****

**ĐẠI HỌC ĐÀ NẴNG**

**KHOA CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

****

**THỊ GIÁC MÁY TÍNH**

ĐỀ TÀI: XÂY DỰNG CHƯƠNG TRÌNH

Phân Loại Hình Ảnh

**Giáo viên hướng dẫn : TS Lê Thị Thu Nga**

**Sinh viên thực hiện : Phạm Văn Định -18IT305**

**: Nguyễn Văn Nam -18IT340**

**Lớp : 18IT5**

**Khóa : 2018 – 2023**

***Đà nẵng, tháng 12 năm 2021***

**NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN**

**………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………**

**Chữ kí của giảng viên hướng dẫn**

**LỜI CẢM ƠN**

Trân trọng gửi lời cảm ơn thầy TS Lê Thị Thu Nga đã tạo điều kiện giúp đỡ cho chúng em có cơ hội được làm đồ án.

Trong học kì vừa qua, nhờ sự chỉ dẫn nhiệt tình của thầy, chúng em đã tiếp thu được những kiến thức cơ bản quan trọng để có thể làm được một ứng dụng hoàn chỉnh.

Cũng xin cảm ơn thầy cô trong khoa Khoa Học Máy Tính đã nhiệt tình hỗ trợ, tạo điều kiện nhóm chúng em làm bài báo cáo này.

# 

# MỤC LỤC

[**MỤC LỤC 5**](#_Toc73047552)

[**MỞ ĐẦU 7**](#_Toc73047553)

[**1 Lý do chọn đề tài 7**](#_Toc73047554)

[**2 Mục đích 7**](#_Toc73047555)

[**3 Đối tượng nghiên cứu 7**](#_Toc73047556)

[**4 Phạm vi nghiên cứu 7**](#_Toc73047557)

[**CHƯƠNG I. CƠ SỞ LÝ THUYẾT 8**](#_Toc73047559)

[**1.1 Python 8**](#_Toc73047560)

[**1.2 Keras 8**](#_Toc73047561)

[**1.3 Tensorflow 9**](#_Toc73047562)

[**1.4 Bộ dữ liệu CIFAR-10 10**](#_Toc73047563)

[**1.5 Deep learning 11**](#_Toc73047564)

[**1.6 Thuật toán CNN 11**](#_Toc73047565)

[**1.7 Tkinter 12**](#_Toc73047566)

[**1.8 Numpy 12**](#_Toc73047567)

[**CHƯƠNG 2. PHÁT TRIỂN BÀI TOÁN 13**](#_Toc73047568)

[**2.1 Mô tả bài toán: 13**](#_Toc73047569)

[**2.2 Các chức năng chính 13**](#_Toc73047570)

[**2.3 Sơ đồ hoạt động 13**](#_Toc73047571)

[**CHƯƠNG 3. KẾT QUẢ THỰC HIỆN 14**](#_Toc73047572)

[**3.1 Kết quả: 14**](#_Toc73047573)

[**3.2. Giới thiệu các chức năng chính 14**](#_Toc73047574)

[**3.3 Các đoạn mã để viết lên chương trình 21**](#_Toc73047575)

[**KẾT LUẬN 24**](#_Toc73047576)

[**1.Kết quả thực hiện được: 24**](#_Toc73047577)

[**2.Ưu điểm 24**](#_Toc73047578)

[**4. Hướng phát triển 24**](#_Toc73047579)

[**Tài liệu tham khảo 25**](#_Toc73047580)

# MỞ ĐẦU

1. **Lý do chọn đề tài**

Một cách thực tế, cho dù kiểm tra biểu đồ đầu tư danh mục đầu tư chứng khoán hoặc nhìn vào sự lan tỏa của một trận bóng đá sắp tới, hoặc đơn giản là tham gia vào nghệ thuật, chúng ta liên tục phân tích nội dung trực quan, diễn giải ý nghĩa và lưu trữ các kiến ​​thức để sử dụng sau này.

Tuy nhiên, đối với máy tính, việc diễn giải nội dung của hình ảnh khó hơn - tất cả máy tính thấy là một ma trận lớn của các con số. Nó không có bất kì ý tưởng nào liên quan đến những suy nghĩ, kiến ​​thức, hoặc ý nghĩa hình ảnh đang cố gắng truyền tải.

Để hiểu nội dung của hình ảnh, chúng ta phải áp dụng [phân loại hình ảnh](https://uniduc.com/vi/blog/phan-loai-anh-co-ban), trong đó  
là nhiệm vụ của việc sử dụng thị giác máy tính và thuật toán học máy để rút ra ý nghĩa từ một tấm ảnh. Hành động này có thể đơn giản như gán nhãn cho những gì hình ảnh hoặc như nâng cao như diễn giải nội dung của một hình ảnh và trả về một câu có thể đọc được.

1. **Mục đích**

Phân loại hình ảnh là một lĩnh vực nghiên cứu rất lớn, bao gồm rất nhiều kỹ thuật-và với sự phổ biến của học tập sâu, nó đang tiếp tục phát triển.

Bây giờ là lúc để tìm hiểu học sâu và phân loại hình ảnh. Phân loại hình ảnh và hiểu hình ảnh hiện đang là lĩnh vực phụ phổ biến của Computer Vision trong mười năm tới. Chúng ta sẽ thấy ngày càng nhiều ứng dụng tiêu dùng trên điện thoại thông minh có thể hiểu được và giải thích nội dung của một hình ảnh.

1. **Đối tượng nghiên cứu**

* Thuật toán CNN (Convolutional neural Network)
* Keras
* Deep learning
* Tensorflow
* Bộ dữ liệu CIFAR-10
* Tkinter
* Numpy
* Python

1. **Phạm vi nghiên cứu**

Trong đồ án này, nhóm em hướng đến nghiên cứu và sử dụng các dữ liệu trong bộ dữ liệu CIFAR-10 của thư viện keras.

# CHƯƠNG I. CƠ SỞ LÝ THUYẾT

* 1. **Python**

[Python](https://quantrimang.com/python) là ngôn ngữ lập trình hướng đối tượng, cấp cao, mạnh mẽ, được tạo ra bởi Guido van Rossum. Nó dễ dàng để tìm hiểu và đang nổi lên như một trong những ngôn ngữ lập trình nhập môn tốt nhất cho người lần đầu tiếp xúc với ngôn ngữ lập trình.

Python hoàn toàn tạo kiểu động và sử dụng cơ chế cấp phát bộ nhớ tự động.

Python có [cấu trúc dữ liệu](https://quantrimang.com/cau-truc-du-lieu-va-giai-thuat" \o "Cấu trúc dữ liệu và giải thuật) cấp cao mạnh mẽ và cách tiếp cận đơn giản nhưng hiệu quả đối với lập trình hướng đối tượng. Cú pháp lệnh của Python là điểm cộng vô cùng lớn vì sự rõ ràng, dễ hiểu và cách gõ linh động làm cho nó nhanh chóng trở thành một ngôn ngữ lý tưởng để viết script và phát triển ứng dụng trong nhiều lĩnh vực, ở hầu hết các nền tảng.

* 1. **Keras**

Keras là một thư viện mạng nơ ron được viết bằng python năm 2015 bởi 1 kỹ sư deep learning của google. Ta có thể kết hợp keras với các thư viện deep learning.

Ưu điểm nổi bật của keras so với các thư viện deep learning khác có lẽ là cú pháp dễ sử dụng và trực quan của nó. Ngoài ra nó còn có thể chạy trên cả CPU và GPU.

Keras được coi là một thư viện ‘high-level’ với phần ‘low-level’ (còn được gọi là backend) có thể là TensorFlow, CNTK, hoặc Theano. Keras có cú pháp đơn giản hơn TensorFlow rất nhiều. Với mục đích giới thiệu về các mô hình nhiều hơn là các sử dụng các thư viện deep learning, tôi sẽ chọn Keras với TensorFlow là ‘backend’.

Những lý do nên sử dụng Keras để bắt đầu:

Keras ưu tiên trải nghiệm của người lập trình

Keras đã được sử dụng rộng rãi trong doanh nghiệp và cộng đồng nghiên cứu

Keras giúp dễ dàng biến các thiết kế thành sản phẩm

Keras hỗ trợ huấn luyện trên nhiều GPU phân tán

Keras hỗ trợ đa backend engines và không giới hạn bạn vào một hệ sinh thái

* 1. **Tensorflow**

[TensorFlow](https://www.tensorflow.org/) chính là thư viện mã nguồn mở cho machine learning nổi tiếng nhất thế giới, được phát triển bởi các nhà nghiên cứu từ Google. Việc hỗ trợ mạnh mẽ các phép toán học để tính toán trong machine learning và deep learning đã giúp việc tiếp cận các bài toán trở nên đơn giản, nhanh chóng và tiện lợi hơn nhiều.

**1.3.1** **Kiến trúc của tensorflow**

Kiến trúc TensorFlow hoạt động được chia thành 3 phần:  
– Tiền xử lý dữ liệu  
– Dựng model  
– Train và ước tính model

**1.3.2 Cách hoạt động**

TensorFlow cho phép các lập trình viên tạo ra dataflow graph, cấu trúc mô tả làm thế nào dữ liệu có thể di chuyển qua 1 biểu đồ, hay 1 sê-ri các node đang xử lý. Mỗi node trong đồ thị đại diện 1 operation toán học, và mỗi kết nối hay edge giữa các node là 1 mảng dữ liệu đa chiều, hay còn được gọi là ‘tensor’.

TensorFlow cung cấp tất cả những điều này cho lập trình viên theo phương thức của ngôn ngữ Python. Vì Python khá dễ học và làm việc, ngoài ra còn cung cấp nhiều cách tiện lợi để ta hiểu được làm thế nào các high-level abstractions có thể kết hợp cùng nhau. Node và tensor trong TensorFlow là các đối tượng Python, và các ứng dụng TensorFlow bản thân chúng cũng là các ứng dụng Python.

Các operation toán học thực sự thì không được thi hành bằng Python. Các thư viện biến đổi có sẵn thông qua TensorFlow được viết bằng các binary C++ hiệu suất cao. Python chỉ điều hướng lưu lượng giữa các phần và cung cấp các high-level abstraction lập trình để nối chúng lại với nhau.

