

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA KHOA HỌC VÀ KỸ THUẬT MÁY TÍNH

---o0o---



BÁO CÁO BÀI TẬP LỚN
KHAI PHÁ DỮ LIỆU

IMAGE CLASSIFICATION
WITH AUGMENTATION

Danh sách nhóm:

<i>Họ và tên</i>	<i>MSSV</i>
Huỳnh Phạm Quốc Anh	1920001
Nguyễn Hoàng Phúc	1927030
Nguyễn Tuấn Kiệt	1927021
Nguyễn Hữu Nghĩa	1927027

Thành phố Hồ Chí Minh, tháng 12/2020

MỤC LỤC

1. Tổng quan tài liệu	1
1.1. K-nearest neighbor	1
1.1.1. Giới thiệu	1
1.1.2. Mô tả kĩ thuật	1
1.1.3. Ưu điểm	2
1.1.4. Nhược điểm.....	3
1.2. Neural network.....	4
1.2.1. Giới thiệu	4
1.2.2. Mô tả kĩ thuật	4
1.2.3. Ưu điểm	6
1.2.4. Nhược điểm.....	6
1.3. Một số kĩ thuật về Augmentation.....	6
1.2.5. Giới thiệu chung.....	6
1.2.6. Một số phương thức data augmentation cơ bản:.....	7
2. Ứng dụng	9
2.1. Ứng dụng phát hiện bất thường để chuẩn đoán bệnh sớm:	9
2.2. Phát hiện lỗi sản phẩm:	10
2.3. Trí tuệ nhân tạo được ứng dụng vào hệ thống camera giám sát và kiểm soát:	11
2.4. Ứng dụng AI để xử lý ảnh viễn thám:.....	12
3. Thực nghiệm:	13
3.1. Giới thiệu về thư viện:.....	13
3.2. Ứng dụng:.....	13
3.3. Thực hiện:	14
3.4. Nhận xét, đánh giá:.....	17

DANH SÁCH HÌNH ẢNH

Hình 1 Bản đồ của KNN.....	2
Hình 2 Mạng nơ ron một lớp, ba đầu vào và một đầu ra	5
Hình 3 Ví dụ về ảnh gốc	7
Hình 4 Ví dụ về ảnh được tăng cường hiệu ứng (I).....	7
Hình 5 Ví dụ về ảnh được tăng cường hiệu ứng (II).....	8
Hình 6 Dùng AI để phân biệt bệnh nhân lao phổi và bệnh nhân nhiễm COVID 19	10
Hình 7 Dùng AI để nhận diện khuôn mặt	11
Hình 8 Dùng AI phát hiện các chi tiết cần quan tâm	12
Hình 9 Thực nghiệm : Ảnh gốc	14
Hình 10 Thực nghiệm: Thử nghiệm khả năng nhận diện	15
Hình 11 Thực nghiệm: Ảnh khuôn mặt sau khi nhận diện	15
Hình 12 Thực nghiệm: Nhận diện tâm trạng của người trong ảnh	16

1. Tổng quan tài liệu

1.1.K-nearest neighbor

1.1.1. Giới thiệu

K-nearest neighbor là một trong những thuật toán supervised-learning đơn giản nhất (mà hiệu quả trong một vài trường hợp) trong Machine Learning. Khi training, thuật toán này *không học* một điều gì từ dữ liệu training (đây cũng là lý do thuật toán này được xếp vào loại lazy learning), mọi tính toán được thực hiện khi nó cần dự đoán kết quả của dữ liệu mới.

