The EfficientNet architecture needs to be described in detail and thoroughly explain all its concepts and innovations.

EfficientNet یک معماری شبکه عصبی کانولوشن و روش مقیاس بندی است که به طور یکنواخت همه ابعاد عمق یا عرض یا رزولوشن را با استفاده از یک ضریب ترکیبی مقیاس بندی می کند. بسیاری ان را قوی ترین معماری سی ان ان میدانند. بر خلاف روش مرسوم که به صورت دلخواه این عوامل را مقیاس بندی می کند، روش مقیاس بندی EfficientNet به طور یکنواخت عرض، عمق و رزولوشن شبکه را با مجموعه ای از ضرایب مقیاس بندی ثابت مقیاس بندی می کند. برای مثال، اگر بخواهیم 2^N برابر از منابع محاسباتی بیشتری استفاده کنیم، میتوانیم به سادگی عمق شبکه را با α^N ، عرض را با β^N و اندازه تصویر را با ضریب γ^N افزایش دهیم، به طوری که γ , β , α ضرایب ثابت هستند که با جستجوی شبکه در مدل کوچک اصلی تعیین میشوند.

EfficientNet از تکنیکی به نام ضریب مرکب استفاده می کند تا مدل ها را به روشی ساده اما مؤثر مقیاس کند. به جای افزایش تصادفی عرض، عمق یا وضوح، مقیاس بندی ترکیبی به طور یکنواخت هر بعد را با مجموعه ای ثابت از ضرایب مقیاس بندی مقیاس می کند. با استفاده از این روش توسعه دهنده های ان هفت مدل با ابعاد مختلف را توسعه دادند که از دقت پیشرفتهترین شبکههای عصبی کانولوشن و با کارایی بسیار بهتر پیشی گرفت.

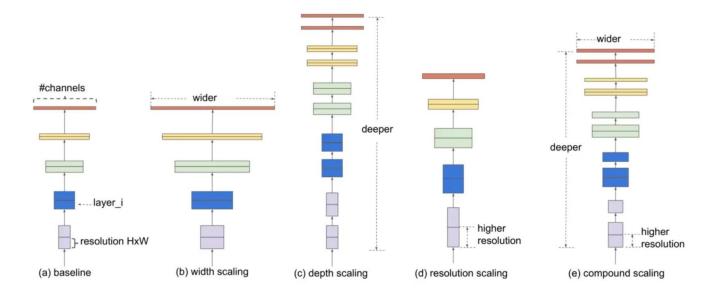
Depth
$$d = \alpha^{\emptyset}$$
, Width $w = \beta^{\emptyset}$, Resolution $r = \gamma^{\emptyset}$, (1)

such that α . β^2 . $\gamma^2 \approx 2$

$$\alpha \geq 1, \beta \geq 1, \gamma \geq 1$$

مقیاس بندی مرکب:

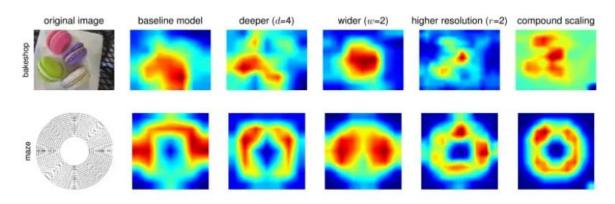
در این روش متخصصان دریافتند که در حالی که مقیاس بندی تک بعدی به بهبود عملکرد مدل کمک می کند، متعادل کردن مقیاس ها در هر سه بعد عرض، عمق و رزولوشن تصویر عملکرد کلی مدل را به بهترین نحو بهبود می بخشد.



مقادیر γ , γ با استفاده از الگوریتم جستجوی شبکه ای تعیین می شود. γ افزایش منابع محاسباتی شبکه را تعیین می کند. این پار امتر توسط کاربر تعریف شده است. اینها به عاملی محدود می شوند که نشان می دهد α . β^2 . γ^2 نزدیک به 2 میباشد. فلاپس (عملیات فلوتینگ پوینت در هر ثانیه) یک عملیات کانولوشن با α , α متناسب است، که نشان می دهد با دو بر ابر شدن عرض یا وضوح شبکه، فلاپس چهار بر ابر میشود. در سی ان ان عملیات عمق شبکه، فلاپس چهار برابر میشود. در سی ان ان عملیات های کانولوشنال هزینه های محاسباتی را کنترل می کنند. مقیاس کردن شبکه با استفاده از معادله بالا، فلاپس را با α . β^2 . γ^2 افز ایش میدهد. از انجایی که α . β^2 . γ^2 میشود، بنابر این برای هر γ جدید کل فلاپس با γ افز ایش میدهد.