TensorFlow 2.0, được ra mắt vào tháng 10 năm 2019, cải tiến framework theo nhiều cách dựa trên phản hồi của người dùng, để dễ dàng và hiệu quả hơn khi làm việc cùng nó (ví dụ: bằng cách sử dụng các Keras API liên quan đơn giản cho việc train model). Train phân tán dễ chạy hơn nhờ vào API mới và sự hỗ trợ cho TensorFlow Lite cho phép triển khai các mô hình trên khá nhiều nền tảng khác nhau. Tuy nhiên, nếu đã viết code trên các phiên bản trước đó của TensorFlow thì bạn phải viết lại, đôi lúc 1 ít, đôi lúc cũng khá đáng kể, để tận dụng tối đa các tính năng mới của TensorFlow 2.0.

**1.3.3 Lợi ích**

Lợi ích dễ thấy nhưng quan trọng nhất mà TensorFlow cung cấp cho việc lập trình machine learning chính là abstraction. Thay vì phải đối phó với những tình huống rườm rà từ việc thực hiện triển khai các thuật toán, hay tìm ra cách hợp lý để chuyển output của 1 chức năng sang input của 1 chức năng khác, giờ đây bạn có thể tập trung vào phần logic tổng thể của 1 ứng dụng hơn. TensorFlow sẽ chăm sóc phần còn lại thay cho bạn.

Ngoài ra TensorFlow còn ung cấp các tiện ích bổ sung cho các lập trình viên cần debug cũng như giúp bạn tự suy xét các ứng dụng TensorFlow. Chế độ [eager execution](https://www.tensorflow.org/programmers_guide/eager) cho phép bạn đánh giá và sửa đổi từng operation của biểu đồ 1 cách riêng biệt và minh bạch, thay vì phải dựng toàn bộ biểu đồ dưới dạng 1 đối tượng độc lập vốn khá mơ hồ hay phải đánh giá chung tổng thể. Cuối cùng, 1 tính năng khá độc đáo của TensorFlow là [TensorBoard](https://www.tensorflow.org/programmers_guide/summaries_and_tensorboard). TensorBoard cho phép bạn quan sát 1 cách trực quan những gì TensorFlow đang làm.

TensorFlow còn có nhiều cải tiến từ sự hậu thuẫn từ các ekíp thương mại hạng A tại Google. Google không những tiếp lửa cho tiến độ nhanh chóng cho sự phát triển đằng sau dự án, mà còn tạo ra nhiều phục vụ độc đáo xung quanh TensorFlow để nó dễ dàng deploy và sử dụng: như silicon TPU mình đã nói ở trên để tăng tốc hiệu suất đám mây Google, 1 online hub cho việc chia sẻ các model được tạo với framework, sự hiện diện của [in-browser](https://js.tensorflow.org/) và [gần gũi với mobile](https://www.tensorflow.org/mobile/tflite/) của framework, và [nhiều hơn thế nữa…](https://siliconangle.com/2018/04/02/new-google-investments-tensorflow-now-ais-leading-development-framework/)

* 1. **Bộ dữ liệu CIFAR-10**

CIFAR-10 là bộ dữ liệu về thị giác máy tính rất phổ biến. Bộ dữ liệu này đã được nghiên cứu kỹ lưỡng trong nhiều loại mô hình Deep learning phục vụ cho nhận dạng đối tượng.

Bộ dữ liệu này bao gồm 60.000 hình ảnh được chia thành 10 target classes, với mỗi class chứa 6000 hình ảnh có kích thước 32 x 32.

Bộ dữ liệu này chứa các hình ảnh có độ phân giải thấp (32 x 32), cho phép các researchers thử các thuật toán mới. 10 lớp khác nhau của tập dữ liệu này là

* Airplane
* Car
* Bird
* Cat
* Deer
* Dog
* Frog
* Horse
* Ship
* Truck

CIFAR-10 đã có sẵn trong datasets module của Keras, chúng ta chỉ việc import trực tiếp từ keras.datasets.

* 1. **Deep learning**

Trí tuệ nhân tạo đang len lỏi vào trong cuộc sống và ảnh hưởng sâu rộng tới mỗi chúng ta, các cụm từ "Artificial Intelligence", "Machine Learing", "Deep Learning" đã không còn quá xa lạ gì. Chúng ta cùng xem hình vẽ để mô tả lại mối quan hệ giữa artificial intelligence, machine learning, và deep learning:

**Deep learning** đã và đang là một chủ đề AI được bàn luận sôi nổi. Là một phạm trù nhỏ của machine learning, deep learning tập trung giải quyết các vấn đề liên quan đến mạng thần kinh nhân tạo nhằm nâng cấp các công nghệ như nhận diện giọng nói, tầm nhìn máy tính và xử lý ngôn ngữ tự nhiên. Deep learning đang trở thành một trong những lĩnh vực hot nhất trong khoa học máy tính. Chỉ trong vài năm, deep learning đã thúc đẩy tiến bộ trong đa dạng các lĩnh vực như nhận thức sự vật (object perception), dịch tự động (machine translation), nhận diện giọng nói,… - những vấn đề từng rất khó khăn với các nhà nghiên cứu trí tuệ nhân tạo.

Trí tuệ nhân tạo có thể được hiểu đơn giản là được cấu thành từ các lớp xếp chồng lên nhau, trong đó mạng thần kinh nhân tạo nằm ở dưới đáy, machine learning nằm ở tầng tiếp theo và deep learning nằm ở tầng trên cùng.

Deep Learning được nhắc đến nhiều trong những năm gần đây, nhưng nền tảng cơ bản đã xuất hiện từ rất lâu​ ​Đã xuất hiện từ khá lâu nhưng kể từ năm 2012, deep learning mới có những bước đột phá lớn, hàng loạt các thư viện hỗ trợ deep learning ra đời. Cùng với đó, ngày càng nhiều kiến trúc deep learning ra đời, khiến cho số lượng ứng dụng và các bài báo liên quan tới deep learning tăng lên chóng mặt.

* 1. **Thuật toán CNN**

Convolutional neural network (tên viết là CNNs mạng nơ ron tích chập) là một trong những mô hình Deep Learning tiên tiến nhất hiện nay. Nó sẽ giúp cho bạn xây dựng hệ thống thông minh nhờ độ chính xác cao nhất.

CNN thường được sử dụng trong các bài toán nhận dạng các object có sẵn trong ảnh. Để bạn đọc có thể hiểu rõ lý do thuật toán này được dùng rộng rãi cho việc nhận dạng. Chúng ta hãy tìm hiểu kỹ lưỡng hơn về nó như sau:

Convolutional là một trong những cửa sổ dạng trượt nằm ngay trên ma trận. Các convolutional layer thường có các parameter được học để điều chỉnh thông tin chính xác mà không cần chọn feature.

Convolution thường sẽ tích chập và nhân từng phần tử có sẵn trong ma trận. Khi đó, Sliding window được gọi là kernel, filter hoặc feature detect. Nó là loại ma trận có kích thước nhỏ.

Ngoài ra, Convolution hay tích chập là các phần tử bên trong ma trận 3×3 với các ma trận ở bên trái. Khi đó, kết quả sẽ cho ra một ma trận gọi là Convoled feature sẽ được sinh ra từ khi nhận được ma trận Filter với ma trận ảnh 5×5.

* 1. **Tkinter**

Tkinter là một gói trong Python có chứa module Tk hỗ trợ cho việc lập trình GUI. Tk ban đầu được viết cho ngôn ngữ **Tcl**. Sau đó Tkinter được viết ra để sử dụng Tk bằng trình thông dịch Tcl trên nền Python. Ngoài Tkinter ra còn có một số công cụ khác giúp tạo một ứng dụng GUI viết bằng Python như wxPython, PyQt, và PyGTK.

* 1. **Numpy**

Numpy là một thư viện lõi phục vụ cho khoa học máy tính của Python, hỗ trợ cho việc tính toán các mảng nhiều chiều, có kích thước lớn với các hàm đã được tối ưu áp dụng lên các mảng nhiều chiều đó. Numpy đặc biệt hữu ích khi thực hiện các hàm liên quan tới Đại Số Tuyến Tính.

# CHƯƠNG 2. PHÁT TRIỂN BÀI TOÁN

## 2.1 Mô tả bài toán:

Tải tập dữ liệu Cifar-10 lưu vào máy, xây dựng model theo thuật toán CNN rồi lưu lại model đó ở file dạng web.h5. Xây dựng giao diện người dùng bao gồm 2 nút bấm và một khung ảnh. Một nút bấm tải ảnh lên từ máy tính và một nút bấm để chương trình phân tích đưa ra dự đoán về bức hình đó là gì.

## 2.2 Các chức năng chính

* Tải ảnh từ các thư mục của máy tính
* Phân tích ảnh và đưa ra dự đoán về bức ảnh đó
* Giao diện cho phép nhìn thấy bức ảnh đã tải lên

