# 18520245 Trần Tuấn Vĩ

**PHẦN 1: THÔNG TIN TÓM TẮT**

|  |  |
| --- | --- |
| **Tên đề tài (tiếng Việt)** | **PHÂN LOẠI ĐƯỜNG CÓ Ổ GÀ (POTHOLE CLASSIFICATION)-**[**https://github.com/Shindora/CS114.K11**](https://github.com/Shindora/CS114.K11) |
| **Họ và tên**  **Ảnh** | Trần Tuấn Vĩ |
| **Số buổi vắng** | 3 |
| **Số buổi đi trễ (không điểm danh)** | 2 |
| **Số lần Comment trên Google Classroom** | 13 |
| **Tóm tắt Bài tập Quá trình** | Bài tập python cơ bản,Linear Regression, Logistic Regression, GradientDescent (BatchGradientDescent, StochasticGradienDescent, Mini-batchGradientDescent),data collection |
| **Tóm tắt Bài tập Cuối kỳ** | Phân loại đường có ổ gà (Pothole Classification) hay không, sử dụng transfer learning, pre-trained network (ResNet-50) |

**PHẦN 2: BÁO CÁO CHI TIẾT ĐỒ ÁN CUỐI KỲ**

1. **XÁC ĐỊNH BÀI TOÁN**

**-**Input: Tập dataset gồm đường có ổ gà và không có ổ gà

-Output: Đường có ổ gà hay không

-Pretrain model: ResNet-50

1. **THU THẬP DATA**

**-** Data phần chủ yếu 85% là lấy từ các cuộc thi trên Kaggle,15% là thực tế.

- 722 samples gồm 367 ảnh đường bình thường (Plain) và 357 ảnh đường ổ gà (pothole)

1. **RESNET-50**
2. **ResNet-50 là gì?**

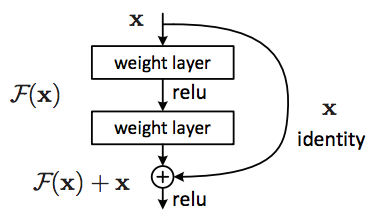
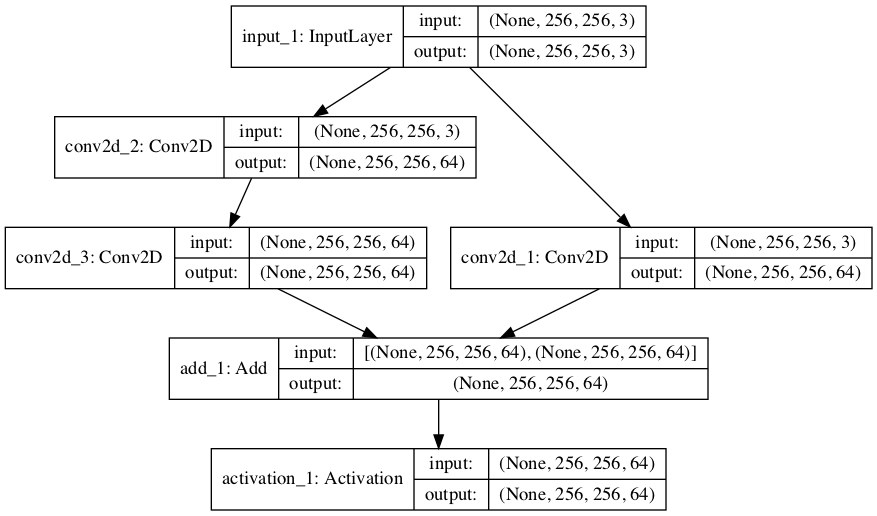
* ResNet-50 là một pre-trained network được dùng cho phân loại ảnh (Image-classification). ResNet-50 có 50 layers, đã được train hơn 1 triệu ảnh với 1000 loại (categories) từ dữ liệu của ImageNet.
* ResNet-50 cho hiệu suất tốt, tỉ lệ thấp đối với bài toán nhận dạng.

