

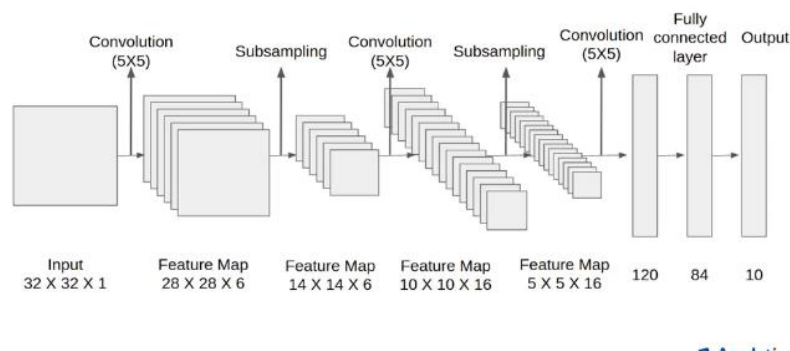
Họ và Tên: Đặng Anh Tiến – 20520800
Lớp: CS431.N11

Bài tập LT11 - Kể chuyện về quá trình phát triển của các mạng học sâu trên dữ liệu ảnh

Bài làm

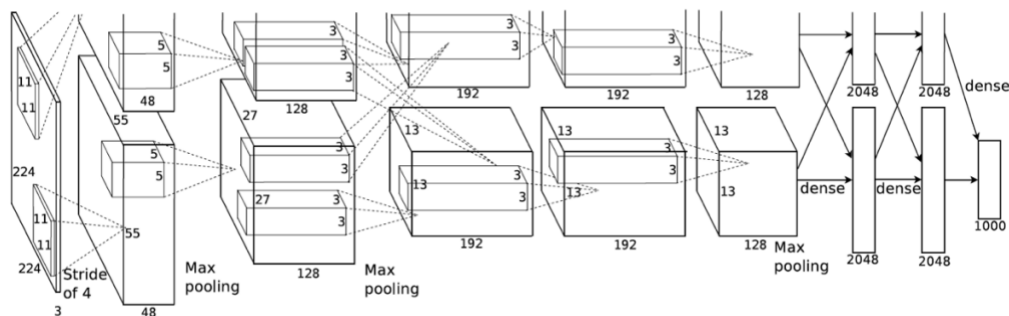
1. LeNet-5 (1998)

- LeNet-5 là kiến trúc đầu tiên sử dụng mạng tích chập.
- Model rất đơn giản, bao gồm 2 lớp Conv2D và 3 lớp Dense.
- Tuy đơn giản nhưng lại đem lại hiệu quả cao hơn các thuật toán Machine Learning khác ở thời điểm hiện tại (SVM, KNN,...) trên bộ dữ liệu MNIST – Hand Written.



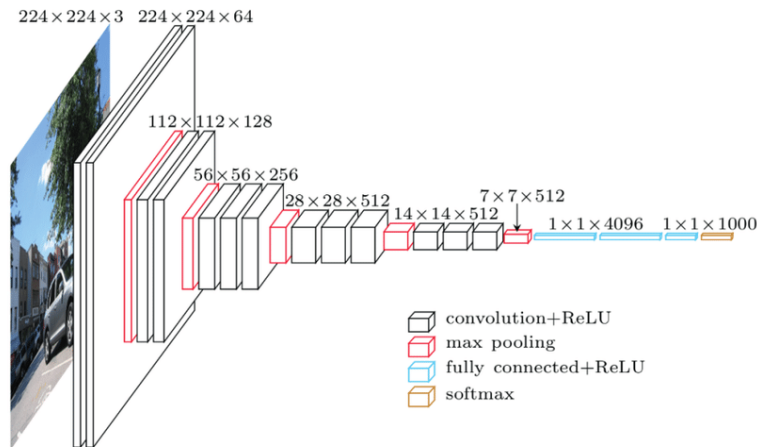
2. AlexNet (2012)

- Chiến thắng cuộc thi ImageNet năm 2012.
- Model có các cải tiến so với LeNet:
 - o Tăng kích thước Input và độ sâu của mạng.
 - o Sử dụng các filter có kích thước càng về sâu càng giảm dần.
 - o Sử dụng ReLU làm activation function.
 - o Sử dụng Dropout để hạn chế overfitting.
 - o Sử dụng GPU để train mô hình.
- Đánh dấu kỉ nguyên của Deep Learning. Làm nền tảng cho các mô hình học sâu sau này.



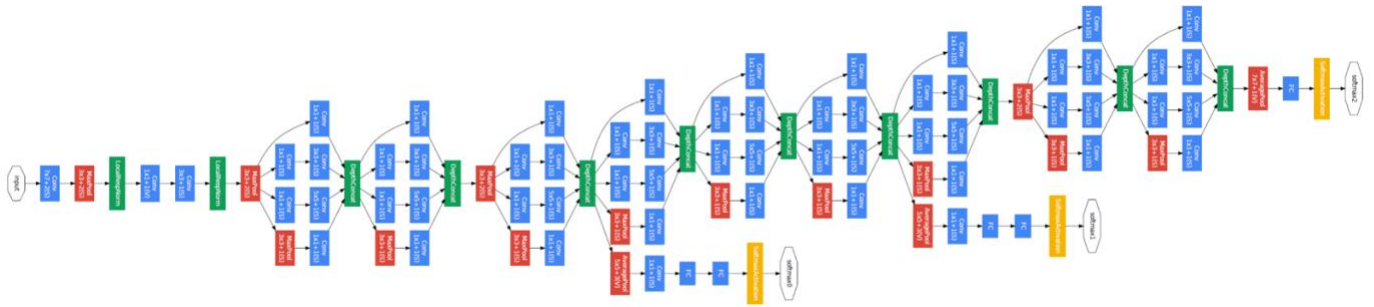
3. VGGNet16(2014)

- Các cải tiến:
 - o Mạng sâu hơn.
 - o Sử dụng các khối tích chập, xen kẽ nhiều Conv2D + MaxPooling.
 - o Chỉ sử dụng 3x3 filter.
- Cho thấy được mạng càng sâu, độ chính xác càng cao. Biến thể VGG19 mang lại độ chính xác cao hơn.