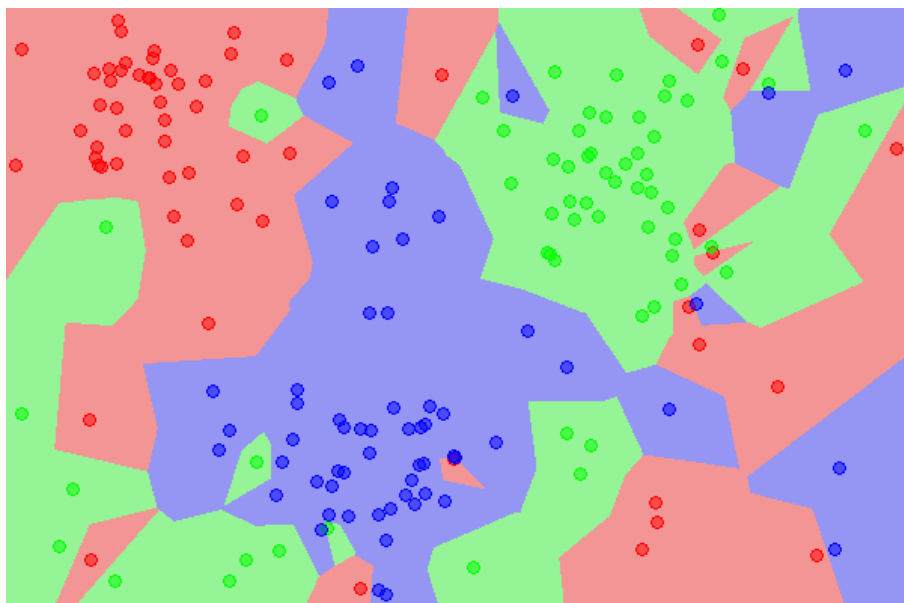
Với KNN, trong bài toán Classification, label của một điểm dữ liệu mới (hay kết quả của câu hỏi trong bài thi) được suy ra trực tiếp từ K điểm dữ liệu gần nhất trong training set. Label của một test data có thể được quyết định bằng major voting (bầu chọn theo số phiếu) giữa các điểm gần nhất, hoặc nó có thể được suy ra bằng cách đánh trọng số khác nhau cho mỗi trong các điểm gần nhất đó rồi suy ra label.

1.1.2. Mô tả kĩ thuật

Trong bài toán Regression, đầu ra của một điểm dữ liệu sẽ bằng chính đầu ra của điểm dữ liệu đã biết gần nhất (trong trường hợp $K=1$), hoặc là trung bình có trọng số của đầu ra của những điểm gần nhất, hoặc bằng một mối quan hệ dựa trên khoảng cách tới các điểm gần nhất đó.

Một cách ngắn gọn, KNN là thuật toán đi tìm đầu ra của một điểm dữ liệu mới bằng cách chỉ dựa trên thông tin của K điểm dữ liệu trong training set gần nó nhất (K-lân cận), không quan tâm đến việc có một vài điểm dữ liệu trong những điểm gần nhất này là nhiễu. Hình dưới đây là một ví dụ về KNN trong classification với $K = 1$.

Ví dụ trên đây là bài toán Classification với 3 classes: Đỏ, Lam, Lục. Mỗi điểm dữ liệu mới (test data point) sẽ được gán label theo màu của điểm mà nó thuộc về. Trong hình này, có một vài vùng nhỏ xem lẫn vào các vùng lớn hơn khác màu. Ví dụ có một điểm màu Lục ở gần góc 11 giờ nằm giữa hai vùng lớn với nhiều dữ liệu màu Đỏ và Lam. Điểm này rất có thể là nhiễu. Dẫn đến nếu dữ liệu test rơi vào vùng này sẽ có nhiều khả năng cho kết quả không chính xác.



Hình 1 Bản đồ của KNN

* Khoảng cách trong không gian vector

Trong không gian một chiều, khoảng cách giữa hai điểm là trị tuyệt đối giữa hiệu giá trị của hai điểm đó. Trong không gian nhiều chiều, khoảng cách giữa hai điểm có thể được định nghĩa bằng nhiều hàm số khác nhau, trong đó độ dài đường thẳng nối hai điểm chỉ là một trường hợp đặc biệt trong đó. Nhiều thông tin bổ ích (cho Machine Learning) có thể được tìm thấy tại Norms (chuẩn) của vector trong tab Math.

1.1.3. Ưu điểm

Trực quan và đơn giản: Thuật toán K-NN rất đơn giản để hiểu và không kém phần dễ thực hiện. Để phân loại điểm dữ liệu mới, thuật toán K-NN đọc toàn bộ tập dữ liệu để tìm ra K lân cận gần nhất.

Không có Training-step: K-NN không xây dựng bất kỳ một mô hình rõ ràng nào, giải thuật này chỉ đơn giản gán thẻ mục nhập dữ liệu mới dựa trên việc học từ dữ liệu lịch sử. Mục nhập dữ liệu mới sẽ được gán thẻ với lớp đa số ở hàng xóm gần nhất.

Dễ phát triển: k-NN là một cách tiếp cận dựa trên bộ nhớ. Việc phân lớp có thể đáp ứng ngay lập tức khi chúng ta cập nhật dữ liệu training mới. Từ đó cho phép thuật toán có thể phản hồi nhanh với những thay đổi của input trong quá trình sử dụng thời gian thực

Dễ thực hiện đối với bài toán nhiều phân lớp: Hầu hết các thuật toán phân loại đều dễ thực hiện cho các bài toán nhị phân và cần nỗ lực để thực hiện cho nhiều lớp trong khi K-NN điều chỉnh thành đa lớp mà không cần nỗ lực thêm.

