

# توجه در یادگیری ماشین

سجاد سیزی  
محمدرضا احمدی تشنیزی

۲۶ آبان ۱۴۰۲

## مقدمه

شبکه‌های عصبی کانولوشنی مبتنی بر توجه به یک معماری شبکه‌های عصبی کانولوشنی (CNN) اشاره دارد که مکانیسم‌های توجه را در خود گنجانده است. مکانیسم‌های توجه با الهام از توجه بصری انسان و امکان شبیه‌سازی توجه مدل به بخش‌های خاص ورودی در حین انجام یک وظیفه، به مدل این امکان را می‌دهند. در زمینه‌ی CNNها، مکانیسم‌های توجه به طور معمول به منظور بهبود قابلیت مدل در گرفتن ویژگی‌ها و ارتباطات مرتبط در داخل یک تصویر، یکپارچه می‌شوند. این مکانیسم‌ها به شبکه کمک می‌کنند تا به‌طور انتخابی به مناطق مهم ورودی توجه کند و وزن بیشتری به ویژگی‌های خاصی بدهد. لایه‌های خاص شبکه در یک CNN مبتنی بر توجه ممکن است بسته به معماری متفاوت باشد، اما یک رویکرد متداول شامل یکی از لایه‌های زیر است:

- Convolutional Layers: این لایه‌ها مسئول استخراج ویژگی‌های محلی از داده ورودی هستند. در یک CNN مبتنی بر توجه، مکانیسم‌های توجه ممکن است بر روی خروجی لایه‌های کانولوشنال اعمال شود تا اهمیت خاصی به برخی از نقشه‌های ویژگی داده شود.
  - Attention Mechanism Layers: این لایه‌ها امتیازهای توجه را محاسبه کرده و وزن‌ها را به اقسام مختلف ورودی اختصاص می‌دهند. مکانیسم‌های توجه مختلفی وجود دارند، مانند توجه فضایی یا توجه به کانال. توجه فضایی به نقاط فضایی خاص در ورودی تمرکز می‌کند، در حالی که توجه به کانال بر روی کانال‌های خاص در نقشه‌های ویژگی تأکید می‌کند.
  - Pooling Layers: لایه‌های پولینگ اغلب برای کاهش ابعاد فضایی نقشه‌های ویژگی استفاده می‌شوند. های CNN مبتنی بر توجه ممکن است از مکانیسم‌های توجه در ارتباط با لایه‌های پولینگ استفاده کنند تا راهنمایی در فرآیند کاهش ابعاد فضایی داشته باشند.
  - Fully Connected Layers: این لایه‌ها که به عنوان لایه‌های چگال هم شناخته می‌شوند، به طور معمول در انتهای شبکه قرار دارند. مکانیسم‌های توجه ممکن است در اینجا اعمال شود تا وابستگی‌های جهانی در فضای ویژگی را گرفته شود.
- طراحی خاص و ادغام مکانیسم‌های توجه ممکن است بین معماری‌های مختلف CNN مبتنی بر توجه متفاوت باشد. برخی از مکانیسم‌های توجه محبوب عبارتند از:
- Self-Attention: همچنین به عنوان توجه داخلی شناخته می‌شود، به مدل این امکان را می‌دهد که به مواقع مختلف ورودی وزن‌های مختلفی اختصاص دهد و این امکان را فراهم کند که در طول پردازش بر روی بخش‌های مرتبط تمرکز کند.
  - Spatial Attention: بخشی از ورودی را به تأکید بر مناطق فضایی خاص آن امکان می‌دهد تا مدل در حین پردازش بر روی مناطق خاصی تمرکز کند.

- Channel Attention: کانال‌های مهم در نقشه‌های ویژگی را برجسته می‌کند و به مدل کمک می‌کند تا بر روی ویژگی‌های خاصی تمرکز کند.

## خلاصه توجه به خود (Self-Attention)

یکی از مثال‌های معماری شبکه‌های عصبی کانولوشنی مبتنی بر توجه، ترانسفورمر است که ابتدا برای وظایف پردازش زبان طبیعی طراحی شد، اما با موفقیت به وظایف بینایی کامپیوتری نیز اعمال شده است.

### خلاصه توجه به خود

توجه به خود، همچنین به عنوان توجه نقطه‌ای مقیاس شده شناخته می‌شود، یک مکانیسم است که به یک مدل این امکان را می‌دهد تا هنگام پیش‌بینی یا کدگذاری اطلاعات، اجزای مختلف ورودی را به طور مختلف وزن دهد. این مکانیسم به طور معمول در معماری ترانسفورمر استفاده می‌شود که ابتدا برای وظایف پردازش زبان طبیعی طراحی شده بود، اما بعدها با موفقیت به بینایی کامپیوتری نیز اعمال شد.