## 2.3 Sơ đồ hoạt động

Tải dữ liệu Cifar-10

Xây dựng model theo thuật toán CNN

Mở giao diện người dùng

Tải hình ảnh muốn nhận diện lên giao diện

Chương trình tiến hành phân tích và đưa ra kết quả dự đoán

# CHƯƠNG 3. KẾT QUẢ THỰC HIỆN

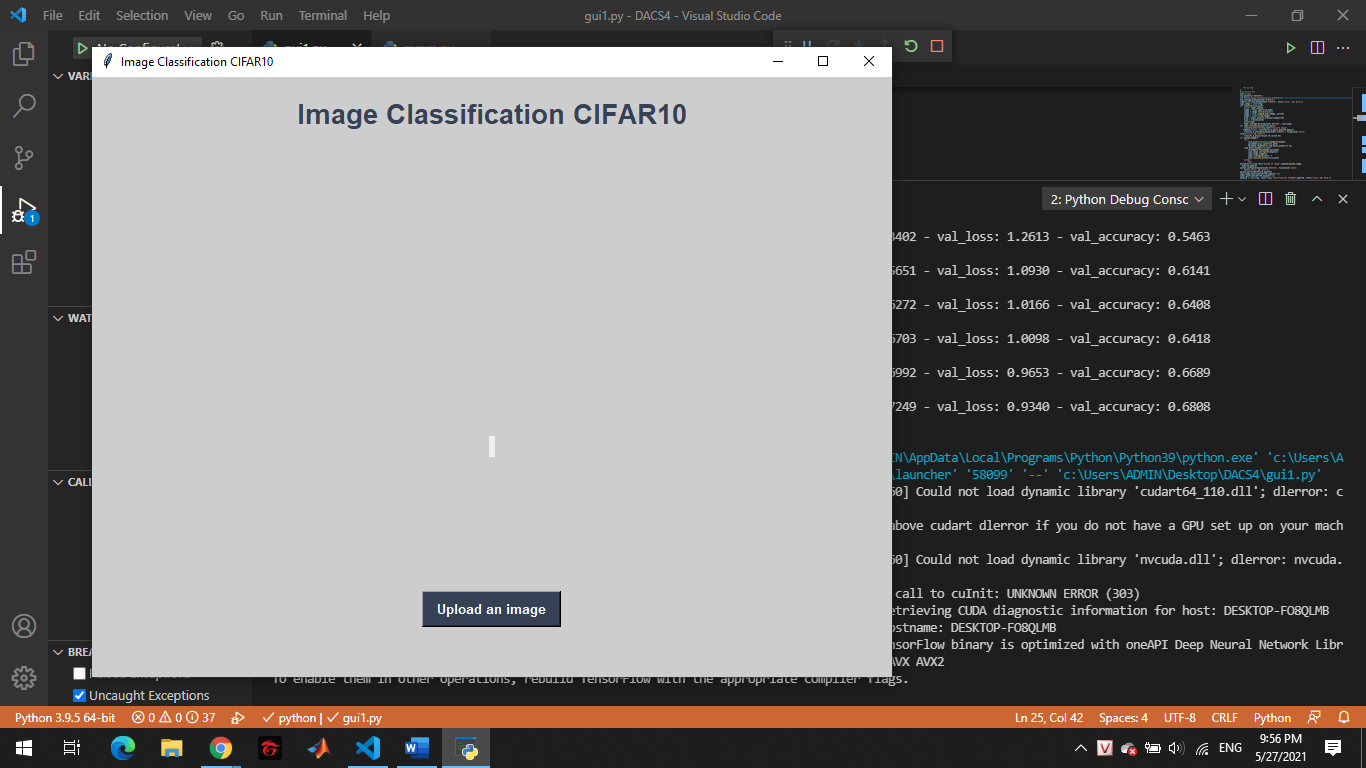
**3.1 Kết quả:**

Thông qua việc xây dựng chương trình phân loại hình ảnh bằng ngôn ngữ lập trình python với keras.

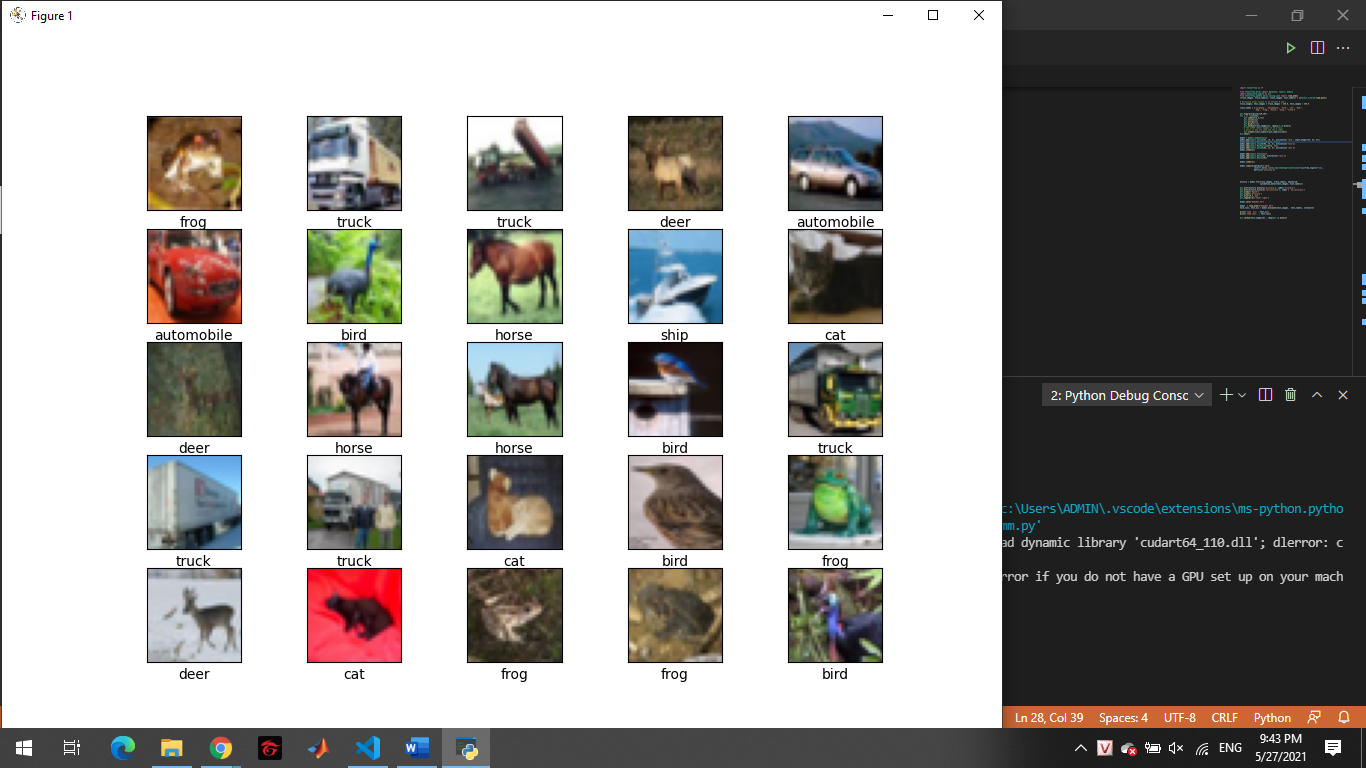
Có thể phân tích dự đoán bức hình đó.