Có thể được sử dụng cho cả Phân loại và Hồi quy: Một trong những ưu điểm lớn nhất của K-NN là K-NN có thể được sử dụng cho cả các bài toán phân loại và hồi quy.

Nhiều tiêu chí khoảng cách để lựa chọn: Thuật toán K-NN cho phép người dùng linh hoạt trong việc lựa chọn khoảng cách trong khi xây dựng mô hình K-NN.

K-NN không có giả định: K-NN là một thuật toán phi tham số. Các mô hình có tham số như hồi quy tuyến tính có rất nhiều điều kiện phải được dữ liệu đáp ứng trước khi có thể được thực hiện, điều này không đúng với K-NN.

1.1.4. Nhược điểm

K-NN là thuật toán xử lý chậm: K-NN có thể rất dễ thực hiện nhưng khi tập dữ liệu tăng hiệu quả hoặc tốc độ của thuật toán giảm rất nhanh.

Bất lợi về số chiều của bài toán: KNN hoạt động tốt với số lượng biến đầu vào nhỏ nhưng khi số lượng biến tăng lên, thuật toán K-NN gặp khó khăn trong việc dự đoán đầu ra của điểm dữ liệu mới.

K-NN cần các đối tượng địa lý đồng nhất: Nếu chúng ta quyết định xây dựng k-NN bằng cách sử dụng một khoảng cách chung, như khoảng cách Euclidean hoặc Manhattan, thì hoàn toàn cần thiết các đối tượng địa lý phải có cùng tỷ lệ, vì sự khác biệt tuyệt đối về trọng số như nhau, tức là, một khoảng cách trong đối tượng địa lý 1 phải có nghĩa tương tự đối với đối tượng địa lý 2.

Số lượng láng giềng tối ưu: Một trong những vấn đề lớn nhất với K-NN là chọn số lượng láng giềng tối ưu được xem xét trong khi phân loại mục nhập dữ liệu mới.

Vấn đề khi dữ liệu không cân bằng: k-NN hoạt động không tốt trên dữ liệu mất cân bằng. Nếu chúng ta xem xét hai lớp, A và B, và phần lớn dữ liệu huấn luyện được gán nhãn là A, thì cuối cùng mô hình sẽ ưu tiên nhiều hơn cho A. Điều này có thể dẫn đến việc phân loại sai lớp B ít phổ biến hơn.

Outliner sensitivity: Thuật toán K-NN rất nhạy cảm với các giá trị ngoại lệ vì nó chỉ chọn các lân cận dựa trên tiêu chí khoảng cách.

Xử lý missing-value: K-NN vốn dĩ không có khả năng xử lý vấn đề về missing-value

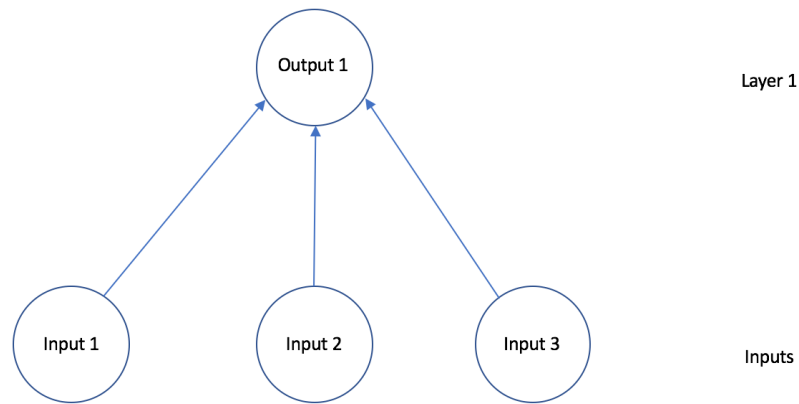
1.2. Neural network

1.2.1. Giới thiệu

Mạng thần kinh là một phương pháp để phân rã một hệ thống phức tạp thành các phần đơn giản hơn để hiện thực hóa nó. Tập hợp các nút và liên kết giữa các nút là thành phần của mạng. Các nút được gọi là đơn vị tính toán của mạng và các liên kết xác định luồng thông tin giữa các nút. NN là một trong những mạng xem các nút là nơron nhân tạo. Tế bào thần kinh nhân tạo là một mô hình tính toán được lấy cảm hứng từ các tế bào thần kinh tự nhiên.