### اجزای کلیدی توجه به خود

- بردارهای پرسش، کلید، و مقدار: توجه به خود شامل سه بردار برای هر عنصر ورودی است - بردار پرسش، بردار کلید، و بردار مقدار.
- امتیازهای توجه: شباهت بین بردارهای پرسش و کلید میزان امتیازهای توجه را تعیین می‌کند. این امتیازها سپس برای وزن دهی به مقادیر متناظر استفاده می‌شوند.
- توجه نقطه‌ای مقیاس شده: امتیازهای توجه با گرفتن ضرب نقطه‌ای بین بردارهای پرسش و کلید محاسبه می‌شوند، که توسط جذر مربعی از ابعاد بردارهای کلید مقیاس شده‌اند. امتیازهای حاصل سپس برای وزن دهی به بردارهای مقدار استفاده می‌شوند.
- توجه چندسر: چندین مجموعه از بردارهای پرسش، کلید، و مقدار به صورت موازی (چندسر) استفاده می‌شوند و نتایج آنها ادغام و به صورت خطی تبدیل می‌شوند تا خروجی نهایی حاصل شود.

### مزایای توجه به خود

- وابستگی‌های دورتاه: توجه به خود به مدل این امکان را می‌دهد که وابستگی‌های دورتاه در دنباله ورودی را گرفته و برای وظایف نیازمند درک متنی استفاده شود.
- موازی‌سازی: مکانیسم توجه به خود بسیار قابل موازی‌سازی است که امکان آموزش کارآمد را بر روی سخت‌افزارهای مدرن فراهم می‌کند.
- انعطاف‌پذیری: توجه به خود توسط میدان‌های تأثیر ثابت اندازه محدود نشده است که این امکان را فراهم می‌کند تا با طول‌ها و ساختارهای ورودی مختلف سازگار باشد.

### چالش‌ها

- پیچیدگی محاسباتی: پیچیدگی زمانی توجه به خود نسبت به طول دنباله ورودی به صورت مربعی است که ممکن است برای دنباله‌های طولانی هزینه محاسباتی زیادی داشته باشد.
- قابلیت تفسیر: امتیازهای توجه توسط توجه به خود تولید شده ممکن است دشوار به تفسیر باشد، به ویژه در مدل‌های پیچیده با چندسر توجه.

## موارد استفاده

- پردازش زبان طبیعی: توجه به خود به عنوان یک مکانیسم بسیار موثر در وظایف پردازش زبان طبیعی مانند ترجمه ماشینی، خلاصه‌سازی متن، و درک زبانی ثابت شده است.
  - بینایی کامپیوتری: معماری‌های مبتنی بر ترانسفورمر با لایه‌های توجه به خود با موفقیت در وظایف دسته‌بندی تصویر، شناسایی اشیاء، و تولید تصویر به کار رفته‌اند، که نشان‌دهنده چندپذیری آنها فراتر از وظایف مرتبط با زبان است.
  - شناسایی گفتار: مکانیسم‌های توجه به خود همچنین در سیستم‌های شناسایی گفتار برای گرفتن اطلاعات زمینه‌ای و وابستگی‌های میان ویژگی‌های صوتی مورد استفاده قرار می‌گیرند.
- به طور خلاصه، توجه به خود یک مکانیسم قدرتمند برای گرفتن روابط در دنباله‌های ورودی است که یک راه حل انعطاف‌پذیر و موثر برای وظایف مختلف یادگیری ماشین فراهم می‌کند. با وجود چالش‌های محاسباتی آن، مزایای آن در گرفتن وابستگی‌های دورتاه منجر به پذیرش گسترده آن در حوزه‌های مختلف شده است.

## پیوند به کد نمونه

https://github.com/huggingface/transformers - کتابخانه Transformers از Face: Hugging یک کتابخانه از Face Hugging که مدل‌های ترانسفورمر پیش‌آموزش داده شده و رابط‌های ساده برای کار با آنها فراهم می‌کند.

## خلاصه توجه فضایی (Spatial Attention)

توجه فضایی یک مکانیسم استفاده شده در شبکه‌های عصبی، به ویژه در شبکه‌های عصبی کانولوشنی (CNN) است که به انتخابی هدفمند بر روی مناطق فضایی خاص داده ورودی تمرکز می‌کند. برخلاف توجه به خود که بر روی کل دنباله عمل می‌کند، توجه فضایی با تأکید یا کاهش توجه به مکان‌های فضایی خاص در یک نقشه ویژگی ورودی تکی مشغول است.

