

سؤال ا

convolution: در شبکه های ماده (Convolutional Neural Network) معمولی و نیز معمولی دارای

متغیرهای داده‌های آزمون و آزمون نیز توزیع متفاوتی دارند covariate shift ② این خواسته نمود که توزیع باعث نشود در نتیجه این داده‌های جبران شوند generalization را داشته باشد.

BN با خواصی داده‌های هویتی دارد، train و test و میانگین و واریانس / تنظیم کننده این داده‌ها برای آن و سریعتر شود. اما اگر داده‌های train و test و calibrate shift، test و train و میانگین و واریانس آنها شده در طبقه train باشند ممکن است تفاوت بزرگی بین آنها باشد. در نتیجه این داده‌ها بزرگی بین آنها دارند و با اینکه میانگین و واریانس آنها متفاوت باشند، این داده‌ها را کمیشن می‌کنند و نتایجی دارند که توزیع متفاوتی دارند

لست بخواهد بود بایان را اینها کند:

$$x_{in} \xrightarrow{cos} x_{out} \xrightarrow{Br} \alpha y \left(\frac{\alpha x_{out} - \alpha k}{N_0} \right) \quad \text{for crowd}$$

plus sign of plus sign after 31

$$\alpha x_{in} \xrightarrow{\text{conv}} \alpha x_{out} \xrightarrow{\text{BN}} \gamma \left(\frac{\alpha x_{out} - \alpha \mu}{\lvert \alpha \rvert \sigma} \right) + \beta = \text{sign}(\alpha) y$$

میں خوب کرے / میرے بھائی کو

ایسا نہیں کرے جو اپنے نتیجے کا نتیجہ نہیں کرے

(4)

BN بازگشایی کردن داده های هوای train در طبقه Batch می باشد و برای اینکه هوای train را بازگشایی کنیم باید نمونه های آماده شده را در یک دسته بندی خاصی مورد استفاده قرار دهیم تا این دسته بندی داری باشد. این دسته بندی را batch می نامیم و شعاع می باشد. نوع داده های مجموعه شود و بگردایم یا نبینیم. حسابی این دسته بندی دوسری داده های مجموعه شود و بگردایم یا نبینیم. تغییرات این دسته بندی دلیل نویز ناشی از تغییرات بین داده های مجموعه شود. داده های مجموعه شود و سواد مجاز اور غایب شدن علائق مری مسدود شوند. در پروژه 6 می بینیم که

Dropout	Batch normalization	عملکرد
نحوی اغیر متعال بودن حسادی نخواهد درکهای زمزمه ایجاد نمی شود	نحوی از طریق تغییر آمارهای محض Batch ایجاد نمی شود	عملکرد
کاهشی واسطگی بین نخواهد علوکری از اورانس شدن	پایدار سازی خود گذاشته کاهشی واسطگی بین کاراول	عدهت
آخرین طبقه ای قویتر و مستقیم	آخرین طبقه ای ملحوظ و نیز پیش	قدرت آخر
واسطه بر توزیع داده ها فقط بر بعنوان تایر در تواریخ است	واسطه بر توزیع داده ها فقط بر بعنوان تایر در تواریخ است	واسطه بر داده ها

در مدل های کوچک یا داده های کم و تکیت $ReLU$ محبوب است اما بزرگ با $tanh$ یا $sigmoid$ بود.

(6)

اختصار گرادیان:

- در چیز معالسانی بدین دلیل بودن داده های ورودی مولایه گردان آسان و بعقب منتشری شود

- در پس فعالسانی هم، اگر جای $ReLU$ بروجی اعمال می شود، اثایق تواند به باره ای گردان گردد

- سه، چیز معالسانی گردان را باید بروجی کند

با این بعینه سازی:

- پیش معالسانی موجب یادگاری نظر بر بعینه سازی قدر چشمی شود؛ زیرا بروجی لایه های دیگر

- در پس معالسانی، شبکه باید بروجی نباشد؛ وزن را نیز از W تنظیم کند که بعینه بگزینی پیشتری دارد

- نتیجه، بعینه معالسانی یادگاری بعینه سازی را بعینه بگزینی بخشد

سازمانی بین آنها و نتیجه:

- آنها را در مدار آنها نیز فحیج افرادی که شوند و برقرار است تابعی مانند در چیز معالسانی

