TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

Khoa Khoa học máy tính



ĐỒ ÁN MÔN HỌC

Nhận Dạng

Đề tài: Nhận dạng và phân loại một số loại quả

Nhóm thực hiện: Teacher:

- Lê Thanh Tiềm 16521214 Ths. Đỗ Văn Tiến

- Nguyễn Đăng Thịnh 16521177

- Trần Văn Quang 16521004

- Nguyễn Hữu Sơn 16521036

- Trần Quốc Dũng 16520263

*Tp Hồ Chí Minh, tháng 12 năm 2019*

**MỤC LỤC**

[**LỜI MỞ ĐẦU** 2](#_Toc28719931)

[**I.GIỚI THIỆU BÁI TOÁN** 4](#_Toc28719932)

[**1. Lý do chọn đề tài** 4](#_Toc28719933)

[**3. Tính ứng dụng của đề tài** 4](#_Toc28719934)

[**II-CÁC CÁCH TIẾP CẬN** 5](#_Toc28719935)

[**1. Phương pháp truyền thống** 5](#_Toc28719936)

[**2. Phương pháp deep learning** 5](#_Toc28719937)

[**III. CÁC BƯỚC THỰC HIỆN** 9](#_Toc28719938)

[**IV. THỰC NGHIỆM** 12](#_Toc28719939)

[**1. Dataset** 12](#_Toc28719940)

[**2.Kết quả** 13](#_Toc28719941)

[**V. KẾT LUẬN & HƯỚNG PHÁT TRIỂN** 16](#_Toc28719942)

[**1. Kết luận** 16](#_Toc28719943)

[**2. Hướng phát triển** 16](#_Toc28719944)

[**VI. PHÂN CÔNG CÔNG VIỆC** 16](#_Toc28719945)

[**VII. REFERENT** 18](#_Toc28719946)

# **LỜI MỞ ĐẦU**

Công nghệ thông tin hiện đang phát triển như vũ bão trên mọi lĩnh vực hoạt động ở khắp nơi trên toàn thế giới. Các lĩnh vực thuộc về máy học mà nhỏ hơn nữa là thị giác máy tính cũng không phải là ngoại lệ. Rất nhiều công nghệ đã được phát triển và áp dụng vào thực tế như nhận dạng biển số xe, phát hiện và cảnh báo giao trong giao thông, nhận diện khuôn mặt,… Những công nghệ đó đã góp phần tự động hóa các vấn đề trong thực tế, giúp nâng cao hiệu suất,…

Chủ đề deep learning hiện nay đang rất hot bởi những ưu điểm vượt trội củanó trong việc giải quyết các vấn đề trong thị giác máy tính. Vì vậy chúng em quyết định sử dụng mô hình deep learning để giải quyết bài toán mà nhóm đã chọn, thêm vào đó là sự so sánh giữa deep và các phương pháp truyền thống.

Trong khuôn khổ môn học, chúng em đã học được khá nhiều kiến thức. Và đây là đồ án để chúng em áp dụng những kiến thức đã học vào giải quyết một vấn đề.

Chúng em gửi lời cảm ơn chân thành nhất đền thầy Đỗ Văn Tiến đã tận tình giảng dạy và giúp đỡ nhóm chúng em nhiều kiến thức lẫn kinh nghiệm để hoàn thành đồ án này.

Trong quá trình làm đồ án, tất nhiên sẽ không thể tránh khỏi những thiếu xót, mong thầy xem xét và bỏ qua.

Nhóm chúng em xin chân thành cảm ơn,.

Tp Hồ Chí Minh,

Nhóm tác giả

# **I.GIỚI THIỆU BÁI TOÁN**

Image classification hay hiểu đơn giản là phân loại hình ảnh là một trong những nhiệm vụ phổ biến trong Computer Vision. Mục tiêu chính của bài toán này đó chính là phân loại một hình ảnh đầu vào(input) thành một nhãn(label) dầu ra(output).

