

بخش اول – سوالات تئوری مربوط به تمرین عملی:

آ) در شبکه‌های کانولوشن عادی از یک grid ثابت برای sampling از تصویر استفاده می‌شود. این grid به صورت منظم و با فاصله ثابت بر روی تصویر حرکت می‌کند و اطلاعات از ناحیه‌های تصویری مشخصی را برای استخراج ویژگی‌ها مورد استفاده قرار می‌دهد. این مکانیزم استفاده از یک grid ثابت به تصویر اجازه مدل بهبود ناپذیر در مقابل تغییرات ناگهانی در تصویر را نمی‌دهد.

در مقابل، در شبکه‌های کانولوشن Deformable، از یک sampling grid با قابلیت تغییر (deformable) استفاده می‌شود. این به شبکه این امکان را می‌دهد که مکان‌های sampling را بر اساس محتوای تصویر و شکل آن تعیین کند. به عبارت دیگر، ماسک کانولوشن در هر لحظه می‌تواند تغییر شکل یابد و به نواحی مختلف تصویر برای استخراج ویژگی‌ها بپردازد. این ویژگی می‌تواند بهبود قابل توجهی در تشخیص اشیاء با شکل‌ها و انحرافات مختلف در تصویر داشته باشد.

بنابراین، تفاوت اصلی بین شبکه‌های کانولوشن عادی و کانولوشن Deformable در نحوه تعیین مکان‌های sampling است. در کانولوشن Deformable، مکان‌های sampling بر اساس محتوای تصویر و شکل آن تعیین می‌شوند و این امکان به شبکه می‌دهد که بهتر با تغییرات ناگهانی و اشکال مختلف در تصویر برخورد کند.

ب) شبکه‌های Deformable از انعطاف پذیری در ترانسفورمیشن‌های جابجایی (Geometric transformations) برای استخراج ویژگی‌های بهتر از تصاویر با اشکال مختلف بهره‌مند می‌شوند. این انعطاف پذیری به دلیل استفاده از یک sampling grid با امکان تغییر شکل (deformable sampling grid) در لایه کانولوشن Deformable به وجود می‌آید. در ادامه توضیحاتی در مورد چگونگی ایجاد انعطاف پذیری در این شبکه‌ها آورده شده است:

Deformable Convolution:

در یک لایه Deformable Convolution، به جای استفاده از یک grid ثابت برای sampling از تصویر، از یک مکانیزم deformable sampling grid استفاده می‌شود. این grid به واسطه یک مدل یادگیری شده تغییر شکل می‌یابد. این مدل شکل‌های مختلفی از sampling grid را یاد می‌گیرد که بتواند بهترین تطابق را با محتوای تصویر ایجاد کند.

Learning Deformable Fields:

انعطاف پذیری در تصویر با استفاده از یادگیری میدان‌های تغییر شکل ایجاد می‌شود. این میدان‌ها به نوعی نشان می‌دهند که در هر نقطه از sampling grid چقدر باید حرکت کند تا بهترین نتیجه را برای استخراج ویژگی‌ها از تصویر بدهد. این میدان‌ها به صورت یادگیری شده توسط شبکه عصبی اصطلاحاً deformable fields نامیده می‌شوند.

Adaptation to Object Shapes:

این امکان به شبکه می‌دهد که به شکل‌ها و انحرافات مختلف در تصویر پاسخ دهد. به عبارت دیگر، اگر تصویر شامل اشیاء با اشکال یا انحرافات مختلف باشد، شبکه با تغییرات در deformable sampling grid به بهترین شکل ممکن با این اشیاء هماهنگ می‌شود.

به این ترتیب، استفاده از Deformable Convolution باعث می‌شود که شبکه عصبی انعطاف پذیرتر به تغییرات هندسی و اشکال مختلف در تصاویر پاسخ دهد، و این امکان را فراهم می‌کند تا ویژگی‌های بهتری از تصاویر با ابعاد و شکل‌های مختلف استخراج شوند.

ج) شبکه‌های کانولوشن ساده ممکن است در مواجهه با تصاویری که دارای تغییرات چرخش یا تغییرات فضایی زیادی هستند، دچار مشکلاتی شوند. دلایل این مشکلات شامل عدم انعطاف پذیری به چرخش، حساسیت محدود به تغییرات ترجمه، استفاده از یک grid ثابت برای sampling، و اندازه محدود برای receptive field هستند. در مقابل، شبکه‌های کانولوشن با قابلیت انعطاف پذیری مثل شبکه‌های Deformable Convolution می‌توانند بهتر با تغییرات هندسی مختلف مقابله کنند و بهترین نتایج را ارائه دهند.

د) در لایه‌های شبکه، افست‌ها یا مکان‌های sampling به واسطه deformable sampling grid و میدان‌های تغییر شکل (deformable fields) محاسبه می‌شوند. این افست‌ها مشخص می‌کنند که هر نقطه از grid باید به کدام نقطه از تصویر نگاشت شود تا بهترین نتیجه حاصل شود.

مراحل محاسبه افست در یک لایه Convolution Deformable به صورت خلاصه به شرح زیر است:

ورودی‌ها:

ویژگی‌های ورودی از لایه قبلی (ممکن است با یک کرنل کانولوشن معمولی به دست آمده باشند).

میدان‌های تغییر شکل (deformable fields) که توسط شبکه یادگیری می‌شوند.

محاسبه موقعیت‌های Sampling

از deformable fields برای محاسبه تغییرات مکانی در sampling grid استفاده می‌شود.

مکانیزم یادگیری در deformable fields نقاط grid را به نقاط مختلف تصویر نگاشت می‌کند.

اعمال افست به ویژگی‌ها:

افست‌های محاسبه شده از deformable fields به ویژگی‌های ورودی اعمال می‌شوند.

این عمل باعث تغییر مکان sampling grid می‌شود و این تغییرات بهترین تطابق با محتوای تصویر را فراهم می‌کند.

اعمال کرنل کانولوشن:

پس از محاسبه افست‌ها، از یک کرنل کانولوشن به ویژگی‌های ورودی با افست‌ها اعمال می‌شود.

این کرنل کانولوشن ممکن است یک کرنل معمولی یا یک کرنل تطبیقی باشد.

به این ترتیب، افست‌ها در Convolution Deformable با استفاده از deformable sampling grid و میدان‌های تغییر شکل محاسبه می‌شوند. این مکانیزم به شبکه امکان می‌دهد تا به تغییرات هندسی در تصویر با انعطاف پذیری بیشتری پاسخ دهد و از اشیاء با اشکال و انحرافات مختلف بهترین ویژگی‌ها را استخراج کند.