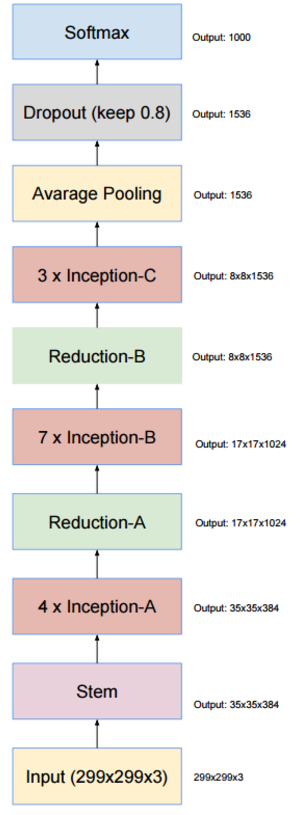
자료 출처 https://poddeeplearning.readthedocs.io/ko/latest/CNN/Inception.v4/



Inception v4 모델의 구조

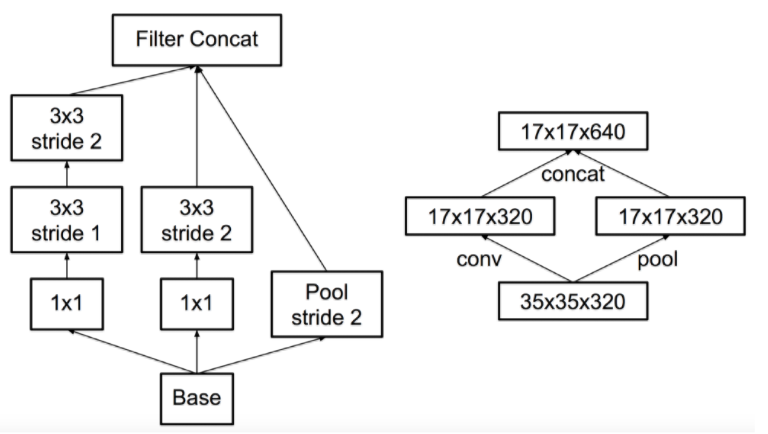
크게 ‘인셉션 모듈, 리덕션 모듈으로 나뉜다.

인셉션 모델의 중요한 두 개념(내 생각)> GoogLeNet 에서 inception으로 바뀌면서 가장 개선된 두 부분

커널의 인수분해(예를 들어 5x5 convolution(파라미터 25개)을 3x3 convolution을 두 번에 걸쳐 연산(파라미터18개)하는 방법으로 적은 파라미터로 같은 결과와 더 깊은 층을 형성하는 방법) 쇼지, 건물이(가) 표시된 사진

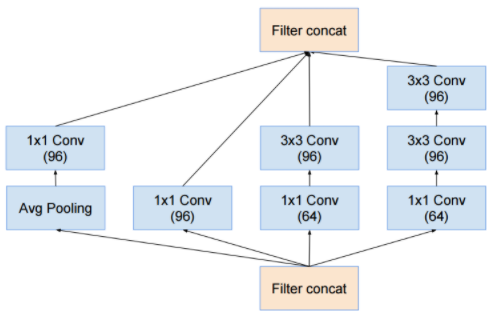
자동 생성된 설명

그리드(해상도)를 줄이기 위해 병렬 연산 구조를 도입한 모델



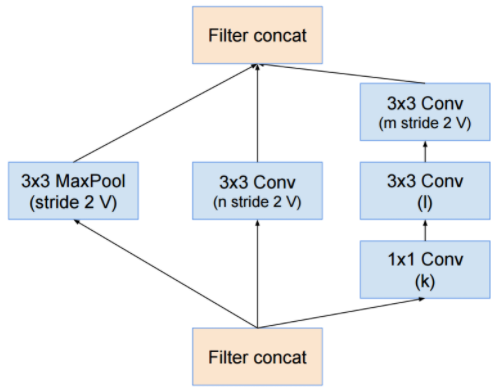
스트라이드의 컨볼루션과 함께 풀링 연산이 처리되고 있다. 두 가지 방식을 병렬로 적용하면 연산량이 줄어듬과 동시에 맵의 정보를 잃지 않을 수 있다.

인셉션 모듈

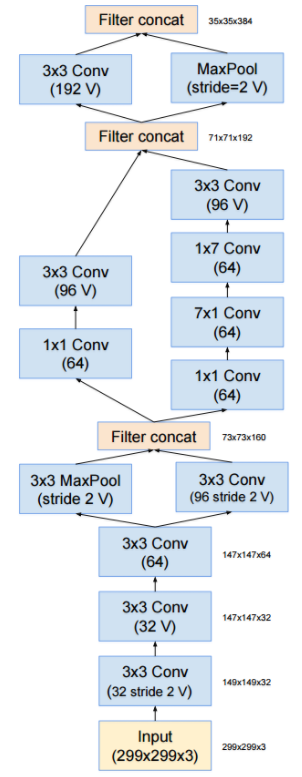


그리드는 줄어들지 않지만 커널의 인수분해 개념을 적용해 효율성이 올라갔다.

리덕션 모듈



병렬 연산이 적용된 모습이다. 이 리덕션 모듈을 통과하면서 그리드가 절반으로 줄어 든다.



Stem(전처리 영역)에도 병렬 연산을 적용하여 그리드 사이즈를 줄이는 것을 알 수 있다.