1. **Residual block**

Thực tế cho thấy khi layer trong CNN tăng thì performance thường tốt hơn, nhưng khi layers >50 thì performance lại giảm đi.

Residual block có thể giúp train các CNN phức tạp mà không sợ bị exploding gradient.

Cứ sau 2 layers, cộng input với output: F(x)+x

ResNet là một mạng CNN được tạo thành từ nhiều Residual block.

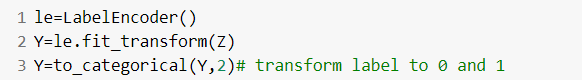
*Hình. Convolutional Neural Network Architecture With an Residual Module*

1. **CÁC BƯỚC TIẾN HÀNH**

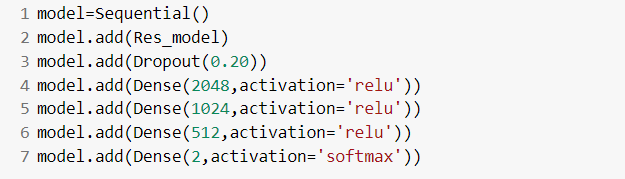
**Xử lý dữ liệu**

Resize tất cả tấm ảnh về cùng một kích thước là 256x256x3

**Label Encoder**

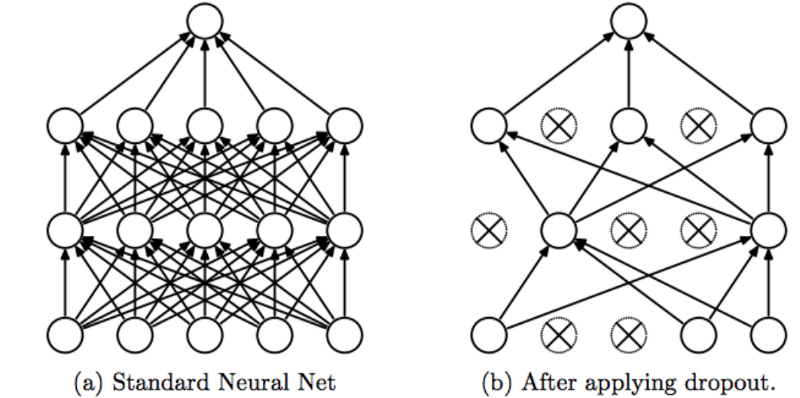
Chuyển đổi nhãn plain->0 và pothole ->1

**Tạo model**

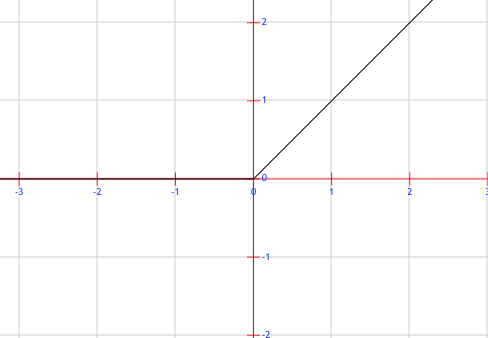
Ngoài ResNet, ta thêm một số layers để phù hợp với mục tiêu output

1. **Dropout**

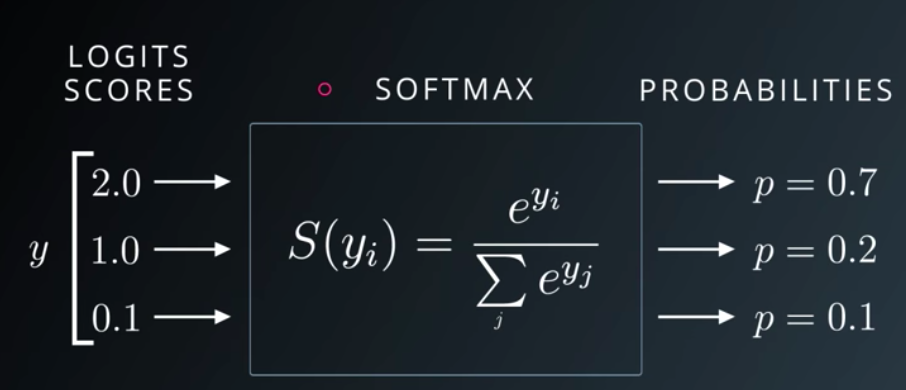
Khi sử dụng full connected layer, các neural sẽ vào trạng thái “phụ thuộc mạnh” lẫn nhau, điều này khiến cho tập train bị over-fitting.