- در پس معالسانی، اگر اون آنها را بولی آنها و نتیجه مقادیر متفاوت باشند (همان درستی ببرور سازی شود) محبوب

- است عملکرد شبکه در نتیجه افت کند یا ناکسرانی را نشود

تفصیل اینها:

- چیز معالسانی: قوارطون BN یعنی $ReLU$ از الگوریتم شدید در روش علوکری می کنند

- بعینه معالسانی: ضرف یا طبقه بسی بون جایله BN می تواند افزایش و ریاضیاتی عینی و ناپایابی شبکه را دری باعث مانند

اینها سهلی عینی:

- چیز معالسانی: BN کمی که شوند توزیع ورودی مولایه یادگاری باشند و شبکه دتواند بسیاری های محبوب

محبوب شود

- بعینه معالسانی در کار توزیع این طرز لایه های توأم و اندیف ایند یادگاری هست که راهنمایی کند

- چیز معالسانی مجموعه ای اندیفی هست که در گرانها و رسمهای سه که این طرز کند

- پس معالسانی نیز در نیم صورت است اما به دلیل وجود میگذاری دیگر BN بسیاری بینشی دارد

ایجاد اولی سادگی خوبی را که کنیم که کارکرد و کارکردن مایک بعیی است (۱۰٪)

از رابطه بازگشی زیر بعدست می آید Receptive field

$$R_L = R_{L-1} + (k_L - 1) \cdot d_L \cdot \prod_{i=1}^{L-1} s_i$$

دوسرا ایں خود بہان صورت اسے کی اگر یہ موقوفت در لای ہے اس کا در تھوکریم ان موقوفت طالع اعمال صلیخ رہا اندکہ اس کا
و، *dilated receptive field* کو فوہد بود کہ *Receptive field* کا مقدار $R_d = R_c \cdot (1 + k)$ اور اپنی فوہد

لطفه غیر مازمتنی به صورت زیر است

$$R_L = 1 + \sum_{l=1}^L (k_l - 1) d_l \prod_{i=1}^{L-l} s_i$$

در این روش R_k در کرنل $K_{k,k}$ و بازست ماتریس \mathbf{A} در معادله $\mathbf{R}_k = \mathbf{A}^T \mathbf{A}$ قرار می‌گیرد.

$$R_E^{(2D)} = (R_E^{(1D)})^2$$

۲) جلی کاچن مفہوم حسابی (F_{loop}) و قیمت خوبی T ، $R^{(i)}$ ، $A^{(i)}$ کا مسئلہ بعین سانی اور جو حسیم کے آن را صورت نیو می تو ایں مقولہ ہے

$$\min_{d_{LSF}} \frac{1}{\prod_{i=1}^n s_i} \cdot k_L^2 \rightarrow \text{gerekli kernel'lerin Flops'ları, strde'lerin sayıları, kernel'lerin boyutları}$$

$$\text{s.t. } \sum_{\ell=1}^L (k_\ell - 1) \leq \prod_{j=1}^{e-1} s_i$$

که می باشد که می باشد که می باشد

یا له نکر کو معینہ یہ کو حستہ است

$$> \int |\psi^{(a)}| \beta^{(a)} > T$$

۵۸۰۵ های ذری ساز خودچی و رادیو الاههای خورخای طاھش می دهد که این هنری در طاھش همینه حسابی اینها بخشد

سوال ۹

۱) درون یابی فضی،
یک بعدی،

فرض کنید دو نقطه (x_0, y_0) و (x_1, y_1) داریم. مقدار y در نقطه x بین آنها به صورت زیر است

$$y = y_0 + \frac{(x - x_0)}{(x_1 - x_0)} \cdot (y_1 - y_0)$$

در درون:

مولی محاسبه مقدار یک نقطه (x, y) در یک شبکه ۲x2 با نقاط (x_0, y_0) , (x_1, y_0) , (x_0, y_1) , (x_1, y_1) می‌باشد.

