

(יש מצגת)

מבנה רשת נייטרונים היא גרם מכון, שמייצג רצף של פעולות על הקלט, וקלטן אלה היה מסס על פונקציה loss כלשהי. המטרה: למצוא פונקציות שיהיו טובים יותר ל loss.

המרכיב המרכזי של ConvNets: שכבה קונבולוציה

שכבה קונבולוציה מורכבת מאוסף פילטרים (שגומים נחמד), אלה "מסעלים" את הקלט.

הפעולה שיש היא פשוט ביצוע קונבולוציה של הפילטר עם הקלט.

קריית activation map - כבר נקראו לפילטר של קונבולוציה עם פילטר קודם.

קצת על גדלים: כשאנחנו מפעילים פילטר על קלט, אנחנו צריכים

לסתור מה קודם הפילטר, כמות צעדים (Stride),

ואצבע חיפוש (padding) נעשה לקלט - כל אלה יסבירו את גודל הפלט.

input shape	7x7	7x7	7x7	למשל:
filter shape	3x3	3x3	3x3	
stride	1	2	3	
output shape	5x5	3x3	error	

קונבולוציה וצד: אפשר לחשוב על זה כמו על חורף מים

בזמן מאלף. מהי נקודה לעשות את זה? כלומר: אחרים

אלה מס' הפונקציות ומס' החישובים - נקטין ק את מס' ה-channels.

מרכיב חשוב נוסף: שכבה Pooling

(היה) מושגים השלכה max pooling כדי לקטין את הקלט

ואיזו אינוויאנטיה אנחנו מקבלים.

שימוש ב-pooling של 3x3 עם Stride של 2, יגרום לחפיפה

ב-pooling, ויכול למנוע aliasing.

Conv Nets: המחקר

אנליטקורט נפתרה:

* LeNet, 1998 - עוצמה בסיוף אליוט וספרה OCR

* AlexNet, 2012 - בסיוף ImageNet

* VGG, 2014 - תחילת: פירוט קטנים של קונבולוציה עם תחילת בסיוף

(הדמיון אל הפירוט נ-11111-3x3)

ומספר שפירוט קטנים זה יורד על חלופה

אימון של רשת קונבולוציה:

1. Preprocessing: (א) נורמליזציה של התמונה וסיוף הקו 1.

(ב) whitening: סיוף אקסטרמליזציה על התמונה

שינוי דוגמא אקסטרמליזציה.

2. אחר ממשקולות:

אסטרטגיה 1: נאמרו את הממשקולות ל-0. אסטרטגיה

אם לא? כי במחיל האימון, כל הממשקולות חול

ממשקולות הממשקולות יאמרו על כל התחילתים יאמרו.

אם יש ReLU כל הממשקולות יאמרו

אם יש tanh הכל יאמרו חול מהממשקולות של הממשקולות

הממשקולות ואל במחיל האימון (עצבן רק אלף)

אסטרטגיה 2: אחר תחילת מהממשקולות קדומה

אסטרטגיה 3: "Xavier": אחר תחילת מהממשקולות אחרת לעולם

אם של Stal-מה או יורד קקו אליו הממשקולות

במחיל אחר: Batch Normalization: אסטרטגיה חלופה (קדומה) אל הממשקולות של הממשקולות

הקדומה

תחילתים

3. Loss: אל ממשקולות עקרונית > תחילתים

(א) Cross-Entropy: אקסטרמליזציה - נורמליזציה מוחק בין הממשקולות

בה ממשקולות על הממשקולות של הממשקולות סיוף $\log(P(X|Y))$

Square Loss (ד)

אימון של רשתות קונבולוציות: המעבר

3. בחירה $loss$: המעבר: הסדר $P(y|x)$ CE

$$\begin{aligned} \max_{\theta} \log(P(y|x)) &= \max_{\theta} \left\{ \log \left(\frac{P(x|\hat{y}) \cdot P(\hat{y})}{\sum_y P(x|y) \cdot P(y)} \right) \right\} \\ &= \min_{\theta} \left\{ -\log \frac{P(x|\hat{y})}{\sum_y P(x|y)} \right\} = \\ &= \min_{\theta} \left\{ \log \sum_y P(x|y) - \log P(x|\hat{y}) \right\} = \\ &= \min_{\theta} \left\{ \log \sum_y \exp(o_y) - o_{\hat{y}} \right\} = \\ &= \min_{\theta} \left\{ \log \sum_y \exp(o_y - o_{\hat{y}}) \right\} = \\ &= \min_{\theta} \left\{ \log \sum_y \exp \left(\log \frac{P(x|y)}{P(x|\hat{y})} \right) \right\} \end{aligned}$$