اگر تصویر ورودی بزرگتر باشد، شبکه به لایه های بیشتری برای افزایش میدان دریافت و کانال های بیشتری نیاز دارد تا الگوهای بیشتری را روی تصویر بزرگتر ثبت کند. تکنیک مقیاسبندی ترکیبی همچنین به بهبود کارایی و دقت مدل های قبلی سی ان ان مثل مدل های MobileNet و ResNet به ترتیب با دقت های 1.4 درصد و 0.7 درصد در مقایسه با بقیه روش های مقیاس بندی تصادفی کمک میکند.

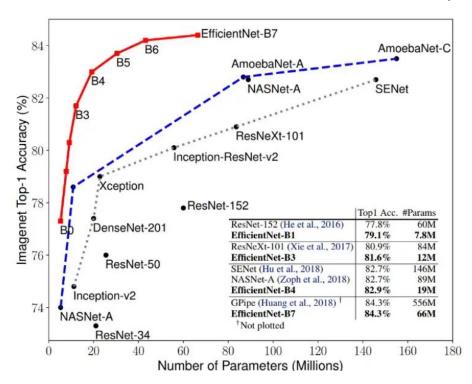
همان طور که در شکل زیر میبینید، این مدل نقشه های فعال سازی کلاس (CAM) بهتری را ارائه می دهد، که بیشتر بر روی مناطق مربوطه با جزئیات بیشتر تمرکز می کند و راه را برای توضیح بهتر مدل هموار می کند.



همان طور که در شکل پیداست روش مقیاس بندی مرکب به مدل مقیاس (ستون آخر) اجازه می دهد تا بر روی مناطق مرتبط تر با جزئیات بیشتر شی تمرکز کند.

عملکرد مدل:

عملکرد این مدل در مقایسه با مدل های دیگر بسیار خوب بوده است به طوری که بزرگ ترین مدل آن یعنی EfficientNet-B7 معلکردی پیشرفته در مجموعه داده های ImageaNet و ImageaNet بدست اورده است. در TOP - 3 و 97.3 درصد (TOP - 5) و TOP - 1) و 97.3 درصد (TOP - 5) و TOP - 3) و 1.5 درصد (TOP - 5) و 1.5 برابر سریع بدست تر از مدل سی آن آن قبلی بود. دقت 91.7% در مجموعه داده CIFAR-100 و دقت 98.8% در مجموعه داده 991.7 امده است.



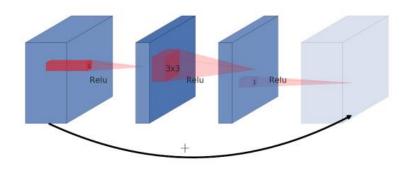
معماري مدل:

بلوک ساختمانی این معماری bottleneck معکوس متحرک MBConv است که بلوک باقیمانده معکوس با یک بلوک SE (فشر دگی و تحریک شدن) اضافی نیز نامیده می شود. این دو بلوک در زیر توضیح داده شده است.

:MBConv

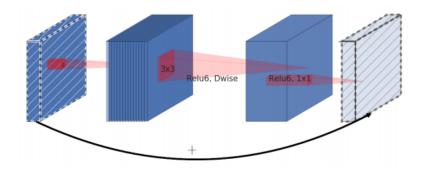
به منظور ایجاد جریان اطلاعات در شبکه های عصبی عمیق، معمولاً از بلوک های رزیجوال استفاده می کنیم. بلوک های رزیجوال ابتدا و انتهای یک بلوک کانولوشن شروع به عریض ابتدا و انتهای یک بلوک کانولوشن شروع به عریض شدن میکنند، سپس در امتداد عمق بلوک باریک تر میشوند تا انتها که به دلیل اطلاعات اضافه شده دوباره عریض میشوند. بنابراین الگوی یک بلوک باقیمانده معمولی از نظر تعداد کانال عریض > باریک > عریض است.

(a) Residual block



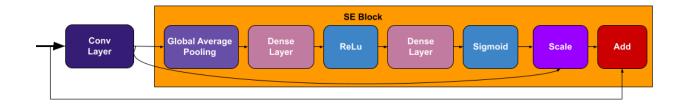
با این حال، بلوک رزیجوال معکوس از الگوی معکوس بلوکهای رزیجوال معمولی پیروی میکند، یعنی باریک > عریض > باریک میباشد.

(b) Inverted residual block



:SE

این بلوک های ساختمانی برای سی ان ان ها ها برای بهبود وابستگی متقابل بین کانال ها با انجام کالیبراسیون مجدد کانال های ویژگی پویا است، به این معنی که به جای وزن دهی یکسان همه کانال ها، شبکه به صورت پویا وزن بالایی را برای مهم ترین کانال ها اختصاص می دهد. شکل زیر اجزای این بلوک را نشان می دهد.



