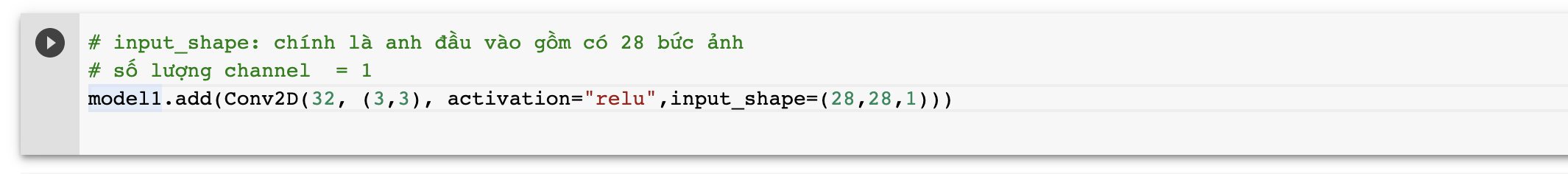
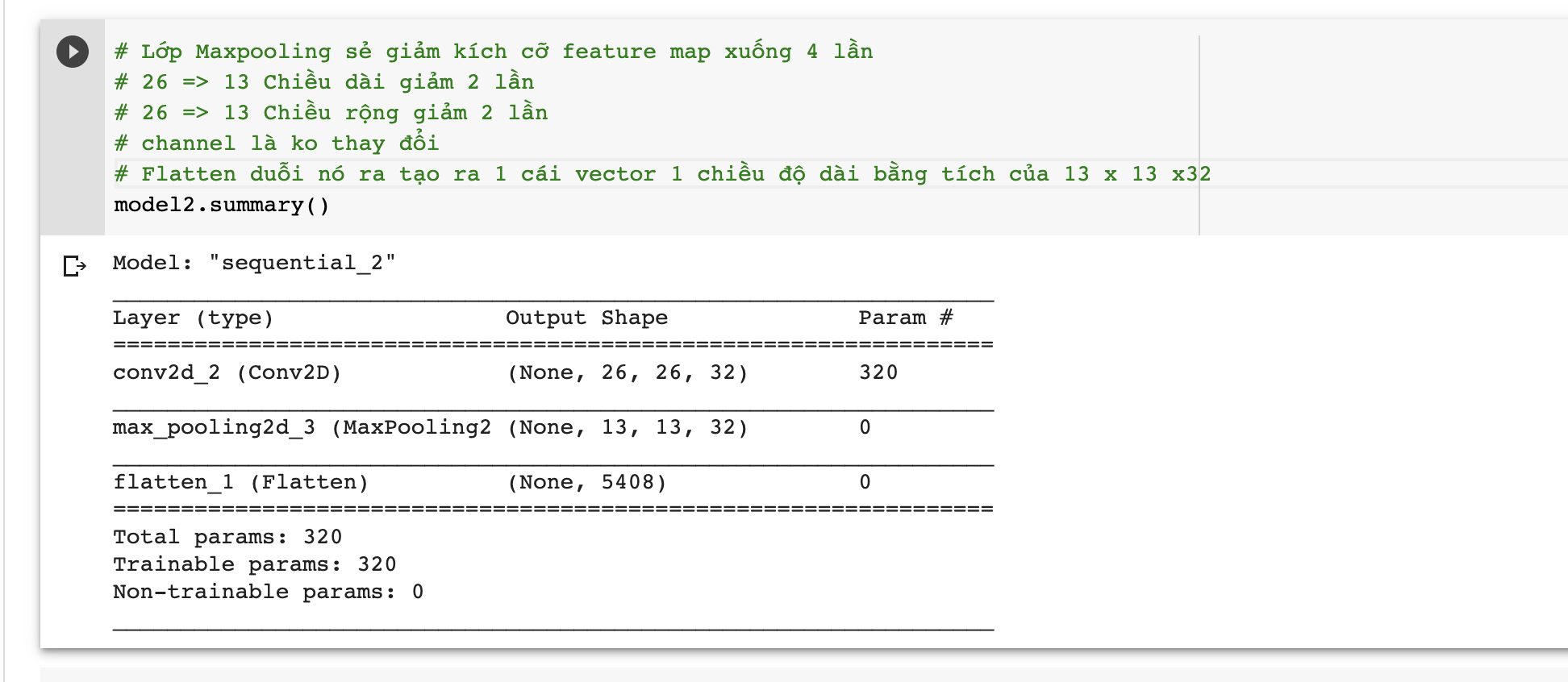
Cho một bộ data MNIST bao gồm 60000 ngàn ảnh để huấn luyện 28 \* 28 chữ số viết tay. Tiến hành xây dựng mạng này bao gồm