**3.2. Giới thiệu các chức năng** **chính**

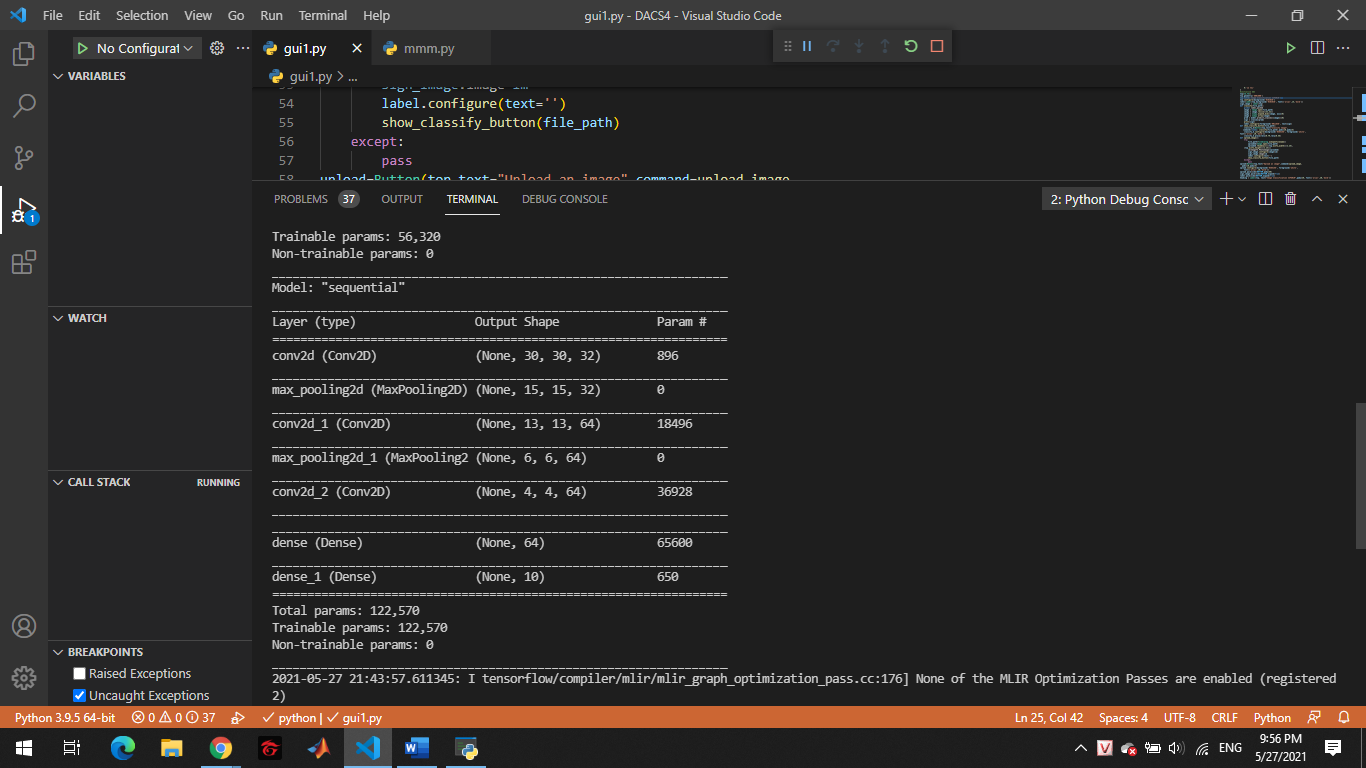
**3.2.1 Giao Diện người dùng**



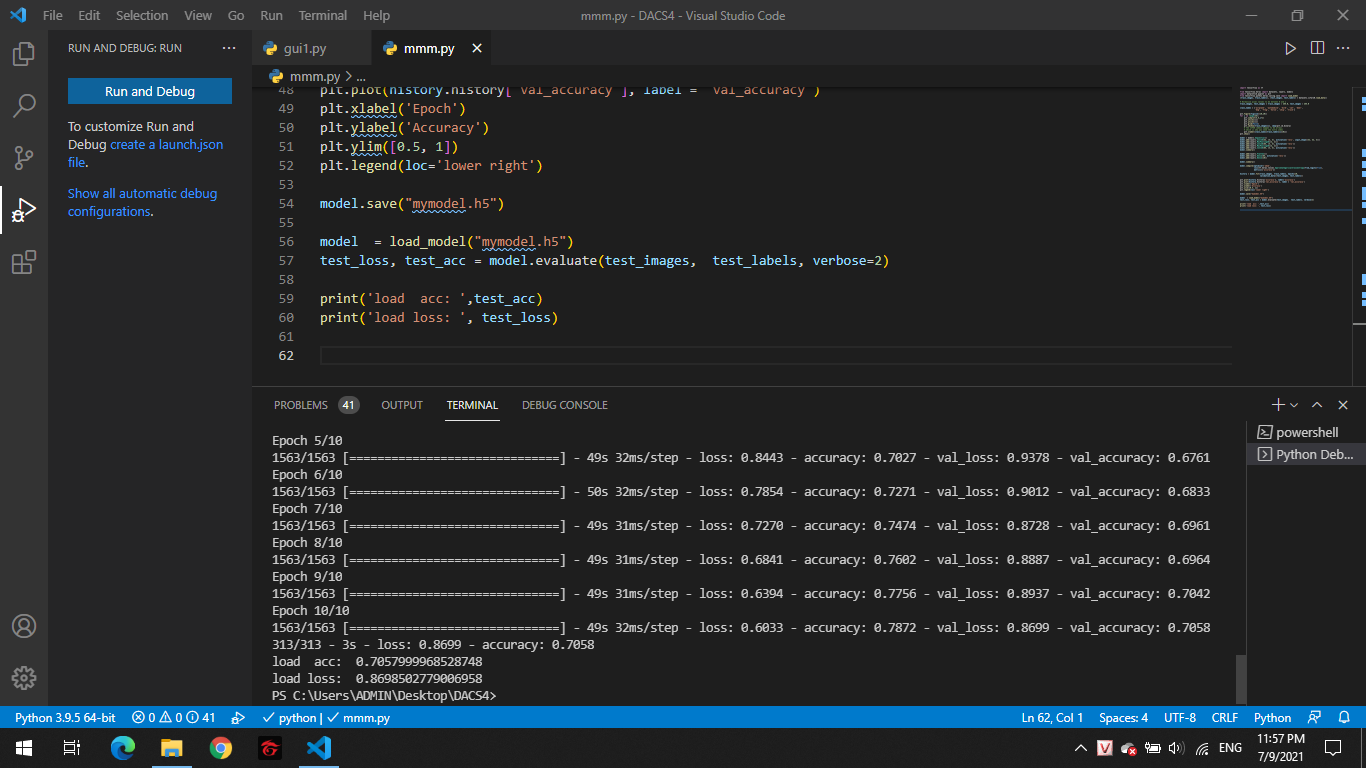
**3.2.2 Test bộ dữ liệu hình ảnh Cifar-10**