4. SGD: (mini batch)

ניקח batch, נחשב עליו את ה-loss ונעביר קטורז הגרדיאנט.
הערה: momentum term: כיאני נוסף טרמיסטים לצדדים העדכון כך
שאם הגרדיאנט ממשיך להיות באותו כיוון מספר צדדים קטן,
נעביר אותו יותר בכיוון הזה - נעזר את העדכון שלנו.
הערה: צדדים לצדדים ולסכום את המינימום:



מקרה זה אם באנו אל מינימום
אז לא נצטרך לשמור, (לסכום את המינימום).

ליסור: nesterov

הערה: Adam: מוסף רעיון של נרמול העדכון עם הממוצעים והסטייה.
של הגרדיאנטים מהעדכון הקודמים.

5. learning rate:

בחרת ה- lr חשובה להימנעות. נדון לעצק את ה- lr חזק כפי האמון.

6. Dropout: מתוך המון, נבחר באופן רנדומי מיקומי

מה זה טופ? 1. forces redundancy

2. הפכה טופה יותר

3. חוסל את הרגש

זה יחידה יחידה עם ק- test מקום משהו monte-carlo

נחש אלו inverted dropout ככה למדן test לא נצרך לשלם כלום

אז לא אלטרנטיבות: (אם אפשר) - פשוט אולי נצטרך להסיר את כל ק

Multiplicative Gaussian Noise- : $h \rightarrow h \cdot r, r \sim N(1, \sigma^2)$

Additive Gaussian Noise- : $h \rightarrow h + r, r \sim N(0, \sigma^2)$

7. Data Augmentation: נאמא סטטיסטיקה על המידע כדי

להקטין את המידע של ילדים priors

חירת hyper-parameters:

אין נאורלן חתה פרוטקט שריר זהותם לקבוע מתוך המידע

אם בחירת האלגוריתם נכונה או לא?

* Grid Search:

ננסה את כל השילובים האפשריים של הפרמטרים ונבחר את

זה שיש לו score הכי טוב על סט איזציה

* Random Search:

מקום כל השילובים האפשריים - אבל פעם נחפש מקום אחד

מרחב החיפוש. ככה נראה (נחש) משהו זהה לטן קיצור

אין טופ של החיפוש

* Coarse to fine: נחיל חיפוש עם קצרה Random Search

ואח"כ נחפש - Grid Search באזור ממוקד שלוקח קצרה טובה

מגן דאָס אידן / מ'גלח נייע

1. איתני שם שבה נסמח אלה אשר כבר קדם ל סוקציה
אל מה פסוק מוכח ?

אמסירה וואס לעב ענדט האט שוין אסער זינגען פאר
העלע רעפארעט פונקציעלע שטעלע אטאקע אטאקע רעפארעט
זינגען אקסעסאריאל.

? פֿאַרשטאַנד פֿון אַרומגעבונג און פֿאַרשטאַנד פֿון אַרומגעבונג
(... receptive fields, channels 'on, p. 118)

[illegible]

הכינו קנייניו שצמוד אלול קופצמאל נחמן, והרש יפה צמח ח סוקר.ה
אל הקצמאל איפה הקצ? כמראה סוקר המצמח אל הקצמאל.

4. אילו SGD מנגנונים יסוגרו בעצמם את התהליך?
בסופו, SGD כן עובד וסוגר בעצמו את התהליך
פי קריסה אלמנטרית.

מס' / א' ססור.

Adversarial Examples: איך מ"צים את הרעל שלפס את מלל?

[illegible]

All-Convolution \rightarrow Stride > 1 \rightarrow pooling \rightarrow 1×1 conv \rightarrow dense

ההסבר: שכן ר"ה של pooling א"כ זה מקובל יותר.

הפרק יגלה כי inception module הוא: GoogLeNet .

המחלקה הכלכלית, משרד המבחן, 1941

הוא מס' ה-channels וככה למצוא את החילוקים.

Opinionale xisxoxoxo lapid meek: ResNet.3