1.2.2. Mô tả kĩ thuật

Bộ não con người bao gồm 100 tỷ tế bào được gọi là tế bào thần kinh, được liên kết với nhau bằng các khớp thần kinh. Nếu đầu vào đủ khớp thần kinh kích hoạt một tế bào thần kinh, tế bào thần kinh đó cũng sẽ kích hoạt. Chúng ta gọi quá trình này là “suy nghĩ”. Chúng ta có thể mô hình hóa quá trình này bằng cách tạo một mạng nơ-ron trên máy tính. Một mạng nơ-ron có các nơ-ron đầu vào và đầu ra, được liên kết với nhau bằng các khớp thần kinh có trọng số. Các trọng số ảnh hưởng đến mức độ lan truyền thuận đi qua mạng nơron. Sau đó, trọng số có thể được thay đổi trong quá trình truyền ngược - đây là phần mà mạng nơ-ron hiện đang nghiên cứu. Quá trình truyền tiến và truyền ngược này được tiến hành lặp đi lặp lại trên mỗi phần dữ liệu trong tập dữ liệu huấn luyện. Kích thước của tập dữ liệu càng lớn và càng có nhiều tập dữ liệu thì mạng nơ-ron sẽ học được càng nhiều và mạng nơ-ron càng có được kết quả dự đoán đầu ra tốt hơn.



Hình 2 Mạng nơ-ron một lớp, ba đầu vào và một đầu ra

Mạng nơ-ron trên hình có một lớp, ba đầu vào và một đầu ra. Trên thực tế, mạng nơ-ron có thể có bất kỳ số lượng lớp, đầu vào hoặc đầu ra nào. Nói một cách đơn giản, mạng nơ-ron là một đồ thị được liên kết với các nơ-ron đầu vào, nơ-ron đầu ra và các cạnh có trọng số. Hãy đi vào chi tiết về một số thành phần sau:

1) Tế bào thần kinh. Mạng nơ-ron là một biểu đồ của các nơ-ron. Một nơ-ron có đầu vào và đầu ra. Tương tự, một mạng nơ-ron có các đầu vào và đầu ra. Các đầu vào và đầu ra của mạng nơ-ron được biểu diễn bằng các nơ-ron đầu vào và nơ-ron đầu ra. Các tế bào thần kinh đầu vào không có tế bào thần kinh tiền nhiệm, nhưng có một đầu ra. Tương tự, một tế bào thần kinh đầu ra không có tế bào thần kinh kế thừa, nhưng có đầu vào.

2) Liên kết và Trọng số. Mạng nơ-ron bao gồm các liên kết, mỗi liên kết chuyển đầu ra của một nơ-ron đến đầu vào của một nơ-ron khác. Mỗi liên kết được gán một trọng số.

3) Lan truyền. Hàm truyền tính toán đầu vào của một nơ-ron từ các đầu ra của các nơ-ron tiền nhiệm. Chức năng truyền giống được tận dụng trong giai đoạn đào tạo truyền về phía trước.

4) Quy tắc học. Quy tắc học là một hàm sửa đổi trọng số của các liên kết. Điều này phục vụ để tạo ra một đầu ra ưa thích cho một đầu vào nhất định cho mạng nơ-ron. Quy tắc học được tận dụng trong giai đoạn truyền ngược của quá trình đào tạo.

1.2.3. Ưu điểm

Mạng nơ-ron rất linh hoạt và có thể được sử dụng cho cả bài toán hồi quy và phân loại: bất kỳ dữ liệu nào có thể ở dạng số đều có thể được sử dụng trong mô hình, vì mạng nơ-ron là một mô hình toán học với các hàm xấp xỉ.

Mạng nơ-ron rất tốt để lập mô hình với dữ liệu phi tuyến với số lượng đầu vào lớn: ví dụ, hình ảnh. Nó đáng tin cậy trong cách tiếp cận các nhiệm vụ liên quan đến nhiều tính năng. Nó hoạt động bằng cách chia nhỏ vấn đề phân loại thành một mạng nhiều lớp gồm các phần tử đơn giản hơn.

- Sau khi được trained, các dự đoán diễn ra khá nhanh.
- Mạng nơ-ron có thể được đào tạo với bất kỳ số lượng đầu vào và lớp nào.
- Mạng nơ-ron hoạt động tốt nhất với khi có càng nhiều dữ liệu.
- Xử lý tốt dữ liệu mơ hồ, không đầy đủ

1.2.4. Nhược điểm

Mạng nơ-ron là hộp đen, có nghĩa là chúng ta không thể biết mỗi biến độc lập đang ảnh hưởng đến các biến phụ thuộc bao nhiêu.

Về mặt tính toán, đào tạo với các CPU truyền thống rất tốn kém và mất thời gian.