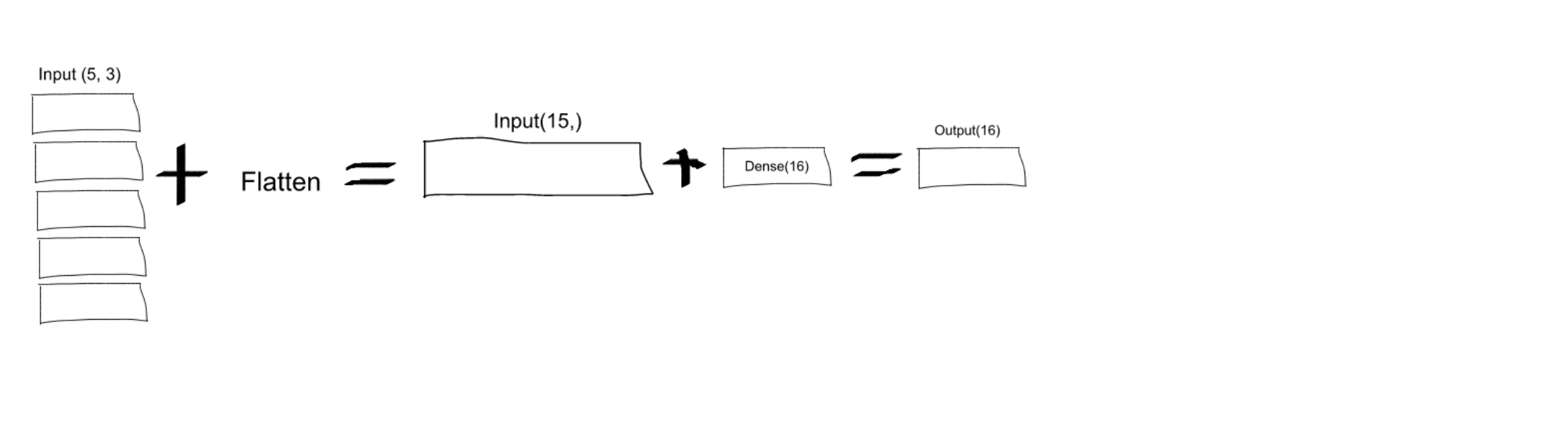
* 1 lớp Conv cỡ 3 \* 3 bao gồm 32 kernels
* 1 lớp Max Pooling cỡ 2 \* 2
* 1 lớp duỗi Flatten
* 1 lớp Dense 128
* 1 lớp Dense 10



Relu là một hàm phi tuyến, y = max(0, x), nếu đầu vào dương nó sẽ trả về chính nó, đầu vào âm thì sẽ trả về 0.



* Do chỉ là ảnh đen trắng nên chỉ có 1 channel => param = (3 \*3 +1) \*32 = 320 params



* Merge layer và Flatten layer khác nhau
* - [Merge](https://forum.vietdesigner.net/tags/merge/) visible: gộp các Layers được chọn lại thành 1 layer, còn Layer nào bị ẩn thì vẫn được giữ nguyên, không bị gộp
* [Flatten](https://forum.vietdesigner.net/tags/flatten/) image: gộp các Layers lại thành 1 cái Background chứ không phải thành 1 Layer, kể cả Layer ẩn cũng bị gộp

Theo như CNN (Convolutional Neural Network) layer là 28x28x32, sao kiểm tra model thì là 26x26x32 ?