## **1. Lý do chọn đề tài**

* Áp dụng kiến thức về xử lý ảnh,deeplearning đã học vào các bái toán.
* Hiểu được quá trình xây dựng và training model
* Cách giải quyết bái toán deeplearning

**2. Mô tả bài toán**

Bài toán đề cập đến việc phân loại hình ảnh hoa quả, nhận dạng xem loại quả đó có tên là gì.

* Input: Hình hoa quả
* Output: Trả về kết quả hoa quả đó có tên là gì.

Ví dụ:



Model

**Dứa**

**Dừa**

**Xoài**

**Chuối**

**Táo**

**…..**

**Dưa**

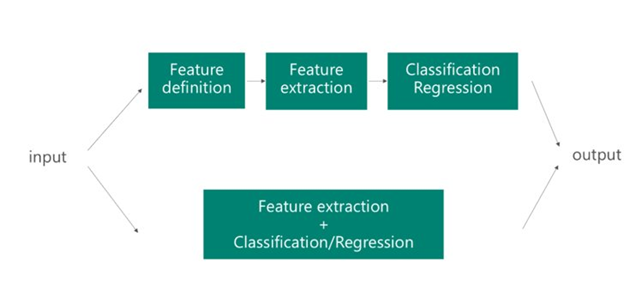
## **3. Tính ứng dụng của đề tài**

Bài toán tuy chỉ còn đơn giản là phân loại nhưng có nhiều ứng dụng khác nhau. Nhóm đề cập đến hai vấn đề chính:

Áp dụng vào phát triển cho giáo dục trẻ em. Giúp trẻ em làm quen và nhận diện các loại hoa, quả một cách dễ dàng, nhanh chóng qua các ứng dụng, trò chơi liên quan.

Ngoài ra, phân loại hoa quả còn giúp cho vấn đề tự động hóa hơn nữa trong các nhà máy hay xí nghiệp, đó là việc tự động hóa phân loại hoa quả trong quy trình sản xuất. Vấn đề này liên quan đến chạy với thời gian thực của băng chuyền tự động

# **II-CÁC CÁCH TIẾP CẬN**



## **1. Phương pháp truyền thống**

Với phương pháp truyền thống thường gồm có 2 phần tách biệt:

- Feature Engineering: quá trình rút trích đặc trưng dựa trên hiểu biết của con người, có giá trị với mô hình. Những đặc trưng này được gọi là hand-crafted features.

- Classifier/Regressor:công đoạn sử dụng các thuật toán Machine learning cho các featuressau quá trình feature engineering để mô hình học.

## **2. Phương pháp deep learning**

Phương pháp deep learning là một phương pháp hiện đại được ứng dụng rộng rãi và mạnh mẽ hiện nay. Đem lại độ chính xác rất cao.

Khác với cách tiếp cận truyền thống,Deep learning kết hợp cả hai quá trình Feature Engineering và Classifier/Regressor lại với nhau

Feature Extractor thường được lấy một cách tự động

**2.1 Lựa chọn phương pháp tiếp cận.**

Bài toán được lựa chọn giải quyết theo hướng tiếp cận deep learning, cụ thể là sử dụng mô hình Convolution Neuron Network.

Lý do:

* + Deep learning là một mô hình end-to-end, việc trích xuất các đặc trưng sẽ được thay đổi cho phù hợp với yêu cầu của bài toán.
  + Đối với dữ liệu hình ảnh, mô hình CNN sử dụng phép convolution có thể trích xuất ra những đặt trưng tốt.

Có nhiều mô hình CNN để lựa chọn như LeNet, GoogLenet, inception, resnet,…

Ở đây nhóm em xây dựng mô hình dựa trên mô hình inceptions V3, Resnet-152 và VGG16 bởi những ưu điểm thực nghiệm của nó và hiệu quả trong quá trình xây dựng.

**2.2. Transfer learning.**

**2.2.1. Transfer learning là gì?**

* Tranfer Learning là một kỹ thuật học máy, nơi một mô hình được đào tạo về một nhiệm vụ được tái sử dụng trên một nhiệm vụ liên quan thứ hai.
* Tranfer Learning là một tối ưu hóa cho phép cải thiện tốc độ và hiệu suất khi mô hình hóa nhiệm vụ thứ hai
* Tranfer Learning liên quan đến các vấn đề như multi-task learning và concept drift.

