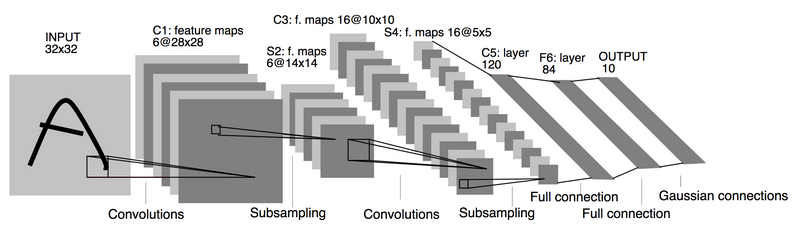
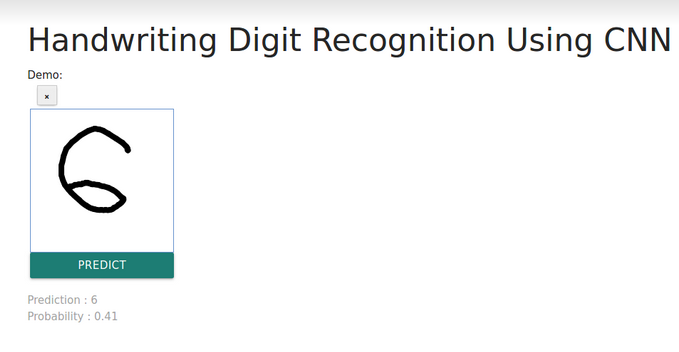
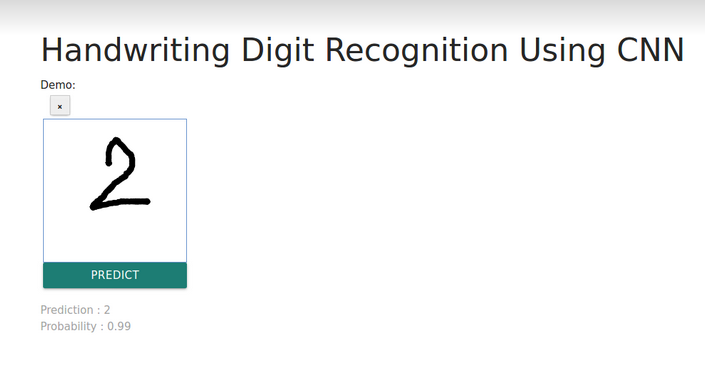
|  |  |
| --- | --- |
| **Author** | gary30404 |
| **Team** | Nguyễn Trần Quốc Thái (qthainguyentran@gmail.com) |
| **Title** | Handwritten Digit Recognition Using Convolutional Neural Network by Python |
| **Applicable** | Ứng dụng trong computer vision, sử dụng thuật toán chính là CNN |
| **Descriptions** | Input sẽ là các ảnh số được viết bằng tay. Sử dụng thuật toán chính là CNN để train, khi train xong sẽ trả ra output là file trọng số weights. Ta sẽ sử dụng trọng số weights đã train để classification số trong hình |
| **Links** | <https://github.com/gary30404/convolutional-neural-network-from-scratch-python> |
| **Models** | Sử dụng mô hình LeNet – 5 tạo ra bởi Yann Lecun năm 1998, để nhận diện chữ viết tay. |
| **Features** | Số lượng tham số và connection đều là: 84 \* (120 + 1) = 10164 |
| **Datasets** | Mô hình được train với bộ dữ liệu [MNIST](http://yann.lecun.com/exdb/mnist/) là bộ chữ số viết tay gồm 60k ảnh số được viết tay khác nhau và 10k ảnh được dùng để test. |
| **Level of difficulty** | Mô hình đơn giản sử dụng nhanh và dễ, có thể thay đổi được model không theo [LENET](http://yann.lecun.com/exdb/lenet/)-5 |

* **Mô hình TEFPA của bài toán**
* Task: phân biệt được ảnh đầu vào là các chữ số viết tay.
* Experience: bộ dữ liệu chữ số viết tay [MNIST](http://yann.lecun.com/exdb/mnist/)
* Performance: Cross Entropy
* Function: dùng các trọng số đã được train. (pre-trained)
  + Model details

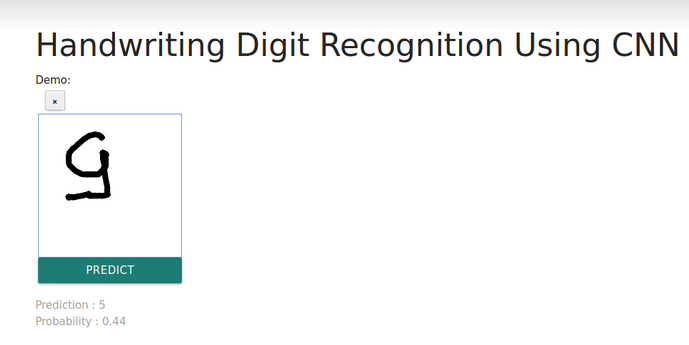


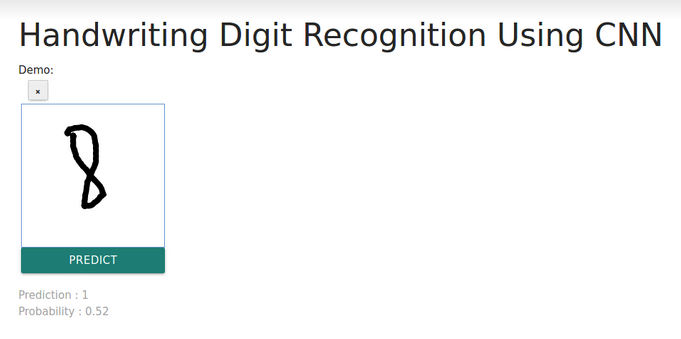
* Input: Ảnh trắng đen (32x32)
* Convolution level 1 (C1): sử dụng 6 kenels (5x5x1), stride = 1 ((6 \* 5 \* 5 \* 1 = 150 weights và 1 \* 6 = 6 bias), activation function là ReLu, padding =2.
* Subsampling (S2): Sử dụng maxpolling với pool\_size = 2 và stride = 2
* Convolution level 3 (C3): sử dụng 16 kenels (5x5x1), stride = 1, padding = 0, activation là ReLu.
* Subsampling (S4): Sử dụng maxpolling với pool\_size = 2 và stride = 2
* Convolution level 5 (C5): sử dụng 120 kenels (5x5x1), stride = 1, padding = 0, activation là ReLu.
* Full Connection (F6): Fully-connected layer, với mỗi unit đều được connect với tất cả 120 unit của C5.
* Output 10: Output layer có 10 unit đại diện cho 10 chữ viết tay
* Algorithm: Sử dụng thuật toán chính là CNN.
* **Body**
* Kết quả:

****

****

Một số kết quả lỗi:





* **Conclusion**
* Phát triển mô hình dự đoán chính xác hơn, và nhận dạng chữ viết tay
* Ứng dụng: có thể sử dụng để scan các văn bản viết tay (thư từ, đơn,…) nhận dạng chữ kí.