

בונוס תרגיל 5

עמית בן ארצי 207448390

20 בינואר 2021

1 מה בעצם קורה כאן?

1.1 Overview

כדי לנקות את התמונה, אנחנו:

1. נראה ביטוי הסתברותי שהמינימום שלו הוא תמונה נקייה
2. נראה שכדי למצוא את המינימום נוכל להשתמש במרחב פרמטרי אחר (כמו שראינו למשל באלגוריתם של hough)
3. נראה כיצד נוכל להשתמש ברשתות נוירונים בתור מרחב פרמטרי

1.2 קצת (הרבה) הסתברות

נסמן ב- x את התמונה הנקייה, וב- \hat{x} את התמונה עם הרעש. אנחנו רוצים למצוא את x^* - תמונה שמנוקה מרעש. מה זה בעצם x^* ? זה הביטוי הבא:

$$x^* = \arg \max_x p(x | \hat{x})$$

ניזכר קצת בהסתברות ונבין את המשמעות של הביטוי הזה.

- הביטוי $p(x | \hat{x})$ הוא ההסתברות לקבל את x בהינתן \hat{x}
- אנחנו רוצים למקסם את הביטוי הזה. אבל לא אכפת לנו מהמקסימום עצמו - אלה מהאיבר שיחזיר לנו את המקסימום, כלומר $\arg \max$.

כעת נשתמש בבייס ונקבל:

$$p(x | \hat{x}) = \frac{p(\hat{x} | x) \cdot p(x)}{p(\hat{x})} \propto p(\hat{x} | x) \cdot p(x)$$

כאשר הסימן \propto מייצג יחס ישר, כלומר הם פשוט שקולים מבחינתנו.

דוגמה

סתם לדוגמה - נניח שיש לנו תמונה x , והוספנו לה רעש $\varepsilon \sim N(0, \sigma^2)$ (כלומר הוספנו לה רעש בהתפלגות נורמלית בין 0 ל- σ^2), וכך יצרנו את התמונה המורעשת $\hat{x} = x + \varepsilon$. אז אנחנו נקבל $p(\hat{x} | x) = N(\hat{x}; x, \sigma^2)$. כלומר את ההסתברות הזו קל לנו לחשב.

1.3 מה זה prior?

בביטוי שעכשיו ראינו - prior זה $p(x)$. אינטואיטיבית prior זה משהו שאנחנו "יודעים" על העולם ועל איך שתמונות נראות. למשל, מניסיון חי, ברוב המקרים לבני אדם לרוב שתי עיניים. לרוב אנחנו נאמן רשתות נוירונים על המון תמונות כדי שיהיה להם מושג מה זה תמונה. אבל כאן אנחנו נראה בהמשך שאנחנו נוכל לאמן רשת נוירונים על תמונה אחת בלבד ועדיין לקבל תוצאות יפות מאוד. מה שמוביל אותנו לשאלה הבאה:

1.4 האם אנחנו יכולים בלי prior?

לא אם אנחנו רוצים לקבל תשובה משמעותית. מדוע? כי מתקיים:

$$\begin{aligned} x^* &= \arg \max_x (x | \hat{x}) = \arg \max_x (\hat{x} | x) p(x) = \\ &= \arg \max_x (x | \hat{x}) = \arg \max_x N(\hat{x}; x, \sigma^2) = \hat{x} \end{aligned}$$

ואז בעצם אנחנו נקבל בחזרה את התמונה המורעשת, אז לא עשינו כאן כלום בעצם. אבל כפי שעוד מעט נראה - אנחנו יכולים להשתמש במבנה של רשת נוירונים כprior בפני עצמו!!

1.5 שימוש במרחב הפרמטרי

1.5.1 שינוי קל בסימונים

מכאן ואילך אנחנו נשנה את הסימונים ונקבל:

$$x^* = \arg \max_x p(\hat{x} | x) p(x) = \arg \min_x -\log p(\hat{x} | x) - \log p(x) = \arg \min_x E(x; \hat{x}) + R(x)$$

השינוי בסימונים יאפשר לנו שימוש במרחב הפרמטרי, ולכן אנחנו עושים אותו. **חשוב חשוב חשוב! הפונקציה E אינה פונקציית התוחלת!** אנחנו נראה בהמשך שלמטרות שונות (הפחתת רעש, השלמה של תמונות, Super Resolution ועוד) יש פונקציות שונות שמתאימות לנו.