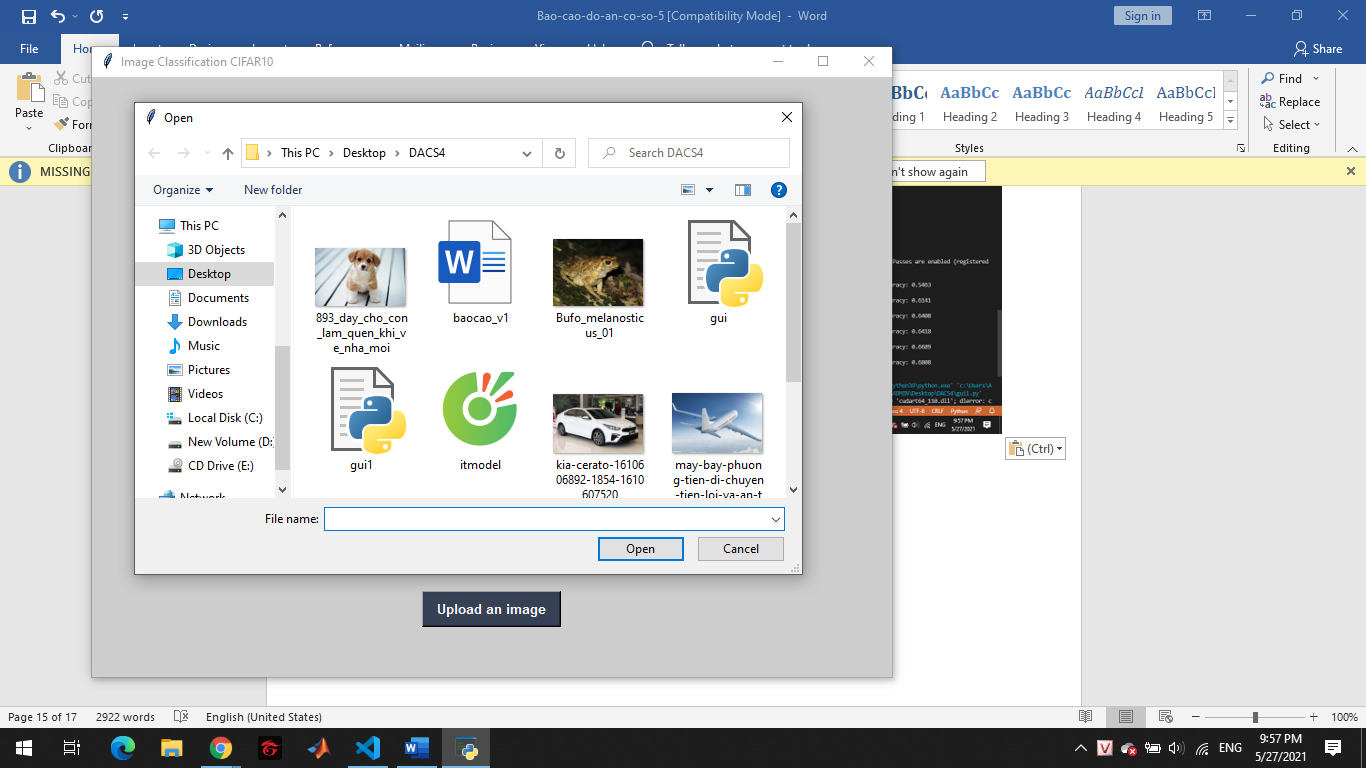
**3.2.3 Các lớp layer trong model**

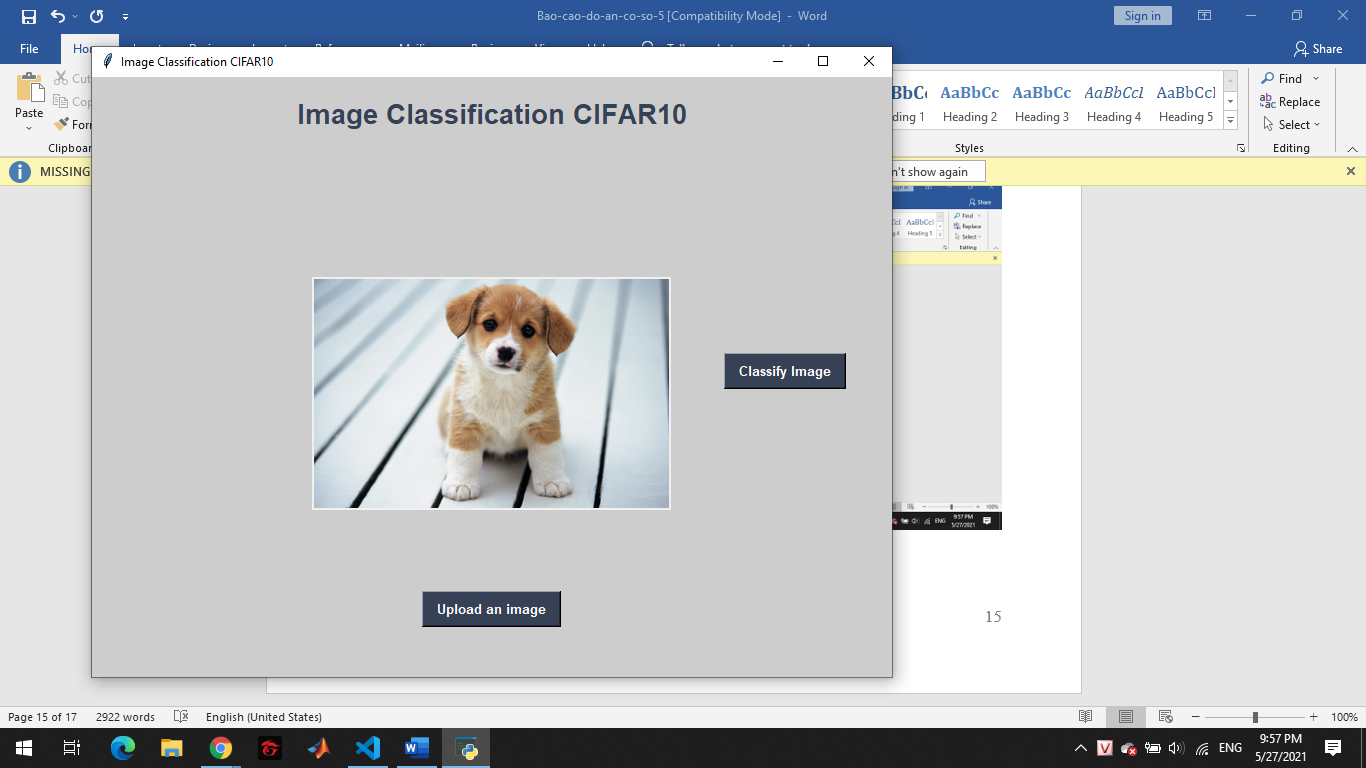


**3.2.4 Train model**

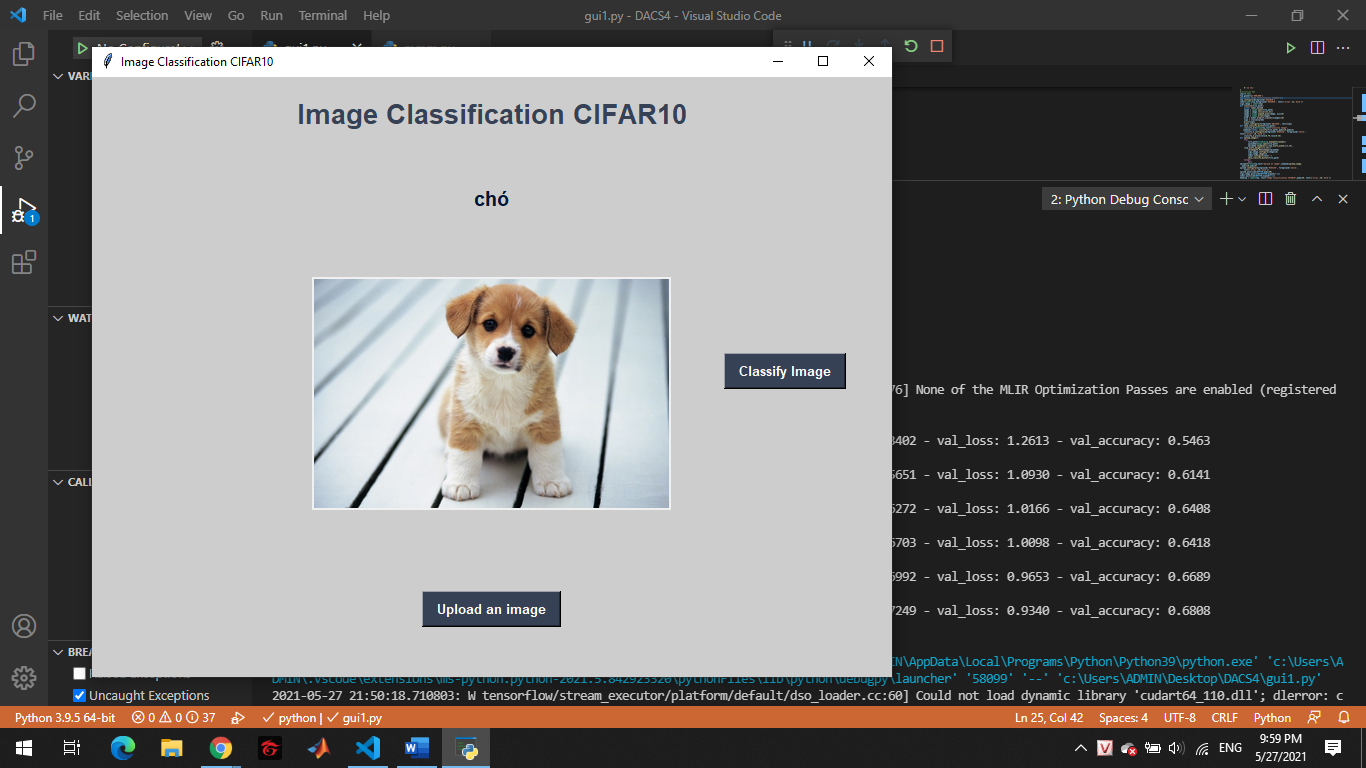


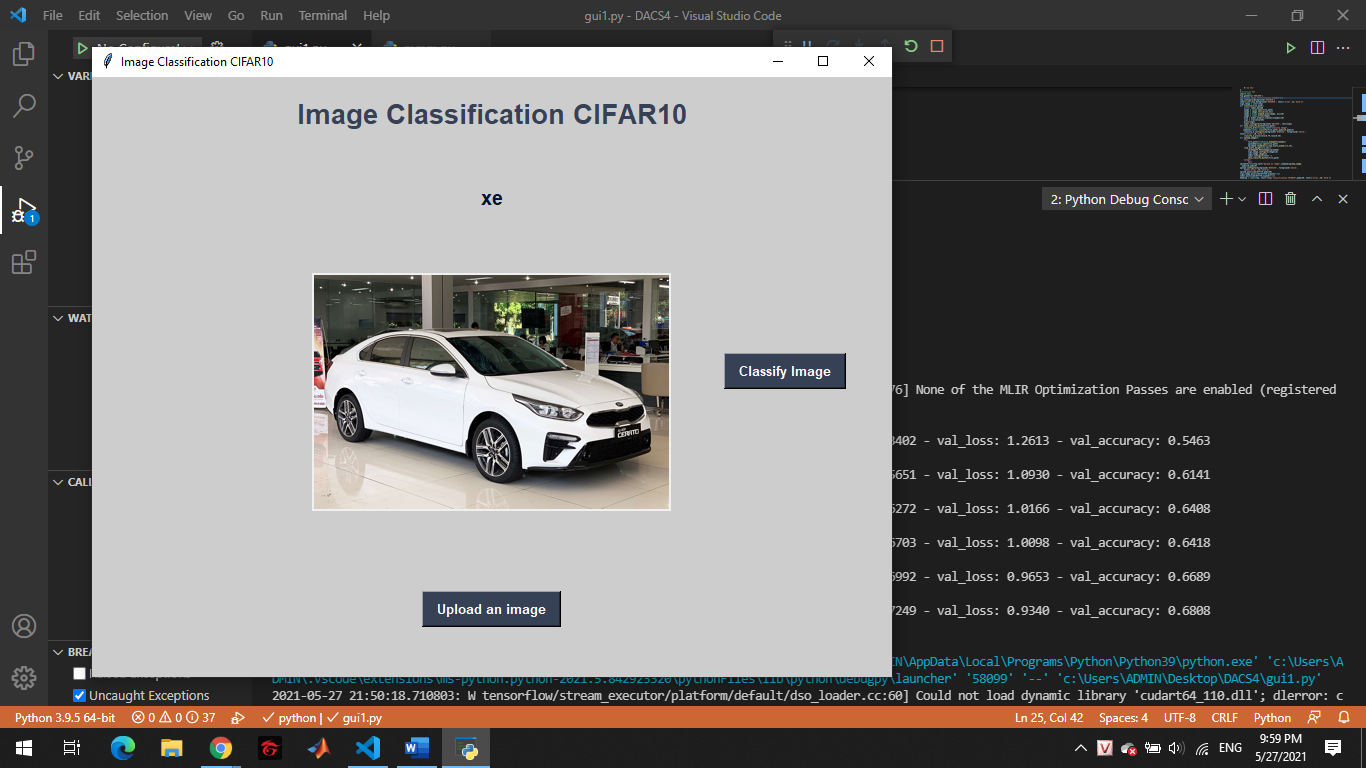
**3.2.5 Mở file ảnh ở thư mục máy tính**