Dropout sẽ bỏ qua một số unit (được chọn ngẫu nhiên dựa trên xác suất).

1. **ReLU**

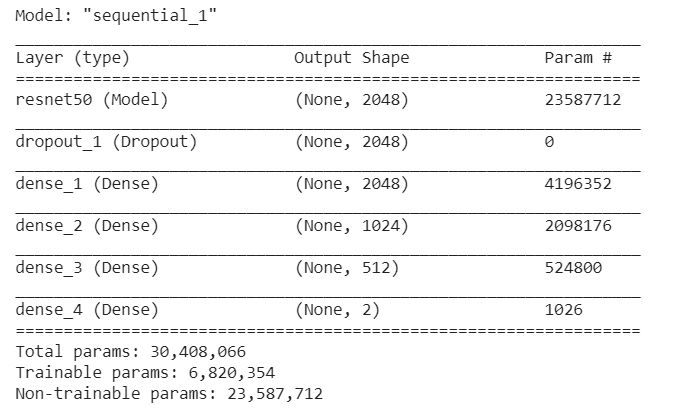
Hàm ReLU được sử dụng phổ biến vì tính đơn giản với hàm toán học, lọc các giá trị <0. Có khả năng tăng tốc tức thời trong gradient, chi phí tính toán ít phức tạp hơn so với sigmoid và tanh (sử dụng exp)

1. **Softmax**

Softmax thường xuất hiện ở layer cuối cùng trong bài toán classification. Nó cho ra phân phối xác suất của list label.

*Hình. Udacity Deep Learning Slide on Softmax*

**Compile model**

1. **Adam**

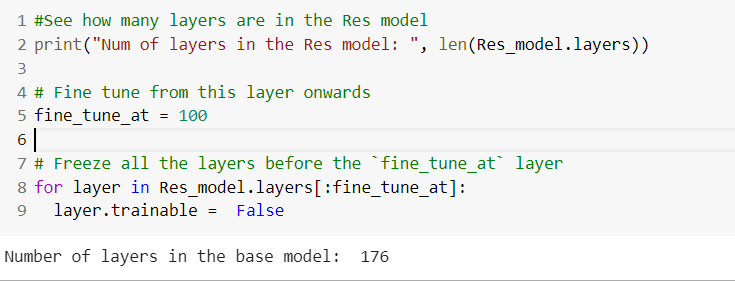
Adam là một thuật toán tối ưu , đặc biệt trong Deep Learning. Có thế xem nó là sự kết hợp giữa RMSprop và Stochastic Gradient Descent (SGD)

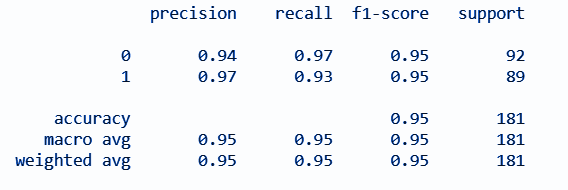
1. **Cross Entropy**

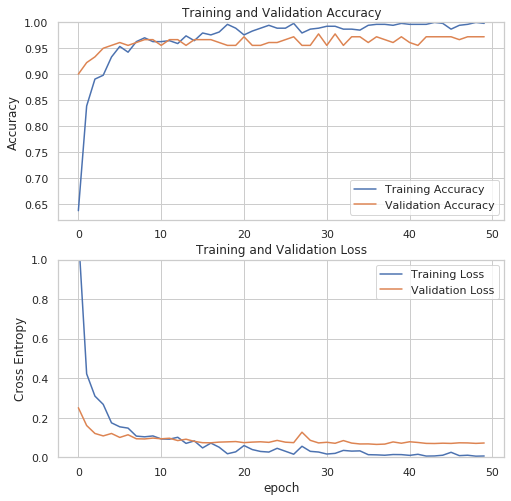
y là true value, là predicted value, N là số class

Cross entropy dùng để tính toán độ chênh lệch giữa 2 phân phối xác suất predicted và của nhãn. Nhãn được transform thành một vector one-hot , chỉ có 1 vị trí là 1, còn lại là 0. Output là một phân phối xác suất (p1,p2,..pn). Xác suất của vị trí 1 trong vector tương ứng là pi.