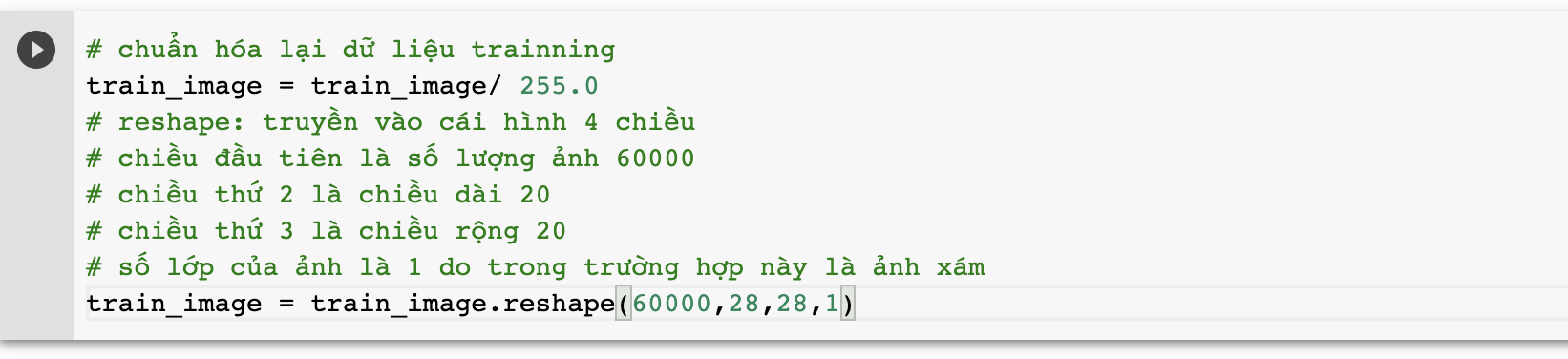
* Vì trượt trên ảnh nên ảnh đầu ra sẽ nhỏ hơn ảnh hiện tại nhé em trường hợp này trượt cách nhau 1 nên mất 1 chút cạnh nên còn 26.

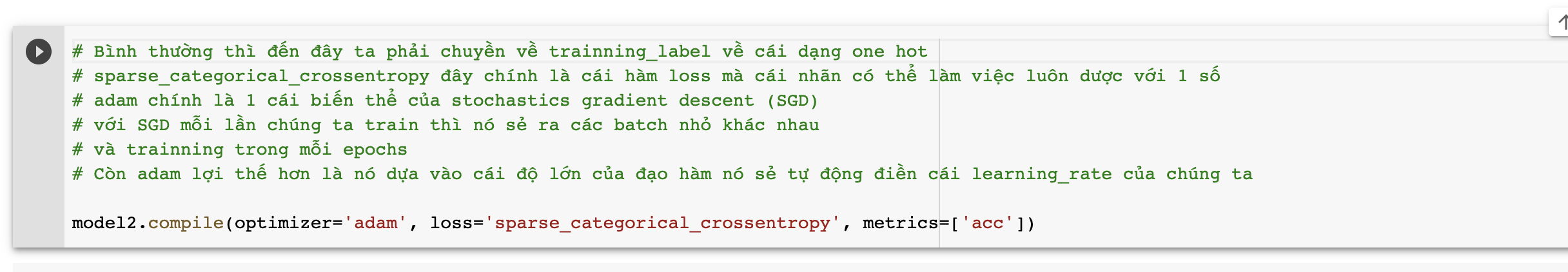
Công thức nào tính ra được số lượng params cần học phù hợp với input bài toán?