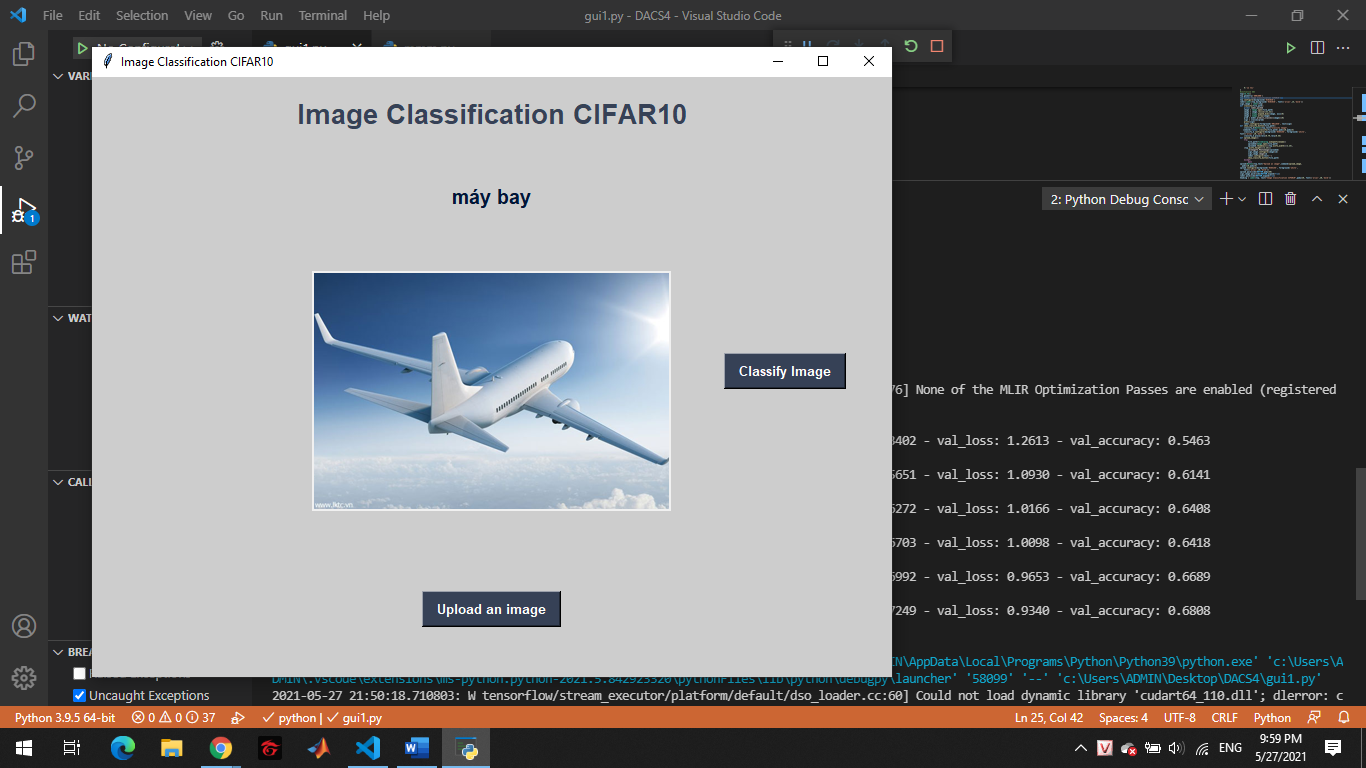


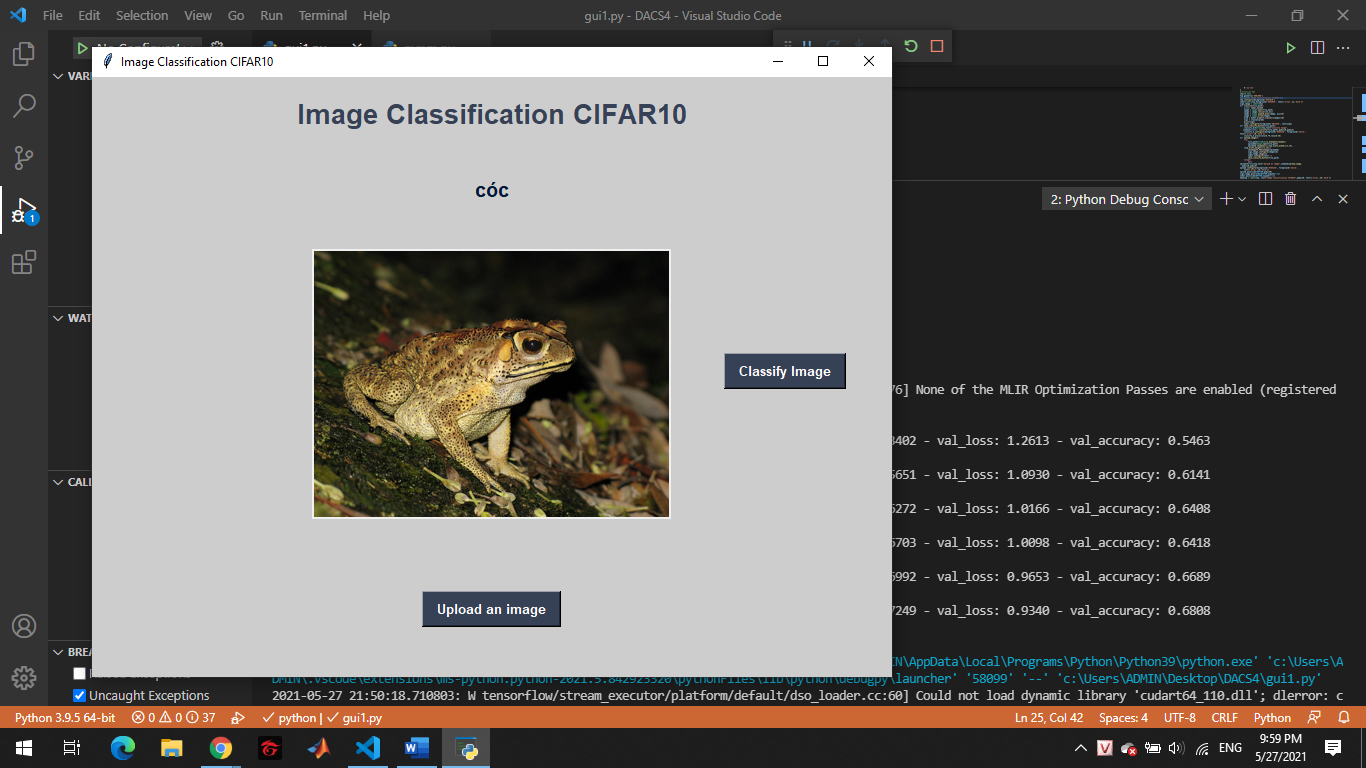


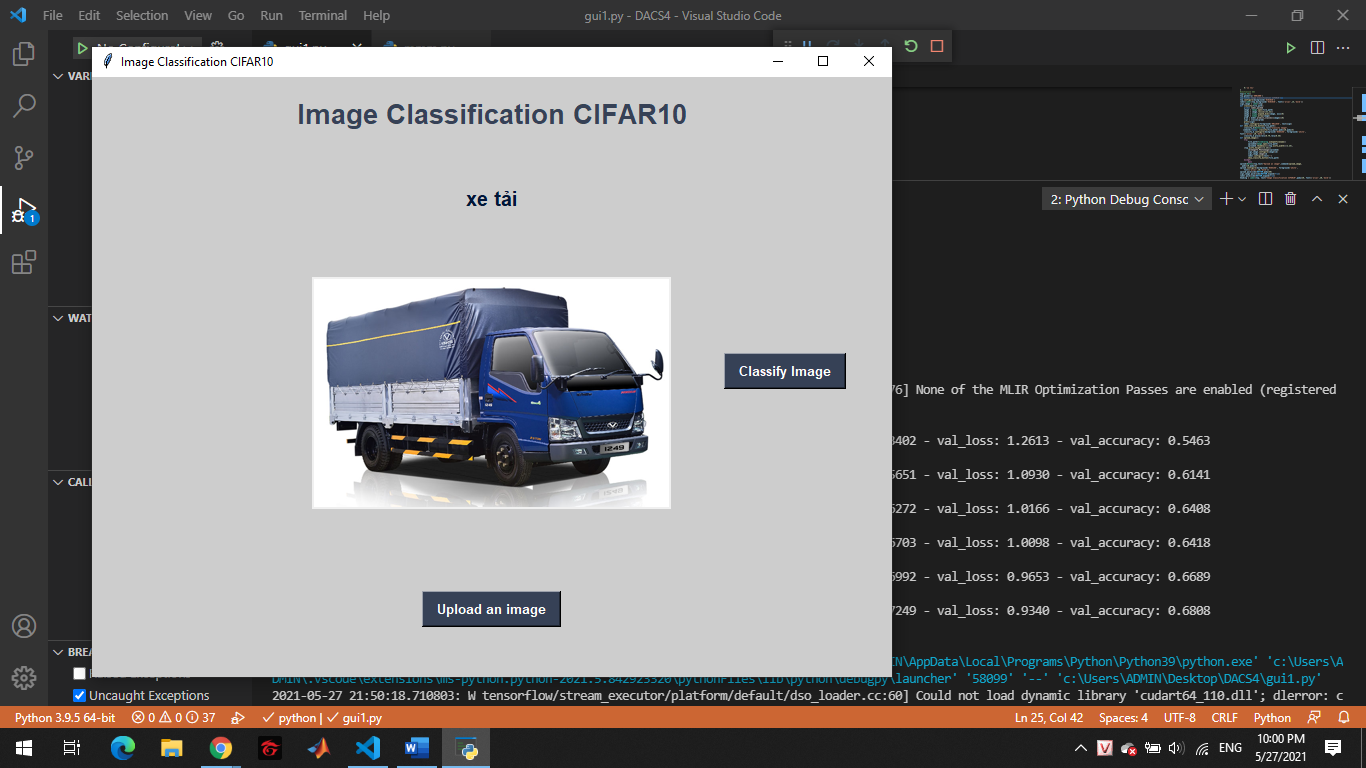
**3.2.6 Đưa ra kết quả dự đoán**

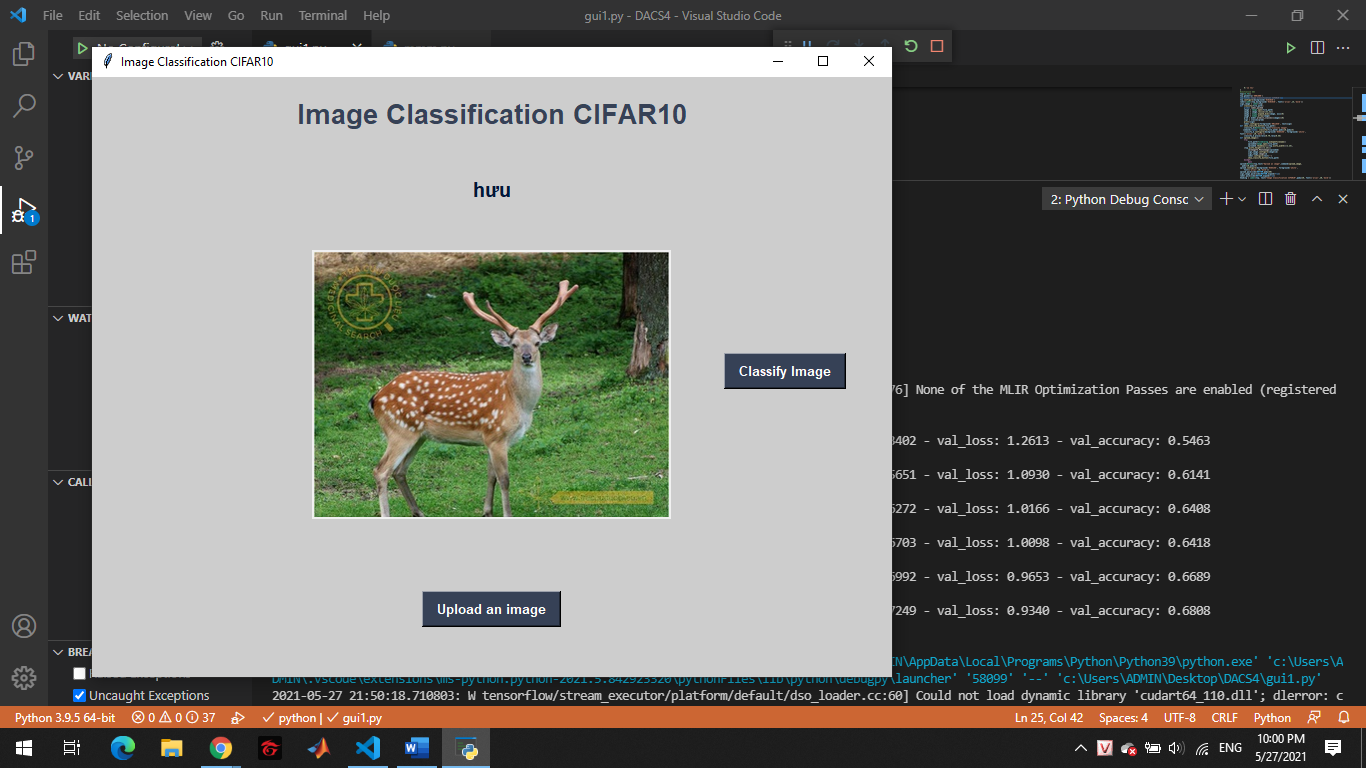




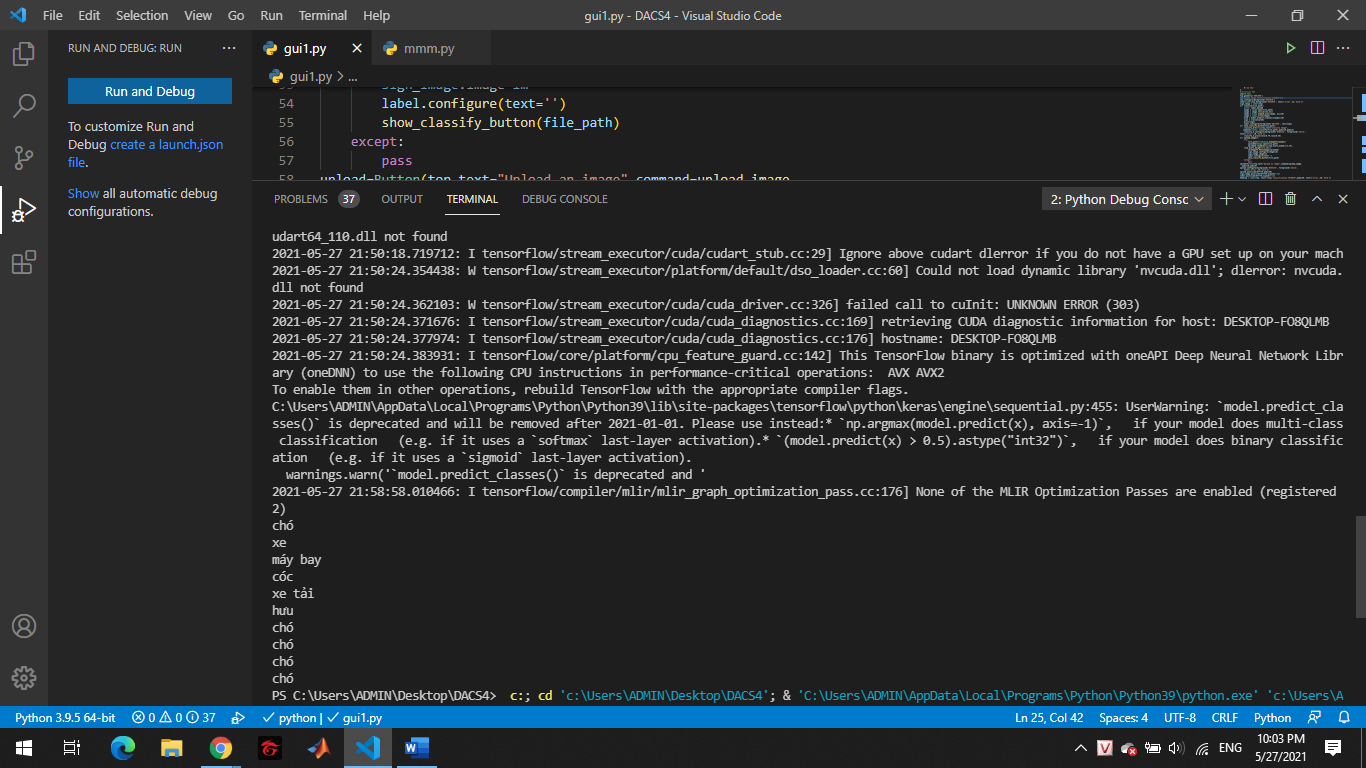