- ابتدا درون یابی خطی در جهت افقی انجام می‌شود:

$$f(x, y_0) = f(x_0, y_0) + \frac{(x - x_0)}{(x_1 - x_0)} \cdot (f(x_1, y_0) - f(x_0, y_0))$$

$$f(x, y_1) = f(x_0, y_1) + \frac{(x - x_0)}{(x_1 - x_0)} \cdot (f(x_1, y_1) - f(x_0, y_1))$$

- سپس درون یابی فضی در جهت عمودی انجام می‌شود:

$$f(x, y) = f(x, y_0) + \frac{y - y_0}{y_1 - y_0} \cdot (f(x, y_1) - f(x, y_0))$$

درون یابی nearest neighbor،
درین روش مقدار هر چهار نقطه پلیر یا مقدار رزون میکسرین پیکسل معهود در دادهای اصلی و کاری گزید

(2)

دندانه هایی بین تصویر اصلی و Activation map داراست با

$$S = \frac{\text{Original size}}{\text{Activation map size}} = \frac{32}{10} = 3.2$$

خطای نید پل تسویه از تصویر $f_{4,8}$ به Activation map

$$x_{\text{activation}} = \frac{4}{3.2} = 1.25 \quad y_{\text{activation}} = \frac{8}{3.2} = 2.5$$

$(1,2), (2,2), (1,3), (2,3)$: فقط یکی از این چهار نقطه در $(1.25, 2.5)$ است. این نقطه را در $f_{4,8}$ بگذار است با:

- درون یابی معنی دارد / نقیض

$$\begin{aligned} f(1.25, 2) &= f(1, 2) + \frac{1.25 - 1}{2 - 1} (f(2, 2) - f(1, 2)) \\ &= 0.75 f(1, 2) + 0.25 f(2, 2) \\ f(1.25, 3) &= 0.75 f(1, 3) + 0.25 f(2, 3) \end{aligned}$$

- درون یابی معنی دارد / نقیض:

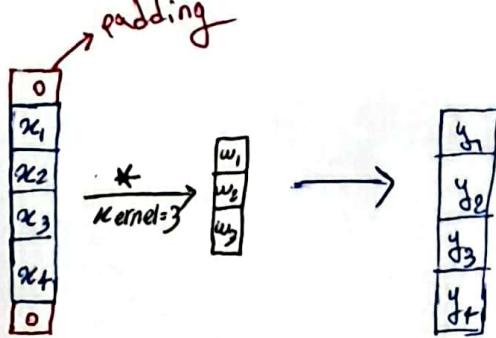
$$f(1.25, 2.5) = 0.5 f(1.25, 2) + 0.5 f(1.25, 3)$$

$$f(4,8) = (0.75 f(1,2) + 0.25 f(2,2)) 0.5 + (0.75 f(1,3) + 0.25 f(2,3)) 0.5$$

$$f(4,8) = 0.375 f(1,2) + 0.125 f(2,2) + 0.375 f(1,3) + 0.125 f(2,3)$$

دستورات

۱



عملیات لایه نشده به فرم:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} w_1 & w_2 & w_3 & 0 & 0 & 0 \\ 0 & w_1 & w_2 & w_3 & 0 & 0 \\ 0 & 0 & w_1 & w_2 & w_3 & 0 \\ 0 & 0 & 0 & w_1 & w_2 & w_3 \\ 0 & 0 & 0 & 0 & w_1 & w_2 \end{bmatrix} \begin{bmatrix} 0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ 0 \end{bmatrix} \rightarrow \begin{cases} y_1 = w_2 x_1 + w_3 x_2 \\ y_2 = w_1 x_1 + w_2 x_2 + w_3 x_3 \\ y_3 = w_1 x_2 + w_2 x_3 + w_3 x_4 \\ y_4 = w_1 x_3 + w_2 x_4 \end{cases}$$

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y_1} \frac{\partial y_1}{\partial x} + \frac{\partial L}{\partial y_2} \frac{\partial y_2}{\partial x} + \frac{\partial L}{\partial y_3} \frac{\partial y_3}{\partial x} + \frac{\partial L}{\partial y_4} \frac{\partial y_4}{\partial x}$$

$$\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial y_1} w_2 + \frac{\partial L}{\partial y_2} w_1 + \frac{\partial L}{\partial y_3} x_0 + \frac{\partial L}{\partial y_4} x_0 = \boxed{\frac{\partial L}{\partial y_1} w_2 + \frac{\partial L}{\partial y_2} w_1}$$

$$\frac{\partial L}{\partial x_2} = \frac{\partial L}{\partial y_1} w_3 + \frac{\partial L}{\partial y_2} w_2 + \frac{\partial L}{\partial y_3} w_1 + \frac{\partial L}{\partial y_4} x_0 = \boxed{\frac{\partial L}{\partial y_1} w_3 + \frac{\partial L}{\partial y_2} w_2 + \frac{\partial L}{\partial y_3} w_1}$$

