# Review 3: Are Deep Neural Architectures Losing Information? Invertibility Is Indispensable?

**Paper: https://arxiv.org/abs/2009.03173v2**

המאמר משך את עיניי כי שמו בתרגום לעברית "האם רשתות נוירונים מאבדים מידע? האם הפיכות נחוצה!"

תחום המאמר: זרימת המידע ברשתות נוירונים

הדקמה: קודם כל אנו יודעים היטב שהתשובה על השאלה הראשונה הינה חיובית מהעבודות של פרופ' תשבי שהראה בצורה רגורוזית שיש אובדן מידע בכל שכבה (הרי בסופו של דבר במרבית המקרים אנו דוחסים את התמונה לוקטור קצר פי 1000 מכמות הפיקסלים בתמונה).

תקציר המאמר על קצה המזלג: המאמר שואל את השאלה המתבקשת האם ניתן לבנות רשת הממזערת את אובדן המידע הזה. למשל רשת המסוג הזה עשויה להיות טובה למשימות עיבוד תמונה low-level כמו שחזור התמונה או colorization.

תקציר: קודם כל הם מגדירים את אובדן המידע המידע הדדי בין התמונה והייצוג הלטנטי שלה. אחר כך הם מוכיחים (למרות שזה די ברור) כי ככל שההסתברות לקבל תמונה X בהינתן הייצוג הלטנטי z קרוב ל 1 אז המידע הדדי הולך עולה. זה הוביל אותם למסקנה כי בשביל לממש את המטרה הזו הם צריכים לקחת את השיטה לגינרוט שנקראת "זרימות מנורמלות"(normalizing flows - NF). העיקרון של NF הינו די פשוט. בוחרים התפלגות לטנטית כלשהי בדרך כלל גאוסי או יוניפורמי. ואז מנסים לבנות פונקציה הפיכה (חד-חד ערכית) ממרחב התמונות למרחב הלטנטי. אז גינרוט התמונות יוצא די פשוט - דוגמים את z ומפעילים עליה את ההופכית של הפונקציה הזו. נשמע לא מסובך אך הבנייה של פונקציות כאלו זו משימה מאוד לא טריוויאלית. אחת הסיבות לכך זה ההופכית של היעקוביאן של הטרנספורמציה (פונקציה) המופיעה בביטוי של הפונקציה ההופכית וכאשר גודל המטריצה הוא עצום החישוב של מטריצה ההפוכה הינו מאוד כבד חישובית. אז בונים אותה בצורה autoregressive היעקוביאן ייצא מטריצה משולשת עליונה הקלה להיפוך.

בקיצור המאמר לוקח את הארכיטקטורה של NF עם יכולת גינרוט הכי גבוהה הנקראת GLOW (מבית היוצר של הממציא של VAE דרך אגב) משנים טיפה את הארכיטקטורה (לדעתי זה לא כזה קריטי למרות שבמאמר נטען אחרת). השינוי המהותי היחיד שהם עושים יחסית ל GLOW זה הוספה של decoder במקום להשתמש בהופכית של ה encoder. פונקציית לוס: L1

הם מוכיחים שעם הארכיטקטורה הזו הם מצליחים לשפר את הביצועים של שורה של משימות עיבוד תמונה low-level כמו שחזור תמונה, colorization, שחזור תמונות דחוסות

הישגי מאמר: שחזור תמונה, colorization, שחזור תמונות דחוסות.

הערות על ניסויים: אני סבור שהניסויים שלהם מאוד מוטים לטובתם ובכלל ההשוואות שהם עשו לא הגיוניות. בגדול הם השוו את הרשת שלהם לרשתות ניקוי רעש בעיבוד תמונה אבל יש שתי נקודות בעייתיות בהשווא הזו.

1. הרשתות שלהם הם השוו הן לא עדכניות. אחת מהן מ- 2016 והשניה מ- 2018 ומאז היו המון שיפורים בתחום.

2. נראה שהם אימנו את הרשת שלהם לדטה סט ספציפי, למשל סלב, ואז בדקו עליו בעוד שהרשתות האחרות מאומנות על דאטה כללי של תמונות טבעיות. ברור הרי שכאשר רשת שאומנה על פרצופים בלבד מחלצת פיצ'רים הרבה יותר רלוונטיים מרשת שאומנה על תמונות טבעיות כלליות. וזה החלק הבעייתי באמת

נ.ב. לא השתכנעתי מהחלק הזה בעבודה, בעיקר כאשר יש תחום שלם של היפוך של generative models שהולך ותופס תאוצה וניתן לעשות השוואות יותר רלוונטיות.