## اجزای کلیدی توجه فضایی

- نقشه‌های پرسش، کلید، و مقدار: همانند توجه به خود، توجه فضایی شامل محاسبه امتیازهای توجه بر اساس نقشه‌های پرسش، کلید، و مقدار است. این نقشه‌ها از نقشه ویژگی ورودی اشتقاق می‌شوند.
- امتیازهای توجه: امتیازهای توجه با اندازه‌گیری شباهت بین نقشه‌های پرسش و کلید برای هر مکان فضایی محاسبه می‌شوند. شباهت بالاتر به امتیازهای توجه بالاتر منجر می‌شود.
- مجموع وزن‌دار: امتیازهای توجه برای وزن‌دار کردن مقادیر متناظر در هر مکان فضایی استفاده می‌شوند. مجموع وزن‌دار این مقادیر خروجی مکانیسم توجه فضایی را تشکیل می‌دهد.

## مزایای توجه فضایی

- تمرکز انتخابی: توجه فضایی به مدل این امکان را می‌دهد که به صورت انتخابی بر روی مناطق مهم ورودی تمرکز کند و توانایی استخراج ویژگی مؤثرتری را فراهم سازد.
- بهبود عملکرد: با اختصاص وزن بیشتر به مکان‌های فضایی مرتبط، توجه فضایی می‌تواند توانایی شبکه را در گرفتن الگوهای معنادار در داده افزایش دهد و به بهبود عملکرد منجر شود.
- انعطاف در برابر تغییرات: توجه فضایی می‌تواند شبکه‌های CNN را در مقابل تغییرات در مقیاس، جهت، و موقعیت اشیاء در یک تصویر مقاومت‌پذیرتر کند.

## چالش‌ها

- هزینه محاسباتی: پیاده‌سازی توجه فضایی هزینه محاسباتی مدل را افزایش می‌دهد که ممکن است نگران‌کننده باشد، به ویژه برای برنامه‌های زمان‌واقعی یا محیط‌های منابع محدود.
- قابلیت تفسیر: همانند توجه به خود، تفسیر امتیازهای توجه در توجه فضایی می‌تواند چالش‌ناپذیر باشد، که سختی تفهیم مناطق فضایی مهم برای تصمیم مدل را افزایش می‌دهد.

## موارد استفاده

- دسته‌بندی تصویر: توجه فضایی به طور معمول در وظایف دسته‌بندی تصویر برای تمرکز بر روی بخش‌های مرتبط یک تصویر و بهبود توانایی تمیزبندی مدل استفاده می‌شود.
  - شناسایی اشیاء: در شناسایی اشیاء، توجه فضایی به مدل کمک می‌کند تا بر روی مناطق مورد نظر در یک تصویر تمرکز کرده و به بهبود موقعیت‌یابی دقیق‌تر اشیاء کمک کند.
  - تقسیم‌بندی معنایی: توجه فضایی برای وظایف تقسیم‌بندی معنایی مفید است که هدف آن تصنیف هر پیکسل در یک تصویر است، زیرا این مکانیسم به مدل کمک می‌کند تا بر روی مناطق مهم تمرکز کند.
  - پرسش و پاسخ بصری (VQA): توجه فضایی در مدل‌های VQA برای نمایان کردن مناطق خاص یک تصویر که با پاسخ به یک سوال داده شده مرتبط هستند، به کار می‌رود.
- به طور خلاصه، توجه فضایی یک مکانیسم ارزشمند در وظایف بینایی ماشین است که شامل پردازش داده‌های دارای ساختار فضایی مانند تصاویر است. این مکانیسم به مدل این امکان را می‌دهد تا بر روی مناطق مرتبط تمرکز کرده و توانایی درک و تفسیر اطلاعات بصری پیچیده را بهبود بخشد. با این حال، هزینه محاسباتی افزایش یافته و چالش‌های تفسیر باید در هنگام استفاده از توجه فضایی در برنامه‌های عملی در نظر گرفته شوند.

## پیوند به کد نمونه

<https://github.com/rwightman/pytorch-image-models> - (timm) Transformers Vision PyTorch  
این مخزن شامل انواع مدل‌های ترانسفورمر بصری است، و می‌توانید پیاده‌سازی‌های لایه‌های توجه فضایی را در آن پیدا کنید.