Mạng nơ-ron phụ thuộc rất nhiều vào dữ liệu huấn luyện. Điều này dễ dẫn đến vấn đề quá khớp (over-fitting) và phổ thông (generalization). Các vấn đề này phụ thuộc vào dữ liệu vào tập dữ liệu training và có thể dẫn đến điều chỉnh dữ liệu

Người dùng có rất ít ảnh hưởng đến chức năng của mạng (trọng số và thresholds).

Có những lựa chọn thay thế đơn giản hơn, nhanh hơn, dễ đào tạo hơn và mang lại hiệu suất tốt hơn (svm, cây quyết định, hồi quy)

Mạng nơ-ron nhiều lớp thường khó đào tạo và yêu cầu điều chỉnh nhiều tham số





1.3. Một số kỹ thuật về Augmentation

1.2.5. Giới thiệu chung

Hiện nay trong deep learning thì vấn đề dữ liệu có vai trò rất quan trọng. Chính vì vậy có những lĩnh vực có ít dữ liệu để cho việc train model thì rất khó để tạo ra được kết quả tốt trong việc dự đoán. Do đó người ta cần đến một kỹ thuật gọi là tăng cường dữ liệu





(data augmentation) để phục vụ cho việc nếu bạn có ít dữ liệu, thì bạn vẫn có thể tạo ra được nhiều dữ liệu hơn dựa trên những dữ liệu bạn đã có. Ví dụ như hình dưới, đó là các hình được tạo ra thêm từ một ảnh gốc ban đầu.

1.2.6. Một số phương thức data aummentation cơ bản:

Original	Flip	Rotation	Random crop
			
<ul style="list-style-type: none"> Image without any modification 	<ul style="list-style-type: none"> Flipped with respect to an axis for which the meaning of the image is preserved 	<ul style="list-style-type: none"> Rotation with a slight angle Simulates incorrect horizon calibration 	<ul style="list-style-type: none"> Random focus on one part of the image Several random crops can be done in a row
image.jpg 904x367 149 KB			

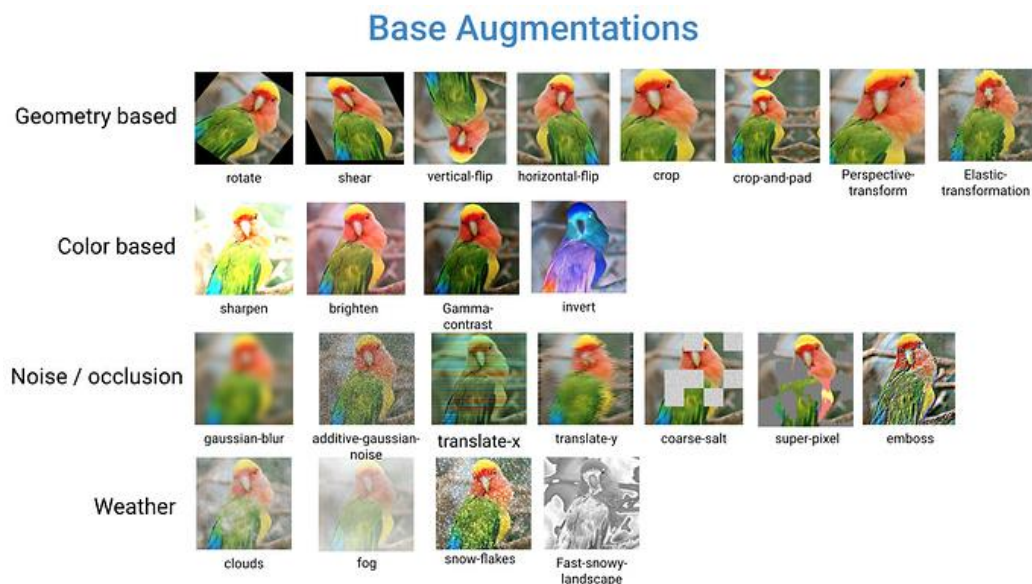
Hình 3 Ví dụ về ảnh gốc

- Original (Ảnh gốc)
- Flip (Lật): lật theo chiều dọc, ngang miễn sao ý nghĩa của ảnh (label) được giữ nguyên hoặc suy ra được. Ví dụ nhận dạng quả bóng tròn, thì mình lật kiểu gì cũng ra quả bóng. Còn với nhận dạng chữ viết tay, lật số 8 vẫn là 8, nhưng 6 sẽ thành 9 (theo chiều ngang) và không ra số gì theo chiều dọc. Còn nhận dạng ảnh y tế thì việc bị lật trên xuống dưới là không bao giờ xảy ra ở ảnh thực tế --> không nên lật làm gì
- Random crop (Cắt ngẫu nhiên): cắt ngẫu nhiên một phần của bức ảnh. Lưu ý là khi cắt phải giữ thành phần chính của bức ảnh mà ta quan tâm. Như ở nhận diện vật thể, nếu ảnh được cắt không có vật thể, vậy giá trị nhãn là không chính xác.

