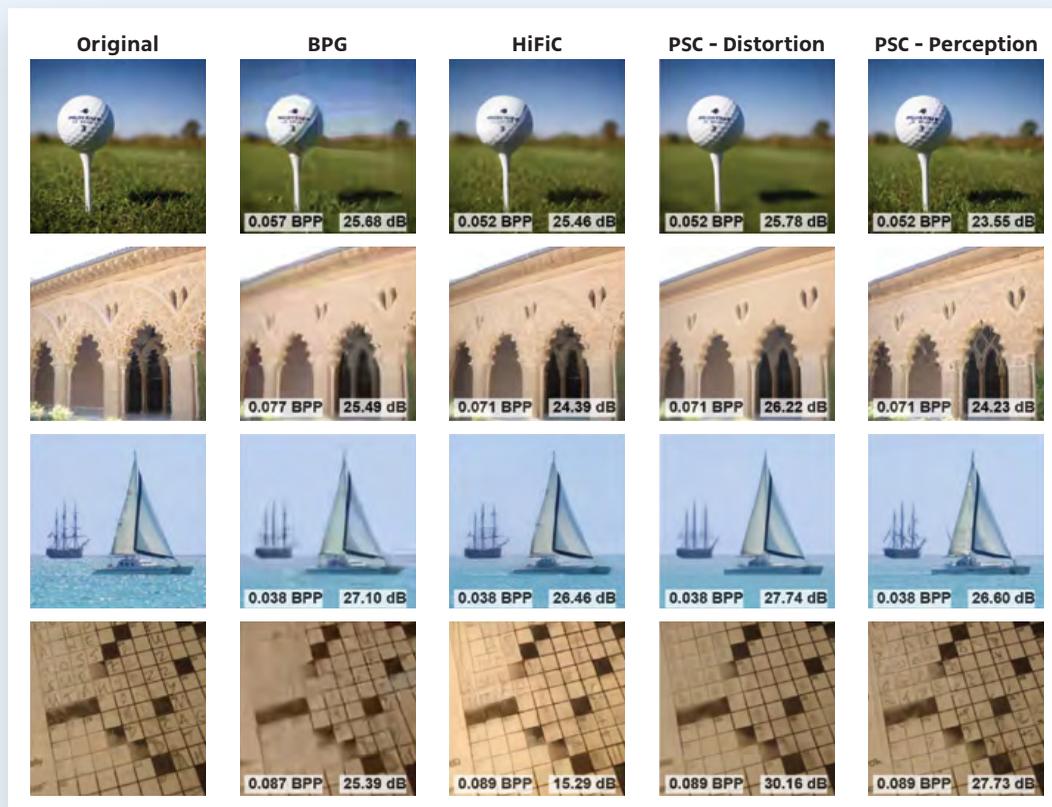
- תהליך זה צובר בהדרגה מידע דחוס שהולך ומייצג את התמונה המקורית טוב יותר ויותר. טכניקה זו נקראת "דחיסה פרוגרסיבית", כי ניתן לעוצרה בכל רגע נתון ולבחון את התמונה המתקבלת, כלומר יחס הדחיסה דינמי ונשלט בקלות.
- המטריצה H נבנית באופן ייחודי ובהתאמה מרבית לכל תמונה – מאפיין ייחודי מאוד שלא הוצעו כמוהו בספרות מעולם. לשם השוואה, שיטת ה-JPEG מפעילה פעולה H קבועה ואחידה לכל התמונות, ומכאן יעילותה המוגבלת.
- הפחתת אי-הוודאות בכל שלב יוצרת סדרה של חוקי פילוג מתונים שמתכווצים במהירות ומתכנסים לתמונה המקורית תוך כדי צבירה קצרה של מדידות – דבר המעיד על יעילות השיטה לפי המדדים של קצב ועיוות.
- ודבר אחרון – בהישענות על מודלי דיפוזיה התמונות הדגומות שמתקבלות הן בעלות איכות ויזואלית גבוהה בכל שלב בתהליך הדחיסה, מה שמבטיח שהמוצא של התהליך זה יפיק תמונות נעימות לצפייה אפילו אם יחס הדחיסה גבוה מאוד.