## خلاصه توجه به کانال (channel attention)

توجه به کانال یک مکانیسم رایج استفاده شده در شبکه‌های عصبی، به ویژه در شبکه‌های عصبی کانولوشنی (CNN) است که به صورت انتخابی بر روی کانال‌های مختلف نقشه‌های ویژگی ورودی تأکید یا آن را کم می‌کند. هدف از این مکانیسم، فعال کردن مدل بر روی کانال‌های مهم است که باعث افزایش نمایش و توان تمایزی ویژگی‌ها می‌شود.

## اجزای کلیدی توجه به کانال

- استخراج میانگین جهانی (GAP): توجه به کانال معمولاً شامل استخراج میانگین جهانی است، جایی که مقدار میانگین هر کانال در ابعاد فضایی محاسبه می‌شود.
- وزن‌های یادگیری‌شده: مقادیر میانگین‌گیری شده سپس از طریق یک لایه شبکه عصبی (معمولاً یک لایه کاملاً متصل یا یک MLP کوچک) می‌گذرند تا وزن‌های کانالی یادگیری شود.
- مقیاس‌بندی و ادغام: وزن‌های یادگیری شده برای مقیاس‌بندی کانال‌های ویژگی اصلی استفاده می‌شوند، یک جمع وزن‌دار ایجاد می‌شود. نتیجه با نقشه ویژگی اصلی ادغام می‌شود تا خروجی نهایی به دست آید.

## مزایای توجه به کانال

- تقویت انتخابی ویژگی: توجه به کانال به مدل این امکان را می‌دهد که به صورت انتخابی کانال‌های اطلاعاتی را تقویت و کاهش دهد، که توانایی یادگیری ویژگی‌ها را مؤثرتر می‌کند.
- بهبود عمومی: با تمرکز بر روی کانال‌های مهم، مدل ممکن است بهتر به داده‌های ورودی متنوع تعمیم بدهد و عملکرد را در تعدادی از وظایف بهبود بخشد.
- کارایی محاسباتی: توجه به کانال معمولاً مقایسه‌شده با مکانیسم‌های توجه به خود، کارایی محاسباتی بیشتری دارد و این امر مناسب برنامه‌های زمان‌واقعی می‌شود.

## چالش‌ها

- تفسیرپذیری: مانند سایر مکانیسم‌های توجه، تفسیر وزن‌های توجه به کانال یادگیری‌شده ممکن است چالش‌بر باشد و سختی تفهیم اینکه کدام کانال‌ها برای تصمیم مدل اساسی هستند را افزایش دهد.
- هزینه محاسباتی: هرچند که به طور کلی موثرتر از برخی از مکانیسم‌های دیگر توجه، توجه به کانال همچنان هزینه محاسباتی اضافی نسبت به لایه‌های کانولوشن استاندارد را معرفی می‌کند.

## موارد استفاده

- دسته‌بندی تصویر: توجه به کانال به طور معمول در وظایف دسته‌بندی تصویر به کار می‌رود تا به صورت خودکار کانال‌های مهم مرتبط با ویژگی‌های تمایزی یاد بگیرد و بر روی آنها تمرکز کند.
- شناسایی اشیاء: در مدل‌های شناسایی اشیاء، توجه به کانال به شبکه کمک می‌کند تا کانال‌هایی که برای شناسایی اشیاء خاص یا ویژگی‌های اشیاء مهم هستند را برجسته کند.
- تقسیم‌بندی معنایی: توجه به کانال می‌تواند برای وظایف تقسیم‌بندی معنایی مفید باشد که در آن درک و تأکید بر کانال‌های مهم به بهبود نتایج تقسیم‌بندی کمک می‌کند.
- تنظیم مدل‌های پیش‌آموزش داده‌شده: توجه به کانال مفید است زمانی که می‌خواهید مدل را بر روی وظایف خاص پیش‌آموزش داده‌شده تنظیم کنید، چرا که به مدل این امکان را می‌دهد که توجه خود را به ویژگی‌های خاص وظیفه تطبیق دهد.

به طور خلاصه، توجه به کانال یک مکانیسم ارزشمند برای تقویت نمایش ویژگی‌ها در شبکه‌های عصبی، به ویژه در وظایف بینایی ماشین است. این یک راه برای مدل فراهم می‌کند تا بر روی کانال‌های مهم تمرکز کند و توانایی گرفتن الگوهای معنادار و افزایش عملکرد را در برنامه‌های مختلف بهبود بخشد.

## پیوند به کد نمونه

<https://github.com/zhanghang1989/PyTorch-Encoding> - PyTorch-Encoding: این مخزن اجرایی و انعطاف‌پذیر ماژول‌های توجه، شامل توجه به کانال، ارائه می‌دهد.