$$\frac{\partial L}{\partial x_3} = \frac{\partial L}{\partial y_1} x_0 + \frac{\partial L}{\partial y_2} w_3 + \frac{\partial L}{\partial y_3} w_2 + \frac{\partial L}{\partial y_4} w_1 = \boxed{\frac{\partial L}{\partial y_2} w_3 + \frac{\partial L}{\partial y_3} w_2 + \frac{\partial L}{\partial y_4} w_1}$$

$$\frac{\partial L}{\partial x_4} = \frac{\partial L}{\partial y_1} x_0 + \frac{\partial L}{\partial y_2} x_0 + \frac{\partial L}{\partial y_3} x_0 + \frac{\partial L}{\partial y_4} w_2 = \boxed{\frac{\partial L}{\partial y_3} w_3 + \frac{\partial L}{\partial y_4} w_2}$$

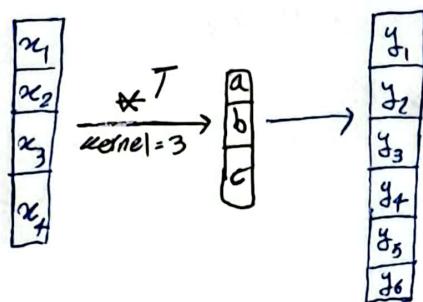
سایر مشتق تابع loss را بحسب بجا ای خروجی ها بدست آوریم که ممکن و این مشتق حداقت است به نتیجه خروجی ها را که خروجی است به ورودی مدقق ضرب کنیم. بنابراین اینکل مشتق تابع loss است به نتیجه روابط متألفی کسب.

$$\frac{\partial L}{\partial y} = \frac{\partial L}{\partial y_1} + \frac{\partial L}{\partial y_2} + \frac{\partial L}{\partial y_3} + \frac{\partial L}{\partial y_4}$$

برای این مرحله بعدی حساب کردن مشتق loss تابع را می‌توانیم با استفاده از قواعد مشتق کردن که در درس پیشین آموخته شده ایم.

$$\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial y_1} \frac{\partial y_1}{\partial x_1} \Rightarrow \frac{\partial L}{\partial y_1} \frac{\partial y_1}{\partial x_1} + \frac{\partial L}{\partial y_2} \frac{\partial y_2}{\partial x_1} + \frac{\partial L}{\partial y_3} \frac{\partial y_3}{\partial x_1} + \frac{\partial L}{\partial y_4} \frac{\partial y_4}{\partial x_1}$$

این مرحله را برای همه عکس کارکرد نمایم و در نهایت مشتق loss را بدست باید



امروزه

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} a & 0 & 0 & 0 \\ b & a & 0 & 0 \\ c & b & a & 0 \\ 0 & c & b & a \\ 0 & 0 & c & b \\ 0 & 0 & 0 & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{cases} y_1 = ax_1 \\ y_2 = bx_1 + ax_2 \\ y_3 = cx_1 + bx_2 + ax_3 \\ y_4 = cx_2 + bx_3 + ax_4 \\ y_5 = cx_3 + bx_4 \\ y_6 = cx_4 \end{cases}$$

ما نیز می‌توانیم padding بود که اگر $\text{padding} = 2$ و $\text{stride} = 2$ باشد
 همان معنی دارد که 4×4 کا ماتریس 4×4 سو فروخته شود و 6×6 ماتریس 3×3 خواهد شد.
 همان معنی دارد که 4×4 کا ماتریس 4×4 سو فروخته شود و 6×6 ماتریس 3×3 خواهد شد.

$$\begin{array}{c|c|c} \begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{matrix} & = & \left[\begin{matrix} a & 0 & 0 & 0 & 0 & 0 \\ b & a & 0 & 0 & 0 & 0 \\ c & b & a & 0 & 0 & 0 \\ 0 & c & b & a & 0 & 0 \\ 0 & 0 & c & b & a & 0 \\ 0 & 0 & 0 & c & b & a \\ 0 & 0 & 0 & 0 & c & b \\ 0 & 0 & 0 & 0 & 0 & c \end{matrix} \right] \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} \\ & = & \left\{ \begin{array}{l} y_1 = 0 \\ y_2 = ax_1 \\ y_3 = bx_1 + ax_2 \\ y_4 = cx_1 + bx_2 + ax_3 \\ y_5 = cx_2 + bx_3 + ax_4 \\ y_6 = cx_3 + bx_4 \\ y_7 = cx_4 \\ y_8 = 0 \end{array} \right. \end{array}$$