Color shift	Noise addition	Information loss	Contrast change
			
<ul style="list-style-type: none"> Nuances of RGB is slightly changed Captures noise that can occur with light exposure 	<ul style="list-style-type: none"> Addition of noise More tolerance to quality variation of inputs 	<ul style="list-style-type: none"> Parts of image ignored Mimics potential loss of parts of image 	<ul style="list-style-type: none"> Luminosity changes Controls difference in exposition due to time of day

Hình 4 Ví dụ về ảnh được tăng cường hiệu ứng (I)

- Color shift (Chuyển đổi màu): Chuyển đổi màu của bức ảnh bằng cách thêm giá trị vào 3 kênh màu RGB. Việc này liên quan tới ảnh chụp đôi khi bị nhiễu --> màu bị ảnh hưởng.
- Noise addition (Thêm nhiễu): Thêm nhiễu vào bức ảnh. Nhiễu thì có nhiều loại như nhiễu ngẫu nhiên, nhiễu có mẫu, nhiễu cộng, nhiễu nhân, nhiễu do nén ảnh, nhiễu mờ do chụp không lấy nét, nhiễu mờ do chuyển động... có thể kể hết cả ngày.
- Information loss (Mất thông tin): Một phần của bức hình bị mất. Có thể minh họa trường hợp bị che khuất.
- Contrast change (Đổi độ tương phản): thay độ tương phản của bức hình, độ bão hòa



Hình 5 Ví dụ về ảnh được tăng cường hiệu ứng (II)