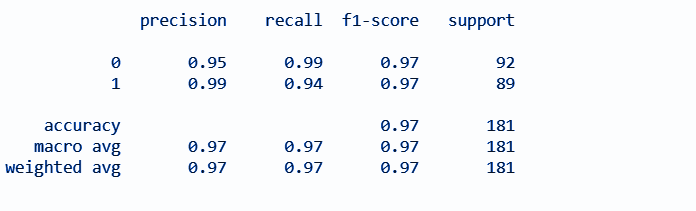
**Fine tune**

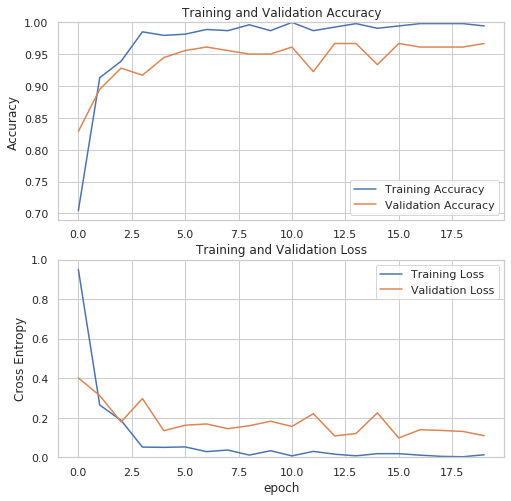
“*Quan sát rằng những layers đầu trong ConvNet thường giúp extract những đặc tính chung của ảnh (các cạnh - edges, còn được gọi là low-level features), các layers cuối thường mang những đặc trưng riêng của cơ sở dữ liệu (CSDL) (và được gọi là high-level features). Vì vậy, việc train các layer cuối mang nhiều giá trị hơn*”\_ Machine learning cơ bản

1. **KẾT QUẢ**
2. **Non-fine tune**



1. **Fine tune**



Theo kết quả có thể thấy, sau khi fine tune cho performance tốt hơn trên tập test (97% so với 95%), đổi lại performance trên tập train lại thấp đi (99,82% so với 100%). Điều này cho thấy dataset có độ dạng tốt, phủ được nhiều trường hợp.

**VI.HƯỚNG PHÁT TRIỂN**

Ta có thể phát triển mô hình lên Classification Pothole real time, với dataset là video được quay từ các camera an ninh, nhằm phát hiện đường có vấn đề và có biện pháp xử lý kịp thời, tránh tai nạn giao thông.

TÀI LIỆU THAM KHẢO

1. Cross-Entropy: [Deep Learning: Phương pháp kết hợp các hàm Loss và Activation](https://ereka.vn/post/deep-learning-phuong-phap-ket-hop-cac-ham-loss-va-activation-5298271585614996324)
2. Softmax-function: [Understand the Softmax Function in Minutes - Data Science Bootcamp](https://medium.com/data-science-bootcamp/understand-the-softmax-function-in-minutes-f3a59641e86d)
3. ReLU:[Các hàm kích hoạt (activation function) trong neural network](https://blog.vietanhdev.com/posts/2019-09-23-cac-ham-kich-hoat-activation-function-trong-neural-networks/)
4. Dropout: [Bài 10: Các kỹ thuật cơ bản trong deep learning](https://nttuan8.com/bai-10-cac-ky-thuat-co-ban-trong-deep-learning/)
5. Resnet : [How to Develop VGG, Inception and ResNet Modules from Scratch in Keras](https://machinelearningmastery.com/how-to-implement-major-architecture-innovations-for-convolutional-neural-networks/)