בשנים האחרונות הוצעו גם שיטות מבוססות AI למשימה הדחיסה, ואלו הן הטובות ביותר המוכרות כיום. ההשוואה בין השיטות השונות נעשית על סמך שלושה מדדים: כמות הביטים הנדרשת לאחסון התמונה (rate), הסטייה בין התמונה הדחוסה-פרוסה לזו המקורית (distortion) והאיכות הויזואלית של התמונה הדחוסה-פרוסה (perception). גם מודלי דיפוזיה מונפו לטובת משימת הדחיסה, ומאמרים ספורים מהשנה האחרונה הציעו סכמות דחיסה חדשניות ויעילות מאוד ברוח זו.

בעבודה משותפת עם פרופ' תומר מיכאלי מהטכניון ועם הדוקטורנט נועם אלתי פיתחנו בשנה האחרונה (2024) שיטת דחיסה חדשה המבוססת על מודלי דיפוזיה. יופייה של שיטה זו הוא ביכולתה להישען על כל מודל דיפוזיה קיים ללא כל אימון נוסף ולמנפו לטובת משימת הדחיסה. ביסודותיה של שיטת דחיסה זו נמצא שימוש בדוגם מחוק הפילוג המותנה $P(x|y)$ שהוזכר קודם, כאשר $y = Hx + n$. הרעיון הבסיסי הוא לצבור את המדידות y וכנגדן את שורות המטריצה H בהדרגה ובאופן "חמדני" שתכליתו הפחתת אי-הוודאות של חוק הפילוג המותנה בכל שלב בצורה מיטבית.

בפירוט רב יותר התהליך נראה כך:

1. נניח כי בדרך כלשהי צברנו עשר שורות למטריצה H וכנגדן יצרנו עשרה ערכים בווקטור המדידות y . נניח שכל אלה ידועים הן במשדר והן במקלט.
2. אלגוריתם הדחיסה יכול לשאול – מה יכולה להיות התמונה שתסביר את y ? המענה על שאלה זו יינתן באמצעות דגימה של מאה תמונות אפשריות מחוק הפילוג המותנה $P(x|y)$.
3. אם נבחן את סט התמונות הדגומות, נוכל לגלות את הציר שבו פיזורן הוא המרבי (זהו חישוב שמוכר בשם PCA) – זהו הכיוון שבו



איור 5. תוצאות דחיסה-פריסה השוואתיות בין השיטה המוצעת ל-HiFiC

ביחידות dB כשגבוה יותר זה טוב יותר. המדד השלישי של איכות ויזואלית אינו מדווח מספרית – כאן בחינה של הצופה תקבע איזו תמונה יפה יותר לעין האנושית.

לסיכום

מודלי דיפוזיה מובילים מהפך של ממש בתחומים שונים במדע, ובכללם עיבוד תמונות. מאמר קצר זה עסק במהותם של מודלים אלה כדוגמים מחוקי פילוג מורכבים של מידע עשיר. קבוצת המחקר שלנו בטכניון עוסקת באינטנסיביות בנושא זה ורותמת מודלי דיפוזיה בדרכים חדשניות ויעילות למשימות של פתרון בעיות היפוך, של ייצור תמונות יעיל יש מאין, של דחיסת תמונות ועוד.

איור 5 ממחיש את תוצאותיה של שיטה זו לעומת שתי שיטות דחיסה תחרותיות מאוד: BPG, שהיא שיטה קלאסית איכותית שנוצרה בשנת 2014, ו-HiFiC, שמבוססת על למידה עמוקה שנוצרה בשנת 2020 בגוגל. בשיטתנו אנו יכולים להתבונן בכל שלב בתוצאת דגימה יחידה, ואז לתקבל איכות ויזואלית גבוהה (עמודה ימנית), או למצע את כלל הדגימות כדי לקבל שגיאה קטנה (עמודה שנייה מימין). יחסי הדחיסה המוצגים כאן הם 461:1 (תמונה עליונה), 338:1 (תמונה שנייה), 631:1 (תמונה שלישית) ו-269:1 (תמונה תחתונה). יחסי דחיסה אלה מתייחסים ישירות לערכים המדווחים בכל תוצאה משמאל (BPP פירושו Bits-Per-Pixel). המספרים מימין מתארים את הסטייה מהתמונה המקורית