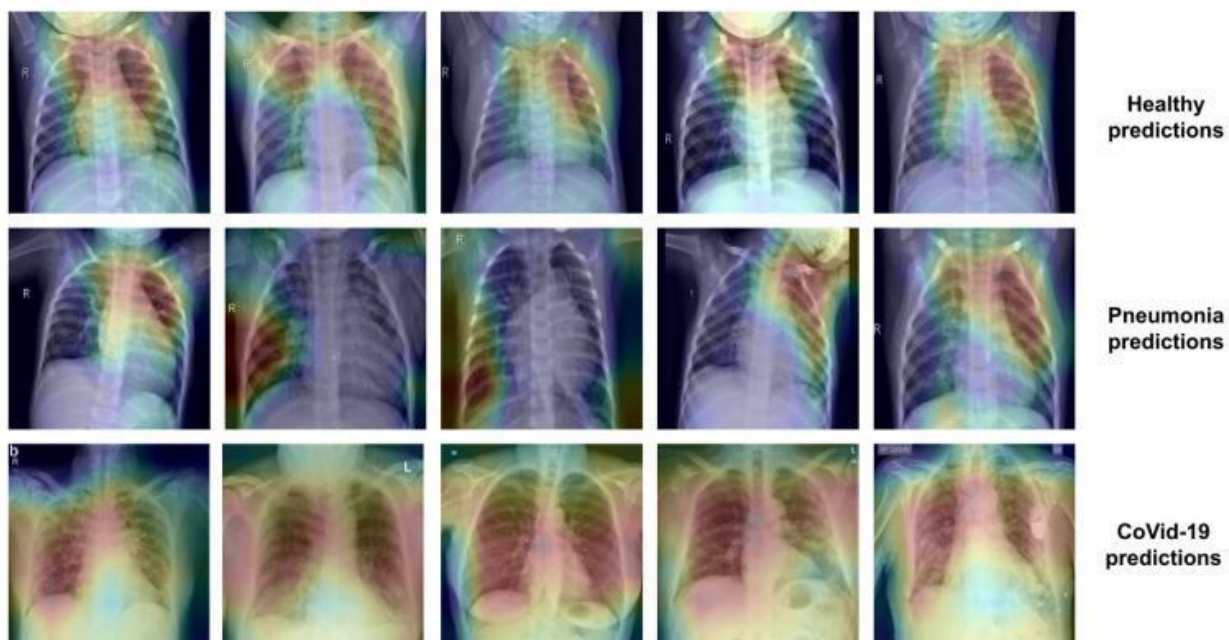
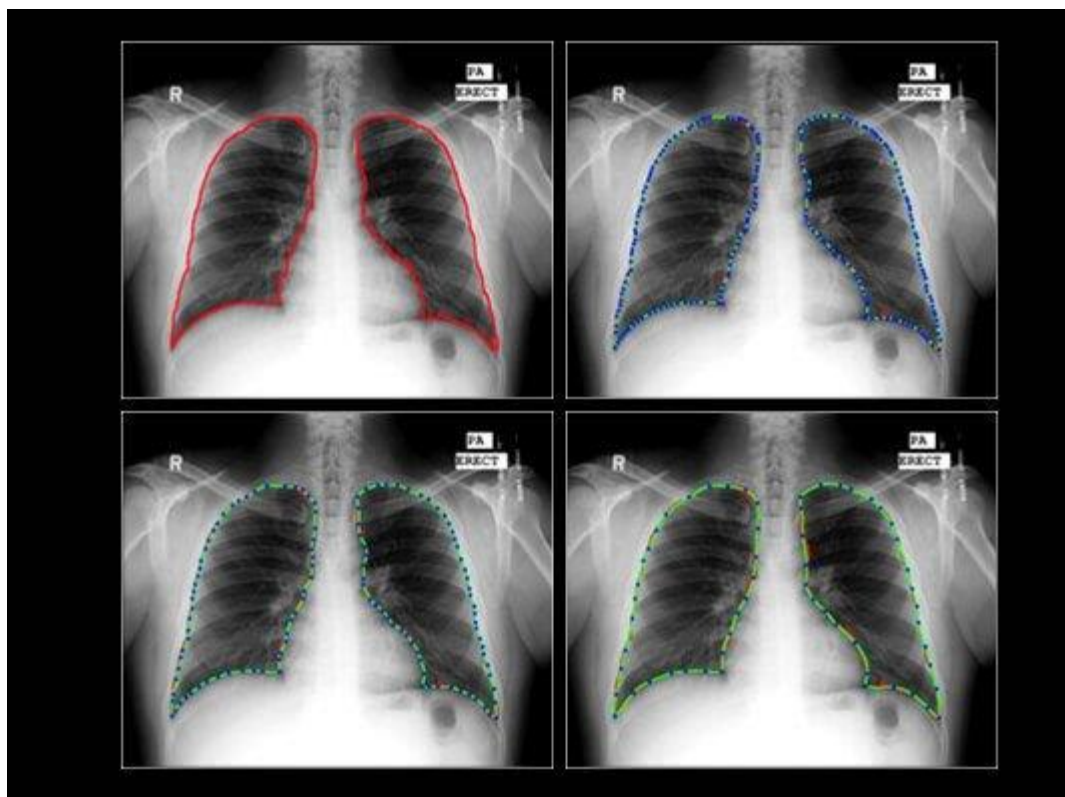
- Geometry based: Đủ các thể loại xoay, lật, scale, padding, bóp hình, biến dạng hình,
- Color based: giống như trên, chi tiết hơn chia làm (i) tăng độ sắc nét, (ii) tăng độ sáng, (iii) tăng độ tương phản hay (iv) đổi sang ảnh negative - âm bản.
- Noise/occlusion: Chi tiết hơn các loại nhiễu, như mình kể trên còn nhiều lắm. kể hết rưng rưng.
- Whether: thêm tác dụng cảnh thời tiết như mưa, tuyết, sương mờ, ...

2. Ứng dụng

2.1. Ứng dụng phát hiện bất thường để chuẩn đoán bệnh sớm:

Việc phát triển mạnh mẽ các kỹ thuật dựng ảnh trong y học đã tạo ra một lượng dữ liệu y học to lớn mà con người khó có thể xử lý tốt được. Vì thế, việc sử dụng AI để hỗ trợ cho con người tìm ra những thông tin hữu ích một cách nhanh chóng là một bước đi cần thiết và quan trọng để phát triển ngành y cũng như tăng khả năng chữa trị thành công cho các bệnh nhân. Cũng giống như bác sỹ cần nhiều năm học tập và làm việc để có thể tích lũy kinh nghiệm trong việc chẩn đoán và đưa ra phác đồ điều trị tốt cho bệnh nhân, các hệ thống AI cũng cần được “huấn luyện” để có khả năng sử dụng các thông tin đã học, từ đó để đưa ra kết luận phù hợp.

Các nhà nghiên cứu sử dụng các bộ ảnh chụp CT nội tạng đã được các chuyên gia “đánh dấu” để huấn luyện cho hệ thống AI. Sau khi được huấn luyện, hệ thống AI có khả năng tự đánh dấu các nội tạng ở các ảnh CT khác với độ chính xác phù hợp (phụ thuộc vào dữ liệu và phương pháp huấn luyện) mà không cần sự can thiệp của con người. Không chỉ dừng lại ở việc phân biệt các vùng nội tạng, mà nhiều nhóm nghiên cứu đã phát triển các chương trình AI khác như: Đánh dấu các phần có dấu hiệu ung thư trong ảnh gan, phổi ... phân biệt các vùng khác nhau hay phát hiện các vùng bất thường của não trong ảnh MRI thông qua phân đoạn các vùng trong não. Đặc biệt trong trường hợp ảnh nhiều chiều, nhiều lớp ảnh, việc xử lý tất cả các lớp ảnh là một công việc tốn nhiều thời gian và công sức đối với bác sỹ. Vì thế, việc sử dụng AI sẽ hỗ trợ họ rất nhiều, đồng thời góp phần mang lại chất lượng chữa trị tốt hơn cho người bệnh.



