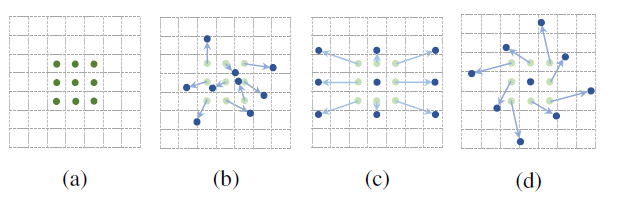
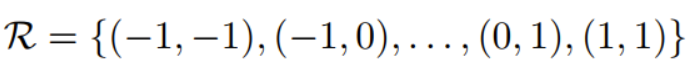
1)

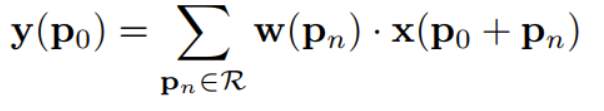
در deformable convolution به منظور داشتن receptive fieldهای متفاوت افست‌های دو بعدی به مکان‌های نمونه‌گیری منظم grid در کانولوشن استاندارد افزوده می‌شود.



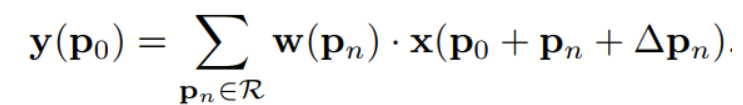
در شکل فوق در قسمت a یک کانولوشن استاندارد را می‌بینیم که به صورت معمولی مطابق با نقاط کانولوشن استاندارد نمونه برداری می‌شود یعنی وزن‌های کانولوشن را را در نقاط منظم (سبز) ضرب و نتایج را جمع می‌کند. اما در deformable convolution به جای نقاط سبز آفست‌هایی دو بعدی به مختصات نقاط کانولوشن استاندارد (نقاط سبز) افزوده می‌شود که این آفست‌ها قابل یادگیری هستند و وزن‌های کانولوشن در نقاط جدید ضرب و نتایج جمع می‌شود (b, c, d)



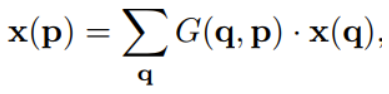
با توجه به کرنل R فوق کانولوشن استاندارد به صورت زیر محاسبه می‌گردد:



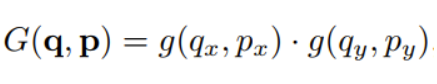
اما در deformable convolution مقدار Δpn که همان آفست هست به p0 + pn افزوده می‌شود و حاصل نهایی به صورت زیر محاسبه می‌شود.



برای اجرای معادله فوق از درون‌یابی دوخطی (bilinear interpolation) استفاده می کنیم زیرا با اضافه کردن آفست به موقعیت‌های نمونه موجود، نقاط اعشاری را به دست می‌آوریم که مکان‌های مشخصی در شبکه نیستند و برای تخمین مقادیر پیکسل آن‌ها از درون‌یابی دوخطی استفاده می‌کنیم که از شبکه 2x2 پیکسل مجاور برای تخمین مقدار موقعیت پیکسل تغییر شکل یافته جدید استفاده می‌کند.

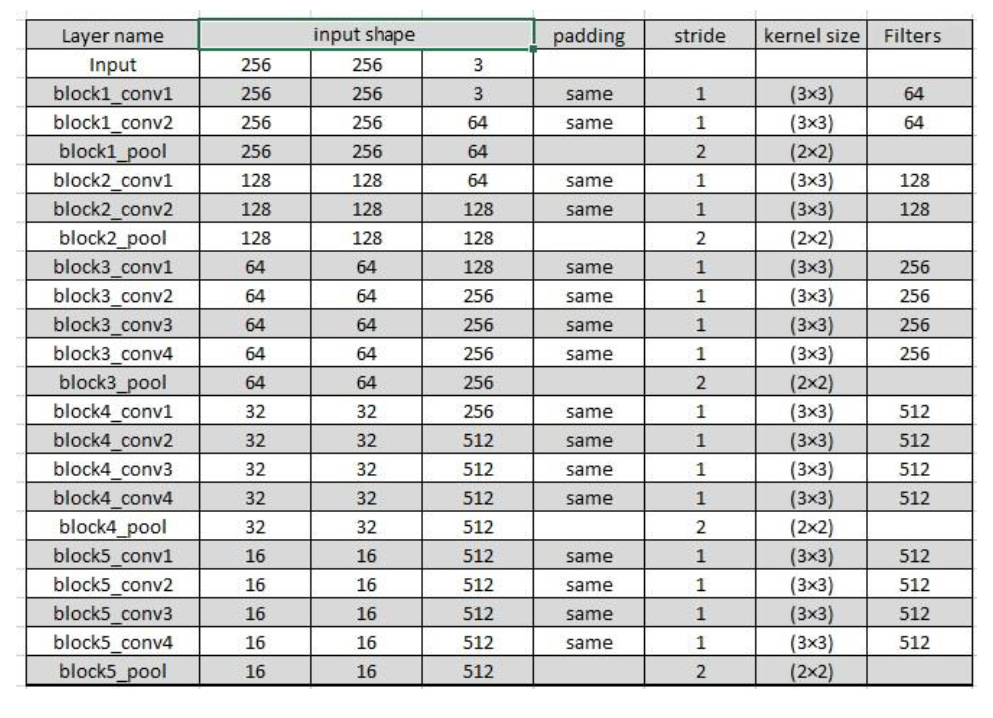


که p = p0 + pn + Δpn و q پیکسل‌های مجاور هستند و G کرنل درونیابی دو خطی است. G دو بعدی است و می‌توانیم به صورت زیر آن را به دو کرنل یک بعدی تقسیم می‌شود