5 Nov

$$\begin{array}{l} a^{1+2} = A(z^{1+2}) \\ z^{1+2} = \omega^{1+2} a^{1+1} \end{array} \quad \begin{array}{l} a^{1+1} = A(z^{1+1}) \\ z^{1+1} = \omega^{1+1} a^1 \end{array}$$

1

$$\frac{\partial \alpha'^{l+2}}{\partial \alpha'} = \frac{\partial \alpha'^{l+2}}{\partial z^{l+2}} \times \frac{\partial z^{l+2}}{\partial \alpha'^{l+1}} \times \frac{\partial \alpha'^{l+1}}{\partial z^{l+1}} \times \frac{\partial z^{l+1}}{\partial \alpha'}$$

$$= A'(z^{l+2}) \times \omega'^{l+2} \times A'(z^{l+1}) \times \omega'^{l+1}$$

$$a^{l+1} = A(z^{l+1}) \quad z^{l+2} = w^{l+2} a^{l+2} \quad a^{l+2} = A(h^{l+2})$$

$$z^{l+1} = w^{l+1} \alpha^l$$

skip com - ! add (200)

$$\frac{\partial \alpha'^{l+2}}{\partial \alpha^l} = \frac{\partial \alpha'^{l+2}}{\partial h'^{l+2}} \times \frac{\partial h'^{l+2}}{\partial \alpha^l} = \frac{\partial \alpha'^{l+2}}{\partial h'^{l+2}} \times \left(\frac{\partial z'^{l+2}}{\partial \alpha^l} + \frac{\partial \alpha'^l}{\partial \alpha^l} \right)$$

کارهای اصلی در کنیت بود

• Caudal vanishing gradient

در شکه های عینق، گودلایان مرضی

$$= \frac{\partial \alpha^{l+2}}{\partial h^{l+2}} \times \frac{\partial z^{l+2}}{\partial \alpha^l} + \frac{\partial \alpha^{l+2}}{\partial h^{l+2}}$$

سالی فربستن در میشیگان

$$= A'(h^{1+2}) \times w^{1+2} \times A'(z^{1+1}) \times w^{1+1} + A'(h^{1+2})$$

کوچک بہ سرعت بھی صاف نہ دیکھ می شو فرد

Lip connection

۳) هم و همیکه قبل از آنچین activation func دلایل کم تابعی همچوی ناتوانی که ازی بدن باعث gradient loss باشود

$$\text{without skip connection} \quad , \quad \frac{\partial z^{l+2}}{\partial a^l} = w^{l+2} \times A'(z^{l+1}) \times w^{l+1}$$

$$\text{with skip connection: } \frac{\partial h^{l+2}}{\partial a^l} = \frac{\partial z^{l+2}}{\partial a^l} + \frac{\partial a^l}{\partial a^l} = w^{l+2} \times A'(z^{l+1}) \times w^{l+1} + 1$$

حالا که خوب کنیم خوب کنیم forward backward loss

باشد. فوی مقدار کوچکتر از اند $A' = 0.2$, $(0.8 \times 0.2 \times 0.8)^2 = 0.0256$ و $w^{l+2} \times w^{l+1}$ توأم تاثر

$$\text{without skip con.} \quad , \quad \frac{\partial z^{l+2}}{\partial a^l} = 0.8 \times 0.2 \times 0.8 = 0.128$$

$$\text{with skip connection: } \frac{\partial h^{l+2}}{\partial a^l} = 0.8 \times 0.2 \times 0.8 + 1 = 1.128$$

می بینیم که با اضافه شدن صفر ثابت به بحثی skip connection مقدار قابل گردایان را درین درجات مکمل می کند اگر سهیها دیگر تراشند این اعداد کوچکتر از ۱ هستند به صورت مدام در دenum ضرب شده و اعداد فیلی کوچکی می سازد که لذت مانع است می شود نسبتاً گزینهای برابری های لولی نزد