**2.2.2. Các loại transfer learning:**

* **Theo chủ đề nghiên cứu:**
* **Theo domains(miền tri thức):**
  + Inductive transfer learning: khác biệt về nhiệm vụ (task).
  + Transductive transfer learning: giống nhau về nhiệm vụ nhưng khác biệt về domain.
  + Unsupervised transfer learning. Giống inductive transfer learning nhưng target task ở dạng un-supervised learning điển hình như clustering (nhóm), giảm chiều dữ liệu (dimension reduction), .v.v.
* **Theo hướng tiếp cận:**
  + Instance transfer: Dùng lại tri thức từ source domain từ một số trường hợp/ví dụ chính.
  + Feature representation transfer: Giảm thiểu sự khác biệt giữa 2 domains.
  + Parameter transfer: dựa trên giả thiết rằng 2 tác vụ có cùng thông số hyper parameters.
  + Relational - Knowledge transfer: giải quyết vấn đề khi dữ liệu không độc lập và phân bố ngẫu nhiên, thường được ứng dụng liên quan tới mạng xã hội.

**2.2.3. Tại sao nên dùng transfer learning trong deep learning?**

* **Không đủ dữ liệu**: Deep Learning cần rất nhiều dữ liệu, và ngốn rất nhiều tài nguyên để học trên tập dữ liệu và ứng dụng đó và việc huấn luyện Deep Learning trên tập dữ liệu ít phần nhiều là không hiệu quả. Vậy ngoài kỹ thuật data augmentation trước đó, transfer learning cũng là một giải pháp cho vấn đề này.
* **Không đủ tài nguyên**: Việc học trên tập dữ liệu lớn rất ngốn nhiều tài nguyên. Transfer learning sẽ góp phần giảm phần nào thời lượng training.
* **Cải thiện chất lượng**. Rất nhiều trường hợp transfer learning cải thiện chất lượng dự đoán của Target Task so với việc train lại từ đầu. Lý do có thể do Source Network được train với dữ liệu lớn và học được tính khái quát hóa tốt hơn, hay việc train với Target Task trong khi mạng vẫn có thông tin của Source Task cho tách động của multi-task learning.

**2.2.4. Sử dụng tranfer learning như thế nào?**

Có thể sử dụng transfer learning trong các mô hình dự đoán của chính mình theo hai cách tiếp cận phổ biến:

* **Develop Model**
* **Pre-trained Model**

**Develop Model Approach**

* **Select Source Task**. chọn một mô hình dự báo có liên quan với sự phong phú của dữ liệu trong đó có một số mối quan hệ trong dữ liệu đầu vào, dữ liệu đầu ra và / hoặc các khái niệm đã học trong quá trình ánh xạ từ dữ liệu đầu vào đến đầu ra.
* **Develop Source Model**. Tiếp theo, phải phát triển mô hình cho nhiệm vụ đầu tiên này. Mô hình phải tốt hơn một mô hình Naive để đảm bảo rằng một số tính năng học tập đã được thực hiện.
* **Reuse Model**. Mô hình phù hợp với nhiệm vụ nguồn sau đó có thể được sử dụng làm điểm bắt đầu cho một mô hình trên nhiệm vụ thứ hai quan tâm. Điều này có thể liên quan đến việc sử dụng tất cả hoặc một phần của mô hình, tùy thuộc vào kỹ thuật mô hình được sử dụng.
* **Tune Model**. Mô hình có thể cần phải được điều chỉnh hoặc tinh chỉnh trên dữ liệu cặp đầu vào-đầu ra có sẵn cho nhiệm vụ quan tâm.

