⚡️🚀המאמר היומי של מייק 01.08.24: ⚡️🚀

IMPROVED TECHNIQUES FOR TRAINING CONSISTENCY MODELS

היום סוקרים קצרות עוד מאמר בנושא קרוב לליבי - המשך של המאמר שסקרנו לפני בערך שבוע הנקרא ״consistency models". אם אתם זוכרים מודל קונסיסטנטי הוא שייך למשפחת מודלי דיפוזיה (כלומר הוא מתואר על ידי משוואת הדיפוזיה). אחת הבעיות של מודלי דיפוזיה קלאסיים (כמו DDPM) היא איטיות של גנרוט דאטה. הדאטה נוצר באמצעות תהליך denoising הדרגתי - מתחילים עם רעש גאוסי ומסירים אותו לאט לאט.

כדי להתמודד עם הבעיה הזו הוצעו כמה שיטות ואחת מהן DDIM סקרנו אתמול. השנייה היא מודלים קונסיסטנטיים(CM) שניתן להגדיר אותם כי מודל שונה (אך דומה) ממודל דיפוזיה קלאסי. בעיקרון ב-CM אנו מאמנים מודל להסיר רעש מכל פיסת דאטה מורעש באיטרציה t כך שהתוצאה תמיד תהיה פיסת הדאטה מקורית (ללא רעש). מכאן בא שם של המודל: קונסיסטנטי.

איך זה למעשה נעשה? יש שתי דרכים עיקריות לאמן CM. דרך אחת מסתמכת על מודל המשערך את מה שנקרא score function שהיא לוגריתם של פונקציית ההסתברות של הדאטה המורעש באיטרציה t. ידוע כי תהליך גנרוט של דאטה במודלי דיפוזיה (כלומר denoising) מתואר על ידי משוואת זרימה (דיפרנציאלית) שמתאר את המסלול של דאטה מהרעש עד הדאטה הנקי. ו- score function מופיע במשוואת זרימה זו. אז השיטה הראשונה ממזערת את המרחק בין שערוך x\_0 (הדאטה הנקי) מ x\_t+1 לבין שערוך של x\_0 מ- x\_t כאשר x\_t מחושב ממשוואת הזרימה (איטרציה אחת של אוילר של משוואת הזרימה שכבר הזכרנו). ו

דרך אגב שערוך של score function די שקול לשערוך של הרעש הנוסף (לדאטה) במודלי הדיפוזיה הסטנדרטיים. הדרך השנייה ״ליצור״ את x\_t+1 היא לשערך את x\_0 מ-x\_t+1 ולהוסיף רעש (כמו באיטרציה t).

המאמר המקורי על CM השתמש במרחק הנקרא LPIPS המודל דמיון סמנטי בין התמונות (דרך השוואה של אקטיבציות של מודלים מאומנים על דאטהסטים ענקיים של תמונות). המאמר המקורי גם התשמש ב-EMA (החלקה מעריכית) של משקלי המודל בתור המודל עבור x\_t. יש כמובן חשיבות לבחירת השונות של האיטרציות.

אז המאמר שסוקרים היום משפר את תהליך האימון. השינוי הראשון הוא משקול של המרחקים כפונקציה של איטרציה t; ככל שמתקרבים ל 0 המשקול עולה. דבר שני זה שינוי של פונקציית מרחק מ-LPIPS לפונקציית הובר (Huber) עם טוויסט קטן. הדבר האחרון והמעניין הוא ביטול של EMA ל-x\_t כלומר ההשוואה מתבצעת בין שני מודלים ״טהורים״ ל- x\_t ו- x\_t+1. גם הייפר פרמטרים אחרים עבור שינוי למשל השוניות של הרעש באיטרציות.

בקיצור יש לנו כאן שכלול מעניין של CM - בקרוב אסקור עוד מאמרים על זה…

https://arxiv.org/abs/2310.14189