سؤال 6

Convolution

نحوی ۱

طبلوشن استاندارد:

فرض کنید ورودی از $H \times W \times C_{in}$ باشد. هر کاپسول $k \times k$ به C_{out} بسته باشد. خروجی از $H \times W \times C_{out}$ باشد.

$$\text{output}(i, j, \text{cout}) = \sum_{k=1}^K \sum_{l=1}^{C_{in}} \sum_{c_n=1}^{C_{in}} \text{Input}(i+k-1, j+l-1, c_n) \cdot \text{Filter}(k, l, c_n, \text{cout})$$

$K \cdot C_{in} \cdot \text{cout}$ بار محاسبه می شود.

طبلوشن عمیق فایل دهنده:

سایر دو مرحله است

Depthwise convolution - ۱

طبلوشن عمیق به طور جداگانه در هر کاپسول $k \times k$ برای C_{in} و C_{out} انجام می شود. بعد از آن مراحل دیگری دارند.

$$\text{Depthwise}(i, j, c_{in}) = \sum_{k=1}^K \sum_{l=1}^{C_{in}} \text{Input}(i+k-1, j+l-1, c_{in}) \cdot \text{Filter}_{\text{depth}}(k, l, c_{in})$$

Pointwise convolution ۲

باید طبلوشن عمیق مراحل قبل را مطلع کرد. همچوکی باشد که C_{in} تعداد پردازنده های ورودی باشد.

$$\text{output}(i, j, \text{cout}) = \sum_{c_{in}=1}^{C_{in}} \text{Depthwise}(i, j, c_{in}) \cdot \text{Filter}_{\text{pointwise}}(1, 1, c_{in}, \text{cout})$$

نوشته است: مراحل مطلع طبلوشن عمیق فایل دهنده و آنها را در درجه اول بازنشانی کن.

(2-1) ~~معنی~~ معنی $\text{W} = \text{U} \Sigma \text{V}^T$ میں عنصر $w_{i,j}$ دریک آن ویسے اسٹریک نہیں دارد بلکہ $w_{i,j} = \sum_k u_{ik} \sigma_k v_{kj}$

$$\text{W} = \text{U} \Sigma \text{V}^T = \sum_{i=1}^r \sigma_i \text{U}_i \text{V}_i^T$$

کوئی رکن r کا W کا رکن k کا $\hat{\text{W}}$ کا رکن k کا $\hat{\text{W}} = \sum_{i=1}^k \sigma_i \text{U}_i \text{V}_i^T$ دار

'depthwise conv' کا معنی
کوئی لسی عکس کا کافی تکمیل کرنے کا سبک W کا رکن k کا $\hat{\text{W}}$ کا رکن k کا $\hat{\text{W}} = \sum_{i=1}^k \sigma_i \text{U}_i \text{V}_i^T$ دار
کوئی خوبی پڑھنے کا سبک W کا رکن k کا $\hat{\text{W}} = \sum_{i=1}^k \sigma_i \text{U}_i \text{V}_i^T$ دار
کوئی خوبی پڑھنے کا سبک W کا رکن k کا $\hat{\text{W}} = \sum_{i=1}^k \sigma_i \text{U}_i \text{V}_i^T$ دار

(3-1) ~~معنی~~

$$\text{standard conv} \rightarrow \text{Flops}_{\text{std}} = H \times W \times C_{\text{in}} \times C_{\text{out}} \times k^2$$

depthwise ~~blops~~ separable convolution.

$$\text{- Depthwise Flops} \rightarrow \text{Flops}_{\text{depth}} = H \times W \times C_{\text{in}} \times k^2$$

$$\text{- pointwise Flops} \rightarrow \text{Flops}_{\text{point}} = H \times W \times C_{\text{in}} \times C_{\text{out}}$$

$$\text{Total Flops} \rightarrow \text{Flops}_{\text{sep}} = H \times W \times C_{\text{in}} (k^2 + C_{\text{out}})$$

Flops Reduction Ratio:

$$\frac{\text{Flops}_{\text{sep}}}{\text{Flops}_{\text{std}}} = \frac{k^2 + C_{\text{out}}}{k^2 \cdot C_{\text{out}}} = \frac{1}{C_{\text{out}}} + \frac{1}{k^2}$$