#### **Pre-trained Model Approach**

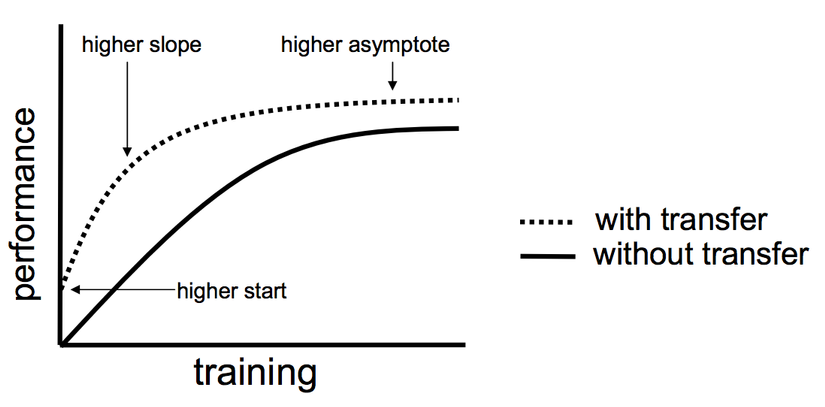
* **Select Source Model**. một pre-trained source model được chọn từ các mô hình có sẵn. Nhiều tổ chức nghiên cứu cung cấp các mô hình trên các tập dữ liệu lớn bạn có thể lựa chọn.
* **Reuse Model**. pre-trained model sau đó có thể được sử dụng làm điểm khởi đầu cho một mô hình trên nhiệm vụ thứ hai quan tâm. Điều này có thể liên quan đến việc sử dụng tất cả hoặc một phần của mô hình, tùy thuộc vào kỹ thuật mô hình được sử dụng.
* **Tune Model**. mô hình có thể cần phải được điều chỉnh hoặc tinh chỉnh trên dữ liệu cặp đầu vào-đầu ra có sẵn cho nhiệm vụ quan tâm.

Cách tiếp cận thứ 2 này rất phổ biến trong deep learning.

**2.2.5. Khi nào thì sử dụng tranfer learning?**

Tranfer Learning là một tối ưu hóa, một phím tắt để tiết kiệm thời gian hoặc nhận được hiệu suất tốt hơn. Tuy nhiên, bạn chỉ nên áp dụng Tranfer Learning trong các trường hợp sau, các trường hợp khác mình nghĩ cứ thực hiện như các bài toán machine learning khác vẫn có thể cho kết quả khả quan đó ạ.

* **Higher start**. The initial skill (trước khi tinh chỉnh mô hình) trên mô hình nguồn cao hơn.
* **Higher slope** Tỉ lệ cải thiện mô hình trong quá trình đào tạo có sự chênh lệch lớn.
* **Higher asymptote**. Tiệm cận hay khả năng hội tụ của mô hình được đào tạo tốt hơn.



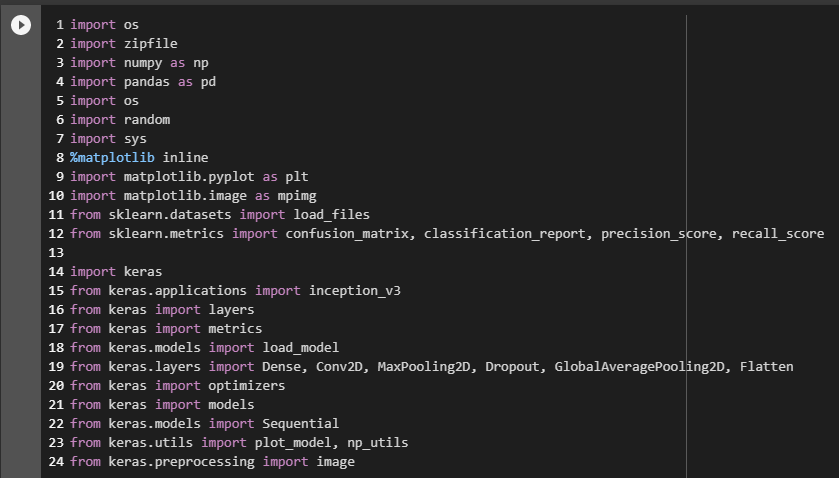
Trong trường hợp lý tưởng nhất, bạn sẽ thấy tất cả ba lợi ích từ một ứng dụng thành công của transfer learning. Đó là một cách tiếp cận để bạn có thể xác định tác vụ có liên quan với dữ liệu, tài nguyên phong phú để phát triển mô hình cho nhiệm vụ đó và tái sử dụng nó như một điểm khởi đầu cho mô hình của riêng bạn. Trong một vài bài toán, nếu bạn không có quá nhiều dự liệu để thực hiện, thì tranfer learning cho phép bạn phát triển mô hình một cách khéo léo, đơn giản, tiết kiệm mà vẫn mang lại hiệu quả cao. Việc lựa chọn dữ liệu nguồn hoặc mô hình nguồn là một vấn đề mở và có thể yêu cầu về miền ứng dụng chuyên môn hóa hoặc thông qua các thực nghiệm.