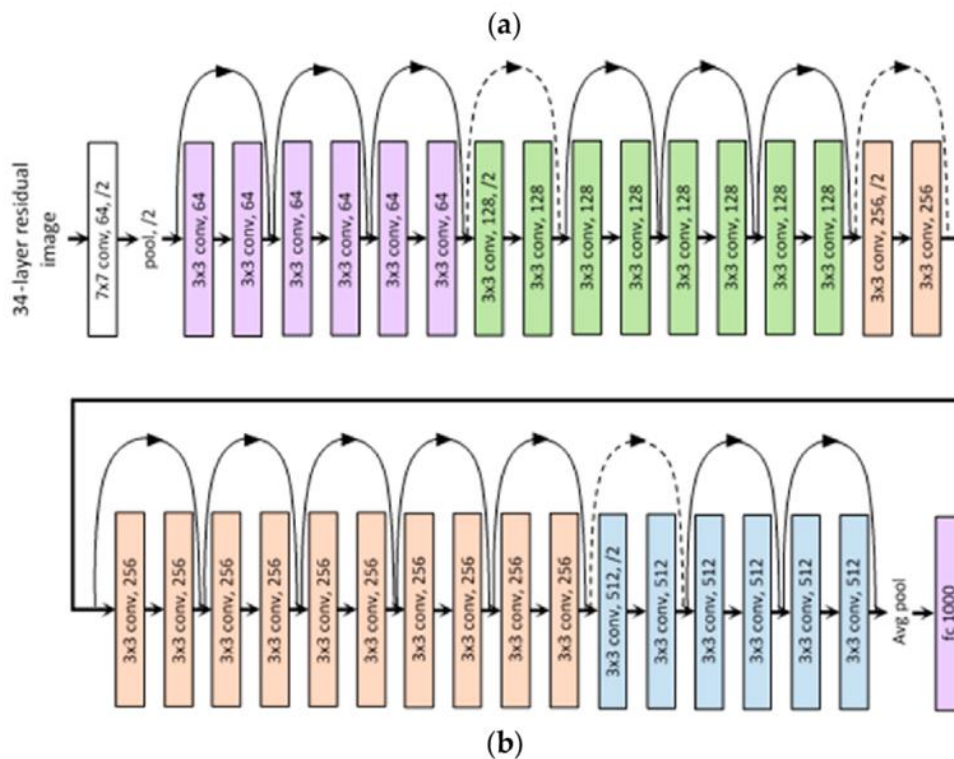
4. GoogleNet – InceptionNet (2014-2015)

- Chiến thắng cuộc thi ImageNet năm 2015.
- Giải quyết vấn đề chọn filter kích thước bao nhiêu là hợp lý. Cho thấy được việc kết hợp các feature từ đa dạng kích thước filter mang lại độ chính xác cao hơn.
- Các cải tiến:
 - o Các Inception block, gồm nhiều nhánh, mỗi nhánh áp dụng các filter có kích thước khác nhau nhằm trích xuất đặc trưng từ các receptive field có kích thước khác nhau.
 - o Các nhánh áp dụng padding và stride sao cho đầu ra có kích thước như nhau để concatenate các khối lại với nhau.
 - o Sử dụng 1×1 filter để giảm chiều sâu của các feature map. Giúp giảm số parameter xuống rất nhiều so với kiến trúc mạng trước.
- Các cải tiến sau của GoogleNet:
 - o Sử dụng thêm BatchNorm, chuẩn hoá feature ở mỗi batchsize giúp hội tụ hiệu quả hơn.
 - o Giải quyết bottle neck do giảm chiều feature đột ngột bằng cách tính toán kích thước filter hiệu quả hơn.



5. ResNet(2015)

- Là kiến trúc thường được sử dụng phổ biến nhất hiện nay.
- Mạng rất sâu, có số lượng layer lên tới 152. Nhưng nhờ các cải tiến nên vẫn đạt độ chính xác cao mà vẫn hạn chế được overfitting.
- Một trong những cách tăng độ chính xác cho kiến trúc mạng là gia tăng chiều sâu của mạng CNN. Tuy nhiên đến một độ sâu nào đó thì độ chính xác bị bão hoà, thậm chí kém đi. Khi kiến trúc mạng càng sâu, thông tin càng bị bão hoà. ResNet giải quyết bằng cách thêm các kết nối tắt (Skip connection).
- Các cải tiến:
 - o Skip connection giúp giữ thông tin không bị bão hoà bằng cách kết nối các layer trước với các layer sau mà bỏ qua các layer trung gian.
 - o Các khối này được gọi là các Identity block, layer trước kết nối tắt với layer sau bằng cách cộng lại với nhau.



6. DenseNet(2016)

- Ý tưởng của DenseNet cũng tương tự ResNet, sử dụng dày đặc các Skip Connection để liên kết các khối với nhau.
- Các cải tiến:
 - o Thay vì cộng các layer, DenseNet sẽ chồng chúng lại với nhau.
 - o Sử dụng các 1x1 filter để giảm số chiều.
 - o Tại mỗi layer nhận thêm các input từ tất cả các layer trước.
 - o Sử dụng BatchNorm để giúp hội tụ tốt hơn.
- DenseNet có ít tham số hơn ResNet nhưng lại có độ chính xác cao hơn trên bộ dữ liệu ImageNet