لیکن $C_{\text{out}} > k^2$ کا نتیجہ کوئی کمی نہیں کرے۔

2 c-تصویر

(جتنی 2-1)

دلوکت باقیمانده مخصوص:
موارد:

ا) افزایش تعداد کمالها با طبقه‌بندی اول و فرمیت اندیساد $t = 6\text{ min}$; Expansion - 1
ب) اعمال کامپکشن (compress) ۳×۳ ، depthwise - 2
ج) اعمال کامپکشن (compress) ۱×۱ ، ابعاد اولی با کامپکشن ایجاد - 3

در این کمالها کامپکشن یا پردازش شبکه را برای mobiles ۷/۲ رسنت (Resnet) می‌گذرانند از طریق
تفصیلی (projection) استفاده شود (بدین تابع مصالحانی ResNet ناگزینی از این دستورات جلوگیری نمود)
اعمال باقیمانده همان حالا بروز نموده که ابعاد معمولی خود را در جلسه باشند

(جتنی 2-2)

کامپکشن عمقی قابل تفکیک

- در این اطلاعات مطابق با این کامپکشن ها ابتدا اعمالات مقسماً (3x3) را بعدها می‌گذارند
- سه مرحله ایجادی دارند که کامپکشن از آنها که کامپکشن از آنها که کامپکشن از آنها که کامپکشن از آنها که
- و محاسبه کمالی: در درجه point wise اطلاعات کمالی کامپکشن شود

کامپکشن عمقی:

با طبقه‌بندی ابعاد کمالها (با طبقه‌بندی ۱×۱) و حفظ اطلاعات ضروری کمالی مسروق شود
ادخلات از تغییر در ترتیب عالی صورت رفته ای کم بعد جلوگیری می‌کند

در نتیجه این ادعا صحیح است

(جتنی 2-3)

ناتیر در معنی اراده ایان
الف) افزایش تعداد کمالها کامپکشن یافته افزایشی برای و تعداد اسبریانی بثمری ای اندسارتراپیان
ب) اسپارچ شدن که ای تو اندسارتراپی (sparsification) / کامپکشن در

$$\text{Var}\left(\frac{\partial L}{\partial w}\right) \propto \frac{1}{t \cdot \text{Cin}} \rightarrow \text{افزایش ای تو اندسارتراپی / کامپکشن بهم} \\ \text{و جزئیات پایه ای کامپکشن داشت}$$

هودینه طامباچی

$$F_{\text{Log}} \propto H.W. (c_{in}^2 \cdot t + c_{in} \cdot t \cdot k^2)$$

بنادران با غرایش بزرگ به حفظ های ملی افزایش می یابند

الرسائل

الآن نحن على بعد خطوة من الانتهاء / المدخلات t=6 على mobilenetv2

ای که وجود در افغانستان و حیان گردانی بقایه و کاهش هزینه trade-off

3 Сюжет

(1-3 جزء)

۱. مشهد مساجدی (Squeezed)

وتحل محل الـ max pooling

$$z_c = \frac{1}{H \cdot W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \rightarrow z_c \in \mathbb{R}^C$$

$C \times H \times W$, but $b_{\text{feature map}} = X$

: (Excitation) پری ۲۵. 2

وأتساع في كل قنواته بحسب مفهوم الـ fully connected \approx $O(2^{\text{number of layers}})$ / channel wise

$$S = \sigma(W_2 \cdot S(W_1, Z))$$

$w, \epsilon R^{F_{XC}}$: First FC layer (reduction)

8. ReLU activation

$w_2 \in R^{C \times \frac{E}{T}}$, second Fe layer (expansion)

σ : sigmoid activation

Questions. 3

(green feature maps per channel-wise case) / 61

$$\tilde{x}_c(i, j) = \varepsilon_c \cdot x_c(i, j)$$

Output: $\tilde{X} \in \mathbb{R}^{CxHxW}$

(2-3) بینت

Metal information perspective

۱. اطلاعات ملکی درونی - خروجی: هرگز اطلاعات ملکی درونی فقط قدر اندیشید

ملکیت اندیشید که اطلاعات ملکی درونی کم مابین ۰ و ۱ است و ملکیت اخراجی اندیشید که اندیشید که اندیشید اخراجی اندیشید (۰ < x < ۱) سود را از خروجی ملکی اخراجی شنیده باشند همانند غنیمتی که ملکیت اخراجی بزرگ است