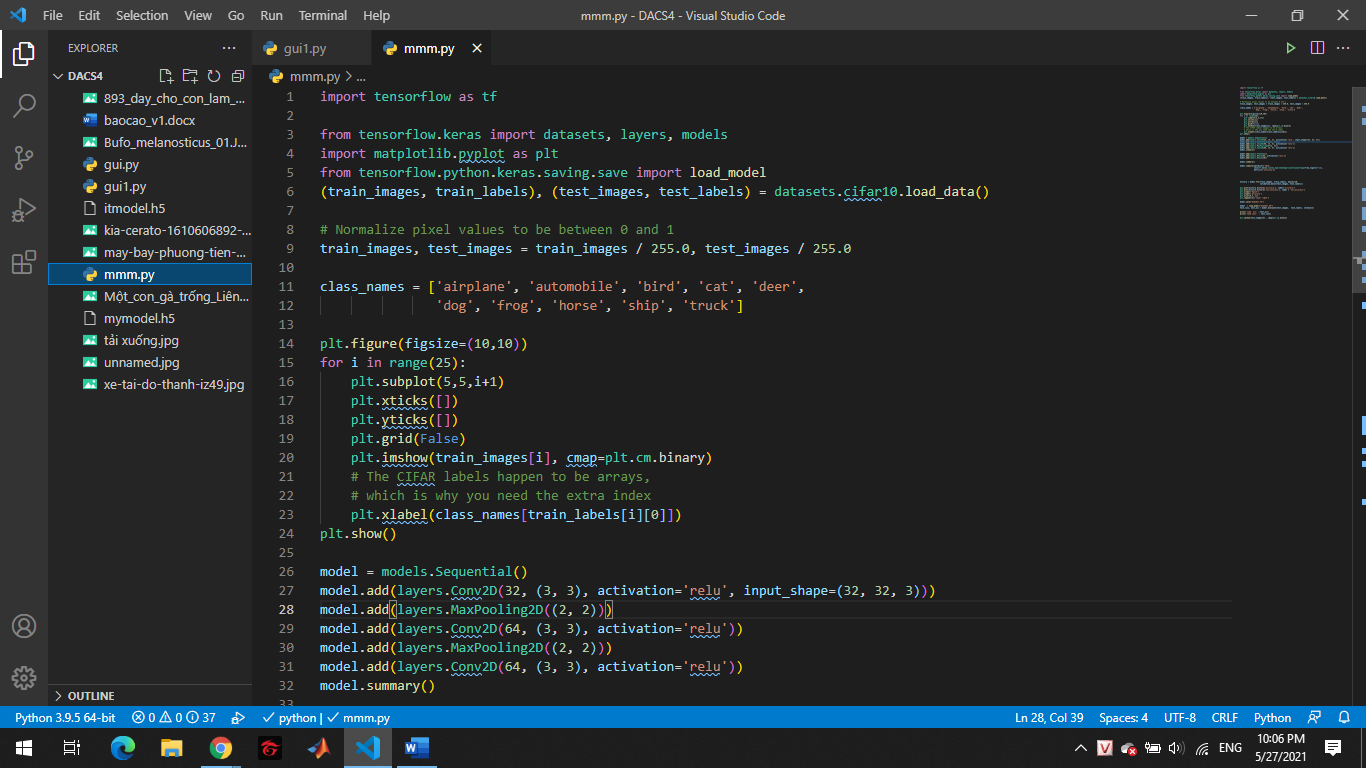


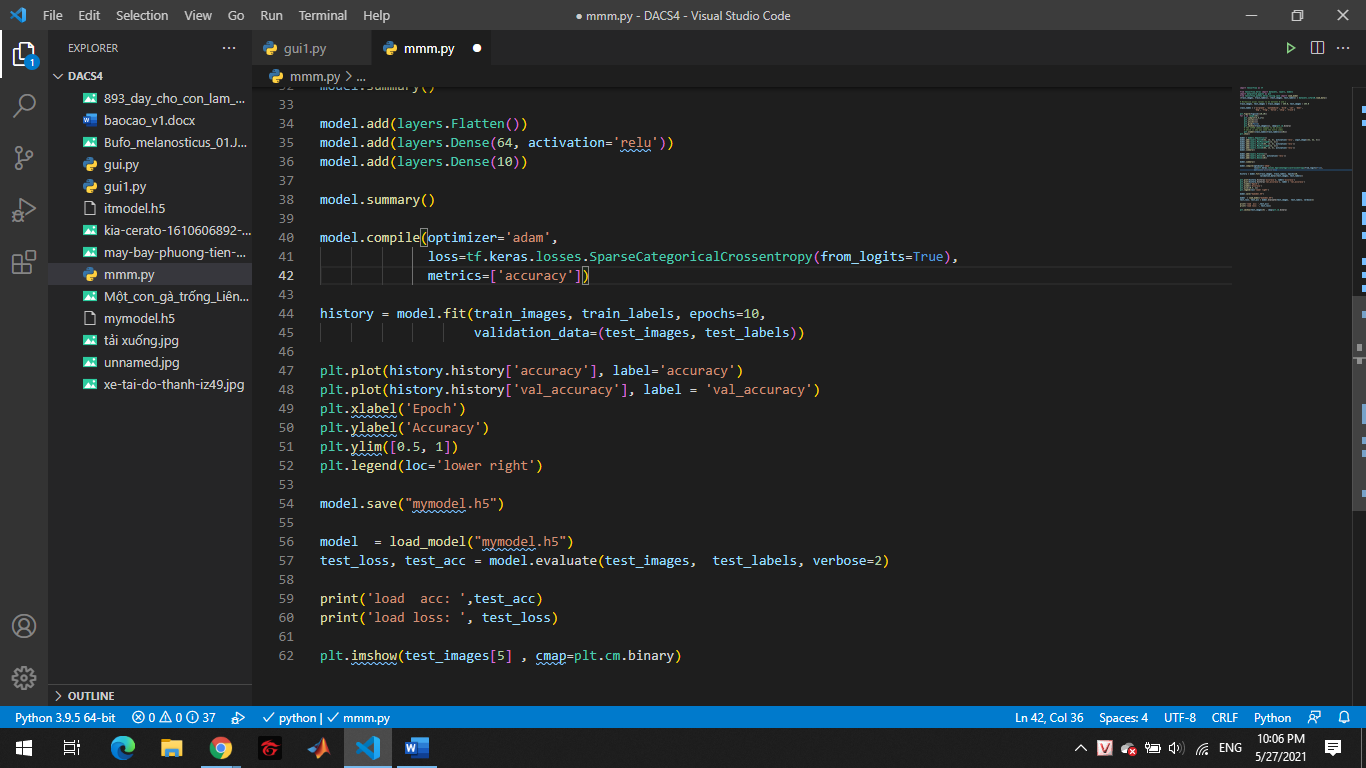
**3.2.7 lưu lại lịch sử dự đoán sau khi khởi chạy chương trình**



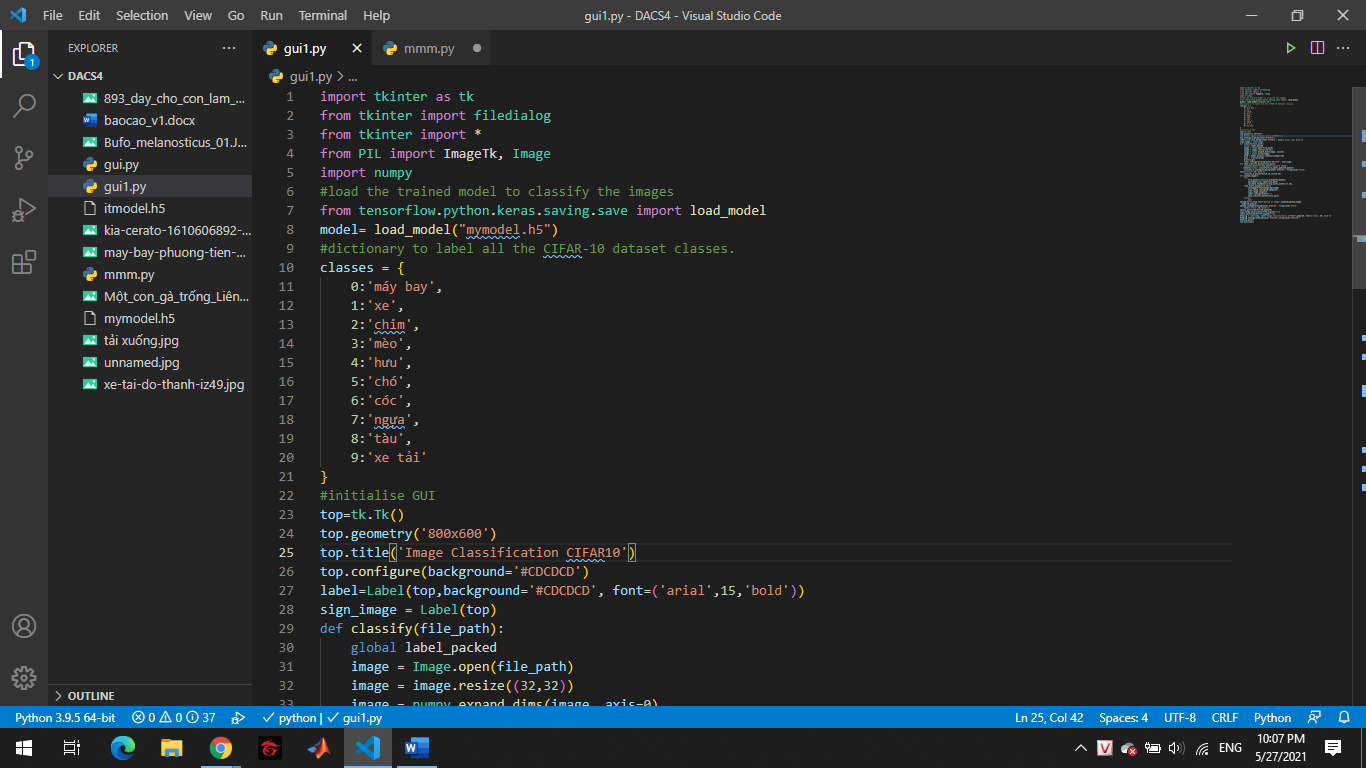
**3.3 Các đoạn mã để viết lên chương trình**