2)

الف)



ابتدا باید shape خروجی block5\_pool را محاسبه کرد:

block5\_pool.shape = 8, 8, 512

Flatten:

flatten\_output.shape = 8 \* 8 \* 512 = 32768

flatten\_number\_of\_parameters = 0

dense\_number\_of\_parameters = 30 \* (32768 + 1) = 983070

GAP:

gap\_output.shape = 512

gap\_number\_of\_parameters = 0

dense\_number\_of\_parameters = 30 \* (512 + 1) = 1530

با وزن یکسان برای تمام کانالها GWAP:

GWAP\_e\_output.shape = 512

GWAP\_e\_number\_of\_parameters = 8 \* 8 = 64

dense\_number\_of\_parameters = 30 \* (512 + 1) = 1530

با وزن متفاوت برای هر کانال GWAP:

GWAP\_d\_output.shape = 512

GWAP\_d\_number\_of\_parameters = 8 \* 8 \* 512 = 32768

dense\_number\_of\_parameters = 30 \* (512 + 1) = 1530

هیستوگرام قابل آموزش با bin = 4:

Histogram\_layer\_parameters: 512 \* 4 \* 2 = 4096

Histogram\_layer\_output.shape = 512 \* 4 = 2048

dense\_number\_of\_parameters = 30 \* (2048 + 1) = 61470

هیستوگرام قابل آموزش با bin = 8:

Histogram\_layer\_parameters: 512 \* 8 \* 2 = 8192

Histogram\_layer\_output.shape = 512 \* 8 = 4096

dense\_number\_of\_parameters = 30 \* (4096 + 1) = 122910

ب)

Flatten:

مزیت: اطلاعات مکانی را دور نمی‌ریزد و می‌تواند با استفاده از آن‌ها عملکرد بهتری داشته باشد و مثلا تشخیص می‌دهد یک چیزی در بالای تصویر است یا پایین تصویر

عیب: تعداد پارامترهای لایه کاملا متصل بعد از آن بسیار زیاد است.

GAP:

مزیت: تعداد پارامترهای لایه کاملا متصل بعد از آن را به شدت کاهش می‌دهد ولی با این وجود عملکرد خوبی دارد.

عیب: اطلاعات مکانی را دور می‌ریزد. ضمن اینکه ممکن است با توجه به داده و وظیفه مورد نظر میانگین گرفتن مناسب نباشد و انتخاب max مناسبتر باشد.

GWAP با وزن یکسان برای تمام کانالها:

مزیت: تعداد پارامترهای لایه کاملا متصل بعد از آن را به شدت کاهش می‌دهد ولی با این وجود عملکرد خوبی دارد. نسبت به GAP اطلاعات مکانی بیشتری را ذخیره می‌کند با اینکه پارامترهای افزوده شده زیاد نیست.

عیب: اطلاعات مکانی را دور می‌ریزد. روی اطلاعات تمام کانال‌ها میانگین گرفته می‌شود و اهمیت هر کانال به یک اندازه در نظر گرفته می‌شود. ضمن اینکه ممکن است با توجه به داده و وظیفه مورد نظر میانگین گرفتن مناسب نباشد و انتخاب max مناسبتر باشد.

GWAP با وزن متفاوت برای هر کانال:

مزیت: تعداد پارامترهای لایه کاملا متصل بعد از آن را به کاهش می‌دهد. نسبت به GWAP با وزن یکسان برای تمام کانالها اطلاعات مکانی بیشتری را ذخیره می‌کند اطلاعات هر کانال مستقل از کانال‌های دیگر است. عملکرد آن به شدت نسبت به GAP و GWAP با وزن یکسان برای تمام کانالها بهتر است.

معایب: ممکن است با توجه به داده و وظیفه مورد نظر میانگین گرفتن مناسب نباشد و انتخاب max مناسبتر باشد. کمی از اطلاعات مکانی دور ریخته می‌شود.

هیستوگرام قابل آموزش با bin=4:

مزیت: به جای میانگین گرفتن هیستوگرام را ذخیره می‌کند بنابراین اطلاعات بهتری را نگه می‌دارد. تعداد پارامترها برای لایه کاملا متصل بعد نسبت به flatten کاهش میابد اما اطلاعات بهتری نسبت به روشهای GAP و حالتهای متفاوت GWAP ذخیره می‌کند.

معایب: پارامترها برای لایه کاملا متصل بعد آن از روش GAP و حالتهای متفاوت GWAP بسیار بیشتر است. اطلاعات مکانی کمتری نسبت به flatten دارد.

هیستوگرام قابل آموزش با bin=8:

مزیت: نسبت به هیستوگرام قابل آموزش با bin=4 با بیشتر شدن bin عملکرد بهتری دارد. تعداد پارامترها برای لایه کاملا متصل بعد نسبت به flatten کاهش میابد اما اطلاعات بهتری نسبت به روشهای GAP و حالتهای متفاوت GWAP ذخیره می‌کند.

معایب: نسبت به هیستوگرام قابل آموزش با bin=4 با بیشتر شدن bin پارامترهای بیشتری دارد. اطلاعات مکانی کمتری نسبت به flatten دارد.

3)

شبکه unet عملکرد خوبی نسبت به dlinknet ندارد اما سریعتر است. شبکه dlinknet به دلیل استفاده از resnet50 حدودا 2 برابر پارامتر دارد و آموزش آن به زمان بیشتری نیاز دارد و عملکرد موفق‌تری دارد.