به عنوان نتیجه نهایی این مدل از تمام مدل های سی ان ان قبلی در اکثر مجموعه های محک بهتر عمل کند. مچنین میتوان از روش مقیاس بندی مؤثر سایر معماری های سی ان ان هم استفاده کرد. این اجازه میدهد مدل های EfficientNet به گونه ای مقیاس شوند که در ImageNet و سایر مجموعه داده های یادگیری ماشین که معمولا استفاده میشوند، به قدرت بیشتری با پارامتر ها و فلاپس های دست یابند.

The ResNext and the Inception-ResNet should be compared and discussed in terms of their advantages and disadvantages.

inception به هدف کاهش بار محاسباتی شبکه های عصبی عمیق و در عین حال به دست آوردن عملکرد پیشرفته ایجاد شده است. همانطور که شبکه عمیق تر می شود، بازده محاسباتی نیز کاهش می یابد، بنابراین توسعه دهندگان آن هنگام توسعه علاقه مند به یافتن راه حلی برای افز ایش مقیاس شبکه های عصبی بدون افز ایش هزینه محاسباتی بودند. میدانیم ResNet بر هزینه ی محاسباتی تمرکز دارد. به طور شهودی، شبکههای عمیق تر نباید بدتر از شبکههای کمعمق تر عمل کنند، اما در عمل، شبکههای عمیق تر بدتر از شبکههای کمعمق تر عمل میکنند، که ناشی از برازش بیش از حد نیست، بلکه به دلیل یک مشکل بهینه سازی است. به طور خلاصه، هر چه شبکه عمیق تر باشد، بهینه سازی شبکه سخت تر می شود.

مدلResNeXt در مقایسه با اینسپشن یک بعد جدید و کار دینالیته را به عنوان یک فاکتور اساسی علاوه بر ابعاد عمق و عرض نشان می دهد.

معماری ResNeXt توسعه ای از شبکه رزیجوال عمیق است که بلوک رزیجوال استاندارد را با شبکه ای جایگزین میکند که از استراتژی «تقسیم-تبدیل-ادغام» (یعنی مسیرهای شاخه دار در یک سلول) استفاده می شود. به سادگی، به جای انجام کانولوشن روی تمام فیچر مپ ورودی، ورودی بلوک به یک سری کانال پایین تر نمایش داده می شود که ما به طور جداگانه چند فیلتر کانولوشن را قبل از ادغام نتایج اعمال میکنیم.

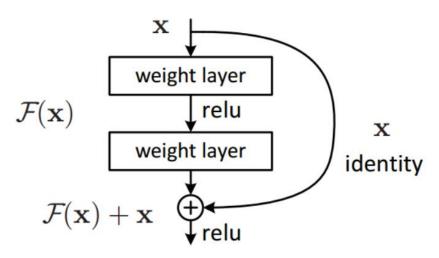
inception یک تکنیک ویژه به اسم Skip Connection دارد که راه فر اری بر ای مشکل نایدید شدن گرادیان ها میباشد. ایده اصلی پشت اجرای این طرح این است که به جای $\mathbf{x} \to \mathbf{F}(\mathbf{x})$ مپی به شکل $\mathbf{x} \to \mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x})$ داشته باشیم به طوری که:

ورودی است :X

خروجی است :F(x)

باقیمانده است :G(x)

وقتی ایکس همان اف ایکس باشد به این معنی که ورودی همان خروجی باشد، جی ایکس صفر خواهد شد. پس سی ان ان یاد میگیرد با صفر کردن وزن لایه های میانی، ورودی و خروجی را ترسیم کند.



علاوه بر این استفاده از تابع فعال سازی رلو نیز یک مزیت اضافی برای این شبکه برای حذف مشکل ناپدید شدن گرادیان ها میباشد .

این مدل عمدتا برای تشخیص ویژگی های مختلف یک تصویر طراحی شده است. این مدل می تواند از هسته های مختلفی برای تشخیص موجودیت های مختلف در تصویر بسته به اهمیت و مساحت آنها در کادر استفاده کند. از دیگر مزایای ان میتوان به این اشاره کرد که فضای شبکه ای را افزایش می دهد که بهترین شبکه از طریق آموزش انتخاب می شود. هر ماژول اولیه می تواند ویژگی های برجسته را در سطوح مختلف ثبت کند. از سوی دیگر این مدل به شدت مهندسی شده است و از ترفند های بسیاری برای بهبود عملکرد چه از نظر سرعت و چه از نظر دقت استفاده شده است. از ویژگی های مصبت ان تکامل مداوم ان خواهد بود.

مدل Inception - resnet به دلیل معماری و ساختار پیچیده ماژول های داخلی ان زمان اجرای ترین و تست تقریبا از مدل های دیگر بیشتر است.هم چنین در مقایسه با مدل موبایل نت و الکس نت هزینه محاسباتی نسبتاً سنگینی دارد.