# **III. CÁC BƯỚC THỰC HIỆN**

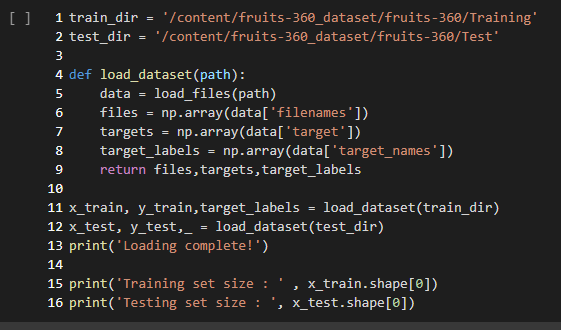
Các bước xây dựng mô hình được dựa trên phương pháp trasfer learning. Theo cách tiếp cận là fine-tuning cho phù hợp với yêu cầu của bài toán và tập dữ liệu.

Thực nghiệm được được tiến hành bằng cách sử dụng lại một pre-trained model của Inception v3, VGG16, Resnet-152 được huấn luyện trên bộ dữ liệu Image net vì tính đương đồng về đặc trưng.

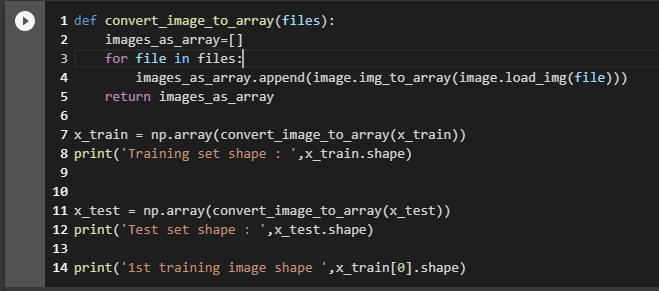
**\* Import thư viện cần thiết:**



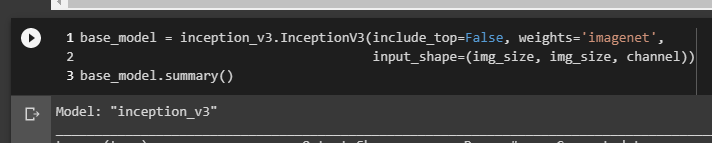
* **Load và xử lý data:**
* Data được chia thành tập train và test. Ta cần load các thư mục



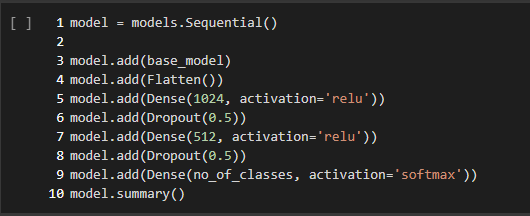
* Convert ảnh thành array:



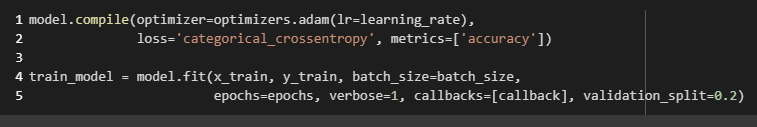
* Load model inception pre-train:
  + Đóng băng các top layer của mô hình lại ( đóng băng các layer FC)