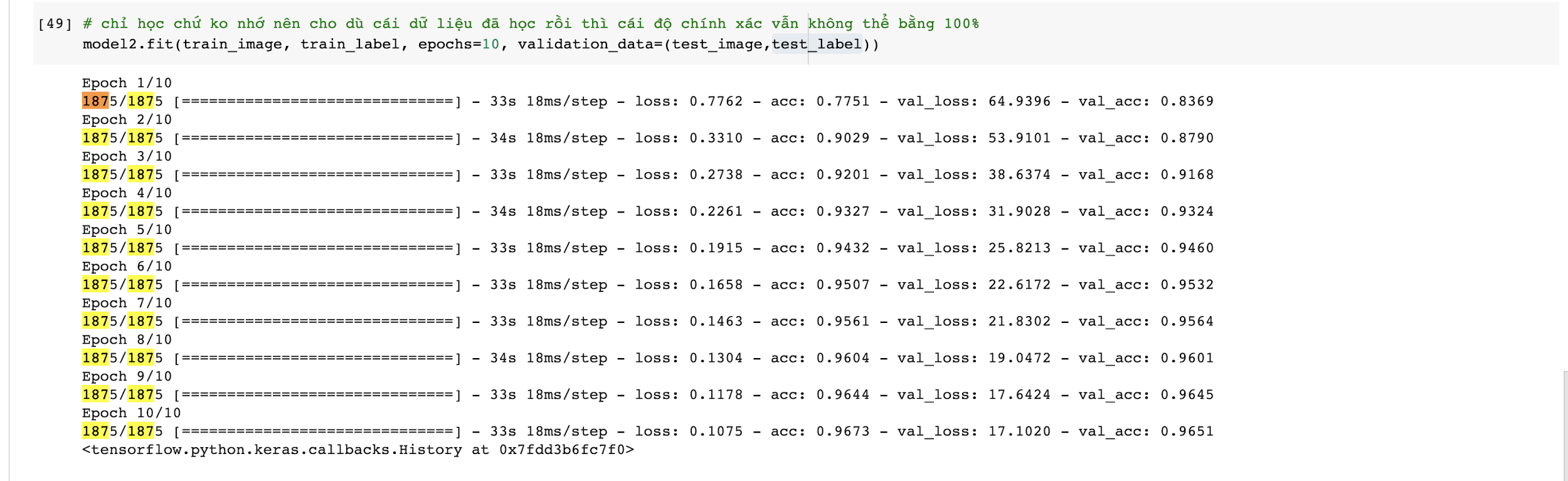
* CNN thì số kernels \* width \* height là số params cần học tại một lớp params
* Fully connected là tích chiều giữa 2 lớp ví dụ em muốn giảm 1024 xuống 10 thì em cần 1024 \* 10 params

Để nhận dạng nhiều chữ trong hình thì liên quan đến R-CNN (regions with CNN features)

* BatchNorm layer vì ít lớp nên ko cần chuẩn hóa lại dữ liệu







* sparse\_categorical\_crossentropy thì dành để tiết kiệm thời gian không cần chuyển thành one hot
* Adam fine tunning cho mình cái learning rate dựa trên đạo hàm, các thông số của Adam cũng mang tính chất thực nghiệm.
* Nếu để mà nói là nên dùng cái gì trường hợp nào sẽ rất khó với SGD và Adam, trong nhiều trường hợp phải thử tất cả các optimizers,
* Mỗi một bài dùng một loại riêng nên đang có hướng nghiên cứu sẽ tạo ra mô hình khác tự học được fine tunning nhưng mà quãng đường này còn rất xa.
* 1875/1875 là steps trong một epoch em nha
* Tuần sau tìm hiểu
* Tìm hiểu CNN
* Tìm hiểu về Maxpooling
* Cách chuyển 1 hình ảnh thành cái dataset thành cái ma trận nhiều chiều