1.5.2 מעבר למרחב הפרמטרי

בעצם מה שאנחנו נעשה זה לחפש את

$$x^* = \arg \min_x E(x; \hat{x}) + R(x)$$

במרחב פרמטרי (דוגמה שראינו לשימוש במרחב פרמטרי זה האלגוריתם של hough למשל). כלומר נחפש במקום את הביטוי:

$$x^* = \arg \min_{\theta} E(g(\theta); \hat{x}) + R(g(\theta))$$

כאשר g היא פונקציה שממפה לנו את מרחב התמונות למרחב פרמטרי אחר. במקום לחפש את ה-argmin עבור תמונות x נחפש את ה-argmin עבור מרחב פרמטרי עם θ .

1.5.3 כמה נקודות ממש חשובות לגבי המעבר הזה

1. מן הסתם חשוב לנו מאוד ש- g תהיה הפיכה, כלומר ח"ע ועל, כדי ששתי הבעיות יהיו באמת שקולות.
2. מכאן ואילך נתייחס לפונקציה g בתור הייפר פרמטר ונשחק איתו.
3. נשים לב לכך שהפונקציה g היא בעצם prior בעצמה!

1.6 שימוש ברשתות נוירונים בתור מרחב פרמטרי

אנחנו למעשה נרצה שהמשקולות של רשת הנוירונים יהיו המרחב הפרמטרי שלנו. כיצד נעשה זאת?

1. נסמן ב- z תמונה מסויימצת של רעש רנדומלי כלשהו
2. בכל איטרציה נכניס את z לרשת נוירונים, ו"נגיד" לרשת נוירונים שאנחנו רוצים לקבל משהו שקרוב לתמונה המורעשת
3. למעשה הפעולה שרשת הנוירונים תעשה על הרעש היא פונקציה מהמרחב הפרמטרי של משקולות רשתות הנוירונים אל מרחב התמונות

1.6.1 איזה פונקציות E יש לנו?

להפחתת רעש:

$$E(x; \hat{x}) = \|x - \hat{x}\|^2$$

נקודה חשובה! במקרה הזה חייבים לעצור יחסית מוקדם - כי אחרת נקבל overfitting לתמונה המורעשת. Inpainting-:

$$E(x; \hat{x}) = \|(x - \hat{x}) \cdot m\|^2$$

לסופר רזולוציה:

$$E(x; \hat{x}) = \|\text{downsampled}(x) - \hat{x}\|^2$$

2 האלגוריתם עצמו

אלגוריתם 1 אלגוריתם לשחזור תמונות בעזרת Deep Prior

1. נבחר רשת נוירונים כלשהי
2. נאתחל את z להיות תמונת רעש כלשהי
3. נפתור באופן איטרטיבי את המשוואה

$$\arg \min_{\theta} E(f_{\theta}(z); \hat{x})$$

באמצעות שיטת gradient descent כלשהו. (כלומר - זה יהיה ה-loss שלנו)

4. נפעיל את רשת הנוירונים עם הפרמטרים שלה על z
-

3 כמה הערות על המימוש שלי

1. מכיוון שההפרש בין ערכי PSNR לא מאוד גבוה בין האריטקטורות השונות - בחרתי לממש בעזרת ResNet, כאשר הפכתי את Relu ל-LeakyRelu כפי שהוצע במאמר.
2. הפרמטרים שבחרתי הם שילוב בין הפרמטרים שבמאמר לכאלו מניסויי וטעייה; סה"כ במאמר הם השתמשו על תמונות הרבה יותר גדולות מאשר $64 \times 64 \times 1$.

3. אם תשימו לב יש שורה בקוד שלי שאפשר לעשות לה uncommit ולראות כמה משתנה הפלט בכל כמה איטרציות - לדעתי זה ממחיש עד כמה קשה לבחור הייפרפרמטרים עבור תמונה מאוד קטנה. הייפר פרמטרים שונים עבדו טוב יותר לתמונות שונות וההבדלים היו מאוד משמעותיים.

4. ההבנתי מהפורום האישי שבדיקת הבונוס היא ידנית; השארתי כמה טסטים קטנים שכתבתי אם תרצו להשתמש בהם 😊