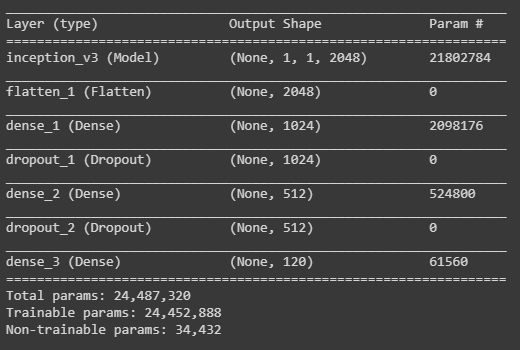
* Khởi tạo các FC mới để train: gồm hai lớp FC, thêm dropout sau mỗi FC



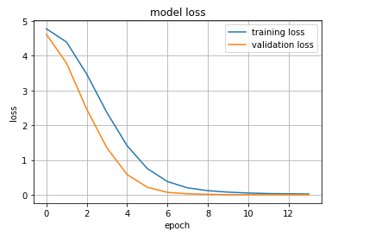
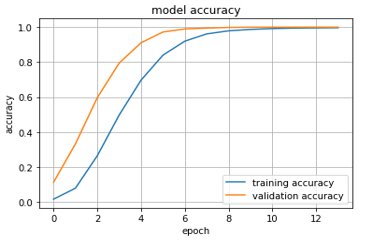
* Tiến hành train:



* + Thông tin về các layer trong model:



* Đánh giá

# **IV. THỰC NGHIỆM**

## **1. Dataset**

Tập dataset được lấy từ trang Kaggle

<https://www.kaggle.com/moltean/fruits>

Dataset gồm 81,120 ảnh​ được chia thành tập train và test riêng biệt như sau:

* 60,498 ảnh cho training​
* 20,622 ảnh cho test.​

Dataset có 120 class bao gồm các loại hoa quả, trái cây khác nhau như: táo, bơ, cherry, nho, kiwi,…​

Kích thước của mỗi ảnh là 100x100​

Một số hình ảnh từ data:



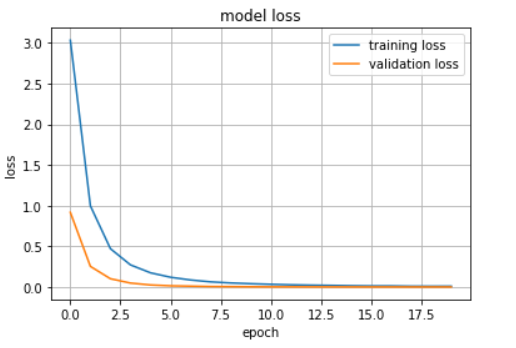
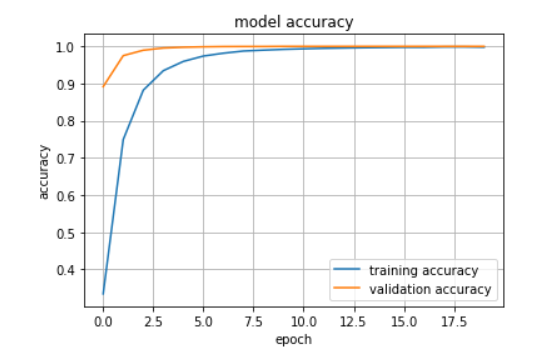
## **2.Kết quả**

**a.Mô hình cổ điển**:

Sử dụng HoG, SIFT để trích xuất đặc trưng: Tuy nhiên việc sử dụng Hog để trích xuất đặc trưntrên tập dữ liệu lớn dẫn đến tốn quá nhiều chi phí lưu trữ và thời gian chạy rất lâu