: Impact of Reduction Ratio r

۲. خروجی: پارامترهای کنترل در الگوریتم (FC و عوامل) وجود دارد که حجم حافظه را کاهش می‌دهند اما توانی این را برای پارامترهای پیشنهادی بسیار بسیار کم کنند

کنترل: پارامترهای پیشنهادی معمولاً اینستین ملکیت اخراجی را بهبودی نمی‌نمایند اما بهبود اینستین ملکیت اخراجی را بهبودی نمی‌نمایند

Trade-off

$r=16$ / $r=4$ / mobile netv3. و عکس از عکس (Flops $\propto \frac{c^2}{r}$) که حجم حافظه را کاهش می‌نماید استفاده کنند

3-3 بینت

پس از $H \times W$ سلسله از سلسله از r depthwise، $\leq E$ فلوبز را کاهش می‌نماید

$$\text{SE Block } (c=64, r=16) \rightarrow \text{Flop}_{\text{SE}} = 2 \cdot \frac{c^2}{r} \rightarrow c \times H \times W = 512 + 64 H \times W$$

Depthwise separable conv $\rightarrow \text{Flop}_{\text{depth}} = 4672 H \times W$

$$\text{Flops Ration} \rightarrow \text{Ratio} = \frac{\text{Flop}_{\text{SE}}}{\text{Flop}_{\text{depth}}} = \frac{512 + 64 H \times W}{4672 \times H \times W}$$

$$\text{small } (H=W=7) \rightarrow \text{Ratio} = \frac{512 + 64 \times 49}{4672 \times 49} = 1.59\%$$

$$\text{medium } (H=W=14) \rightarrow \text{Ratio} \approx 1.43\%$$

$$\text{large } (H=W=224) \rightarrow \text{Ratio} \approx 1.37\%$$

$$H, W \rightarrow \infty \rightarrow \text{Ratio} \approx 1.37\%$$

فصل ۴

پنجم ۴-۱)

موارد NAS در mobilenetV3:

NAS دلیلی خود کار رایج نباید باشد بلکه شبکه های عصبی بسیار شده برای طبقه های خاص (مانند دستگاه های اینترنتی) است در mobilenetV3، NAS ۱ هدف تقابلی بین دقیقت و مخفیت را دارد. این خواهد ماند تا در رسیدن توکیمی از عملیات دستگاه فرآیند و تنظیم پارامترهای محاسباتی را کنسته می کند.

ضدای جستجو:

انواع لایه ها:

- بلوک های رایج مانند مکون (ناپراپ ایمپلیکت و تقاضت)

- بلوک های SE

- Depthwise separable convolutions

- توابع فعالسازی غیر ضعی

باراهمزه های دیگر بینی:

- تعداد کالا ها در هر لایه

- اندازه Kernel های کانولوشن

- پردازش stride بایان نمودن در لایه

- طایفه بلوک های SE در محاسبات

در لایه هایی که جفت دارند دور (DAG)

جفت DAG: جفتی که در آن گراف دوی علیات متسابق / رشتانی را دارند
باشند: جفت جوابان داشتنی که هما را مستقر می کنند. اینا خروجی دوی طبقه بسیار عجیب به واسطه
نه بلوکی SE (که می شود)

جشن ۲-۴

کوچک ساری پیشونده یک استرالی را از خوار و ساده ساری تبدیلی هماری شبک در فوایند کرد این روش با آموزش یک ابو قبک بزرگ آغازی شد و میتوان با حفظ اجزای غیر ضروری هماری را بینیافتد

مراحل :

۱- مقداری اولیه آموزش یک اندیشه انتسابی همانی ممکن است ادامه شود

۲- هرس تدریجی :

پیشنهادی مبنی بر تغییر ارزیابی نیز شبکه ای ساخته افراد دعوه شده اند اولویت دهنده عملیات

نماینده

کامپیوترها کامپیوتر تدریجی تعداد کامپیوترها کامپیوترها

عنوان لایه ها: حرف بلوک ها الایه های اضافی

۳- سلیم دقیق آموزش جدید همانی ساده شده برای ~~لهم~~ از نظر دقت

تا شنبه

عادل یعنی اندیشه ملمس است و دست نارخواه یک سازگار را ساخته اندیشه

کامپیوتری پیش از حد که با تطبیق تدریجی عملی ساده شده باشد