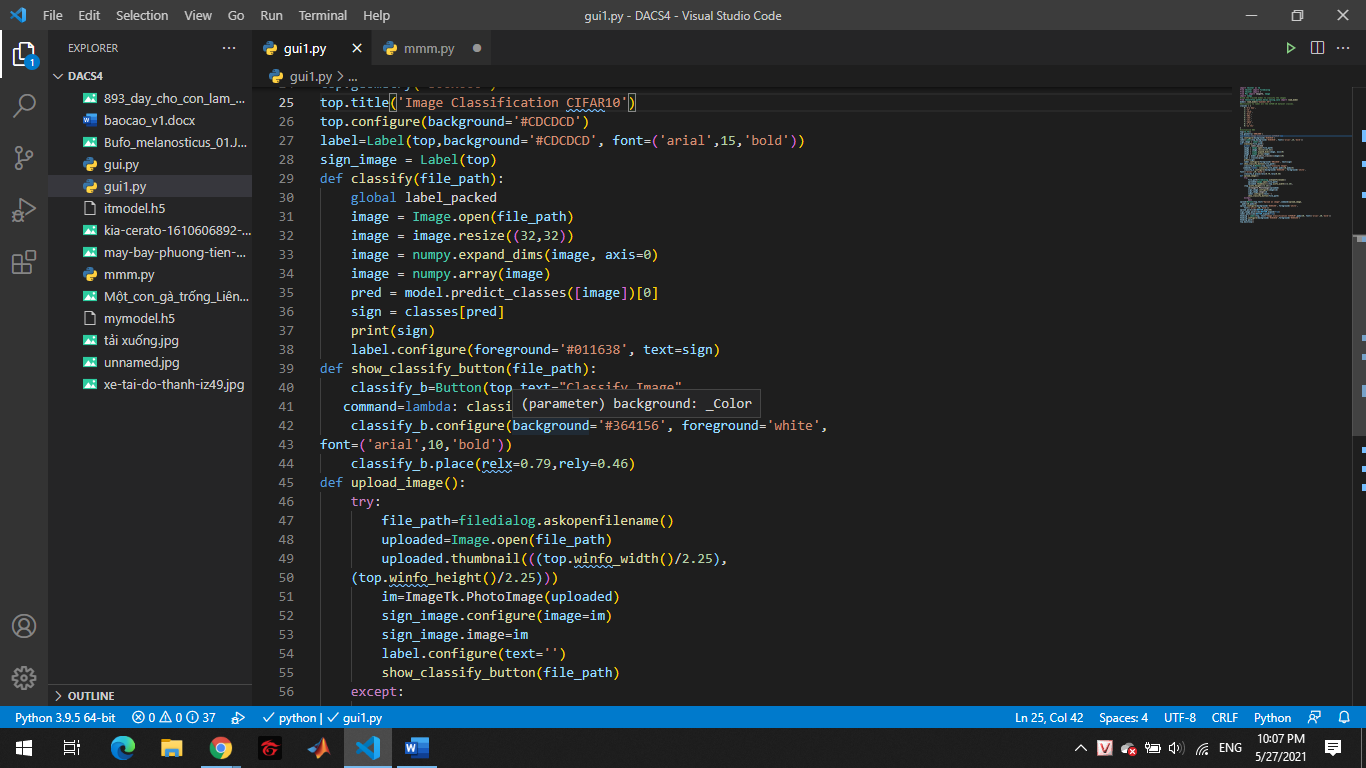
**3.3.1 File tạo model và lưu model**

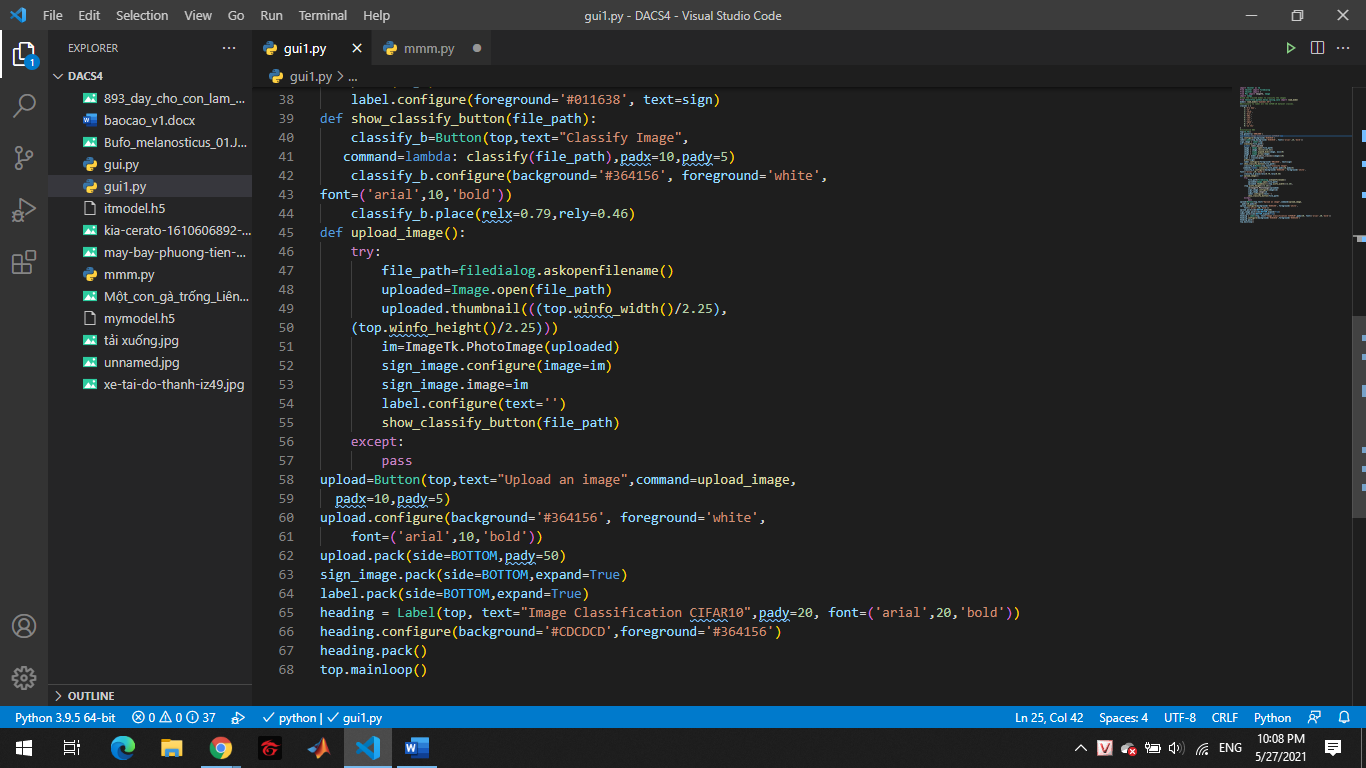




**3.3.2 File giao diện người dùng**







# 

# KẾT LUẬN

## 1.Kết quả thực hiện được:

Thông qua việc thực hiện đồ án cơ sở 5 về đề tài xây dựng chương trình điều khiển máy tính từ xa tôi đã:

- Nghiên cứu

- Ứng dụng

## 2.Ưu điểm

* Đã có sẵn thư viện và dự liệu có thể sử dụng
* Có nhiều sự hướng dẫn có thể thao khảo được từ google

**3.Nhược điểm**

* Phải tìm tài liệu từ nhiều nguồn khác nhau.
* Các hướng dẫn là đồ án có những phiên bản cũ không tương thích với phiên bản hiện tại
* Việc cho máy tính học sâu dẫn đến quá trình máy tính chạy hết công xuất làm cho các thao tác máy tính lúc đó rất chậm
* Áp lực thời gian

## 4. Hướng phát triển

* Sử dụng thuật toán tối ưu hơn cho các chức năng.
* Cải thiện giao diện người dùng
* Thêm các bộ dữ mới để có thể đưa ra được nhiều dự đoán hơn
* Đưa âm thanh vào chương trình sau khi dự đoán và đưa ra kết quả có thể đọc.
* Nâng cấp từ phân loại các sự vật qua hình ảnh lên phân loại các sự vật qua video.

## Tài liệu tham khảo

<http://nghelaptrinh.net/convolutional-neural-network-la-gi/>

<https://topdev.vn/blog/tensorflow-la-gi/>

<https://viblo.asia/p/gioi-thieu-ve-deep-learning-thu-vien-keras-63vKjDGAl2R>

<https://viblo.asia/p/gioi-thieu-ve-numpy-mot-thu-vien-chu-yeu-phuc-vu-cho-khoa-hoc-may-tinh-cua-python-maGK7kz9Kj2>

<https://quantrimang.com/python-la-gi-tai-sao-nen-chon-python-140518>

Và một vài nguồn tài liệu khác.