**b. Mô hình VGG16**:

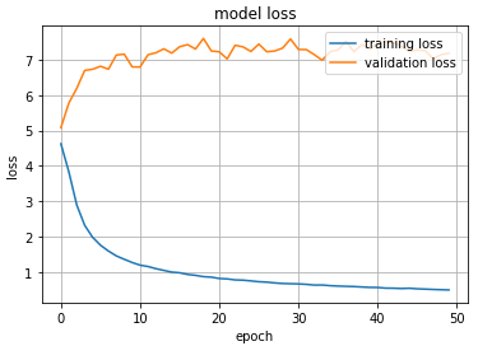
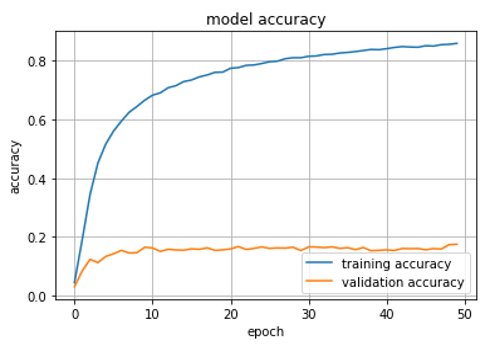
Đóng băng các bottom layer,thêm các FC layer mới để train

=> mô hình cho độ chính xác cao. Accuracy: ~98%

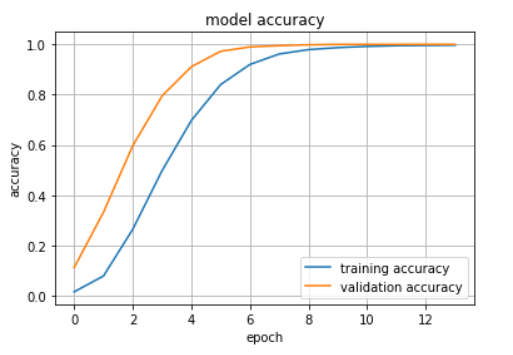
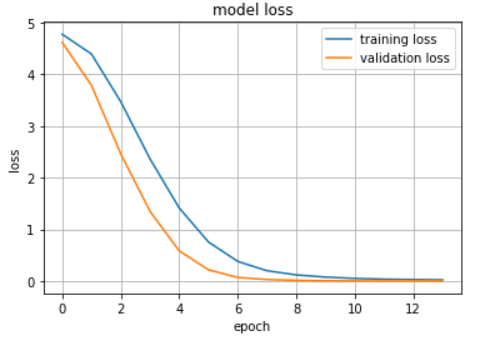
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-score | Support |
| Accuracy |  |  | 0.98 | 20622 |
| Macro avg | 0.98 | 0.97 | 0.98 | 20622 |
| Weighted avg | 0.98 | 0.98 | 0.98 | 20622 |

**c. Mô hình inception v3:**

Khi đóng băng các bottom layer và thêm các FC layer để train, nhận thấy mô

hình bị overfitting

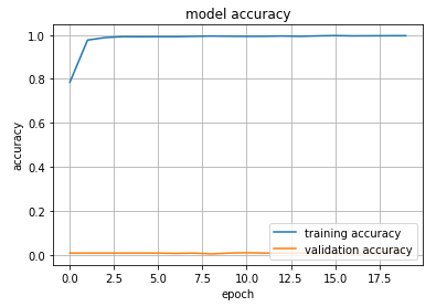
* Mô hình inception v3: Mở tất cả các bottom layer, thêm các FC layer mới để train thì cho kết quả tốt hơn rất nhiều.

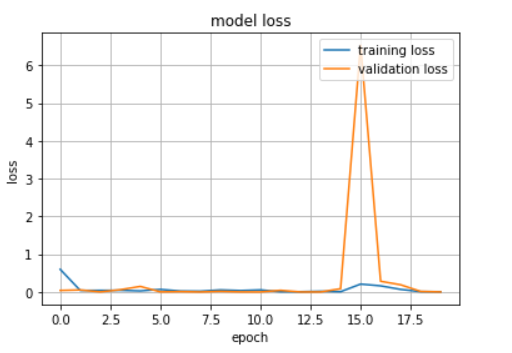
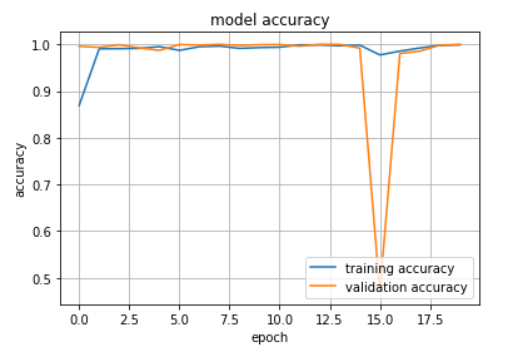
=> độ chính xác cao. Accuracy: ~98%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-score | Support |
| Accuracy |  |  | 0.98 | 20622 |
| Macro avg | 0.98 | 0.97 | 0.97 | 20622 |
| Weighted avg | 0.98 | 0.98 | 0.97 | 20622 |

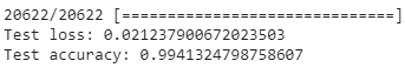


**d. Mô hình resnet152:**

Giống như inception v3,khi đóng băng các bottom layer và chỉ train trên các FC layer mới thêm vào thì mô hình bị hiện tượng overfitting

*  Mô hình resnet152: Khi mở hết các tầng và thêm các FC layer mới để train tất cả thì độ chính xác cao. Accuracy: ~99%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-score | Support |
| Accuracy |  |  | 0.99 | 20622 |
| Macro avg | 0.99 | 0.99 | 0.99 | 20622 |
| Weighted avg | 0.99 | 0.99 | 0.99 | 20622 |



# **V. KẾT LUẬN & HƯỚNG PHÁT TRIỂN**

## **1. Kết luận**

Từ kết quả đánh giá trị thực nghiệm, có thể thấy các model deep learning cho kết quả

rât tốt: precision và recall đều gần bằng 1.

Điều này cho thấy việc sử dụng pre-trained model deep learning đem lại hiệu quả cao.Giúp tiết kiệm thời gian và tài nguyên.

## **2. Hướng phát triển**

Xây dựng thành hệ thống segmentation

Detect được các loại trái cây khác nhau khi có nhiều loại trong cùng một ảnh

# **VI. PHÂN CÔNG CÔNG VIỆC**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Họ Tên | MSSV | Công việc | Trạng thái |
| 1 | Lê Thanh Tiềm | 16521214 | - Tìm hiểu và áp dụng mô hình VGG để giải quyết bài toán  - Đọc data và visualize data  - Viết báo cáo, slide | Hoành thành 100% |
| 2 | Trần Văn Quang | 16521004 | - Tìm hiểu và áp dụng mô hình Resnet để giải quyết bài toán  - Tìm hiểu một số phương pháp cổ điển có thể giải quyết bài toán  - Viết báo cáo, slide | Hoành thành 100% |
| 3 | Nguyễn Đăng Thịnh | 16521177 | - Tìm hiểu và áp dụng mô hình inception để giải quyết bài toán  - Tìm hiểu cách áp dụng SVM vào bài toán  - Viết báo cáo, slide | Hoành thành 100% |
| 4 | Nguyễn Hữu Sơn | 16521036 | - Tìm hiểu transfer learning, cách tiếp cận bài toán  - Viết slide, báo cáo | Hoành thành 100% |
| 5 | Trần Quốc Dũng | 16520263 | - Tìm hiểu transfer learning, cách tiếp cận bài toán  - Viết slide, báo cáo | Hoành thành 100% |

# **VII. REFERENT**

[1] <https://machinelearningcoban.com/2018/06/22/deeplearning/>

[2] <https://www.kaggle.com/lfriedrich/fruits-360-cnn-autoencoder-knn>

[3] <https://www.researchgate.net/publication/321475443_Fruit_recognition_from_images_using_deep_learning>

[4] <https://forum.machinelearningcoban.com/t/kien-truc-cac-mang-cnn-noi-tieng-phan-1-alex-lenet-inception-vgg/2582>

[5] <https://forum.machinelearningcoban.com/t/tong-hop-transfer-learning/5388?fbclid=IwAR04zZ5f0rBYQST8gQ_Wp2Plpv26Dksq55YuYslD3sA9HzTWIbeKMeaCIOg>

- ...