Hình 6 Dùng AI để phân biệt bệnh nhân lao phổi và bệnh nhân nhiễm COVID 19

2.2. Phát hiện lỗi sản phẩm:

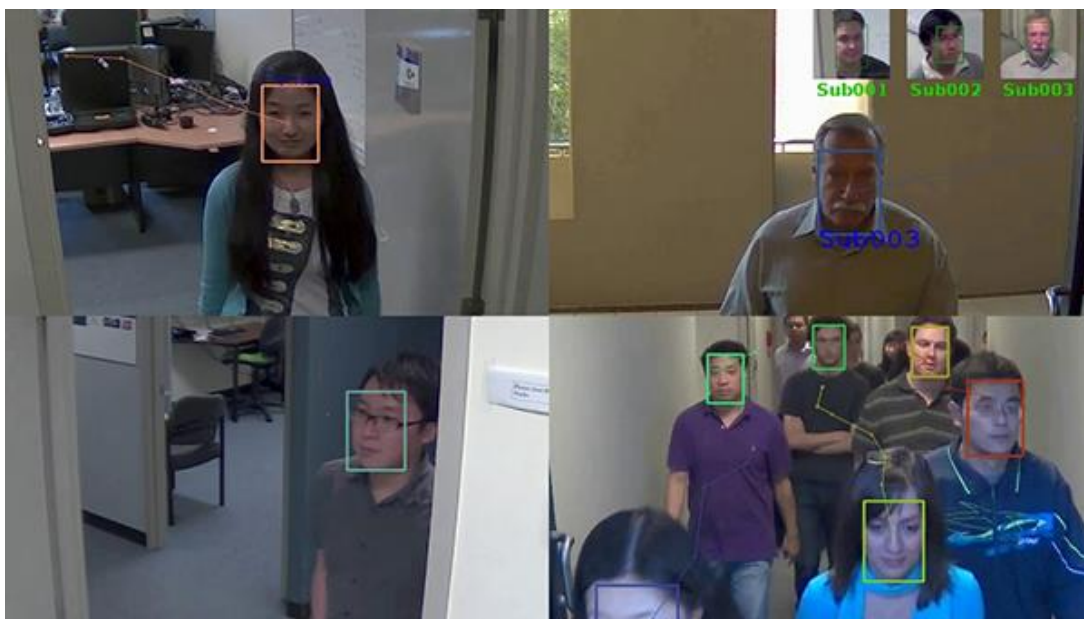
Ngày nay, nhiều dây chuyền lắp ráp không có hệ thống hoặc công nghệ để xác định lỗi trên dây chuyền sản xuất. Ngay cả những thứ có sẵn tại chỗ cũng chỉ rất cơ bản,

đòi hỏi các kỹ sư lành nghề phải xây dựng và hard-code các thuật toán để phân biệt giữa các phần khiếm khuyết và các phần hoàn thiện. Phần lớn các hệ thống này vẫn không thể tự tìm hiểu hoặc dung nạp thông tin mới, dẫn đến vô số các xác nhận sai thực tế (về tình trạng hàng hóa), sau đó phải được nhân viên tại chỗ kiểm tra thủ công.

Bằng cách phát triển hệ thống này với trí thông minh nhân tạo và khả năng tự học, các nhà sản xuất có thể tiết kiệm vô số giờ sản xuất bằng cách giảm đáng kể các thông tin sai lệch và số giờ cần thiết để kiểm soát chất lượng.

2.3.Trí tuệ nhân tạo được ứng dụng vào hệ thống camera giám sát và kiểm soát:

Đối với hệ thống camera an ninh truyền thống thì những hoạt động can thiệp thủ công cần thiết sẽ được thực hiện khi có những hành vi vi phạm xảy ra. Công nghệ trí tuệ nhân tạo được tích hợp trên camera sẽ thay đổi tất cả điều này bằng cách nhận ra các mối đe dọa tiềm tàng trước khi ảnh hưởng đến con người và tài sản, cho phép đội quản lý an ninh can thiệp ngay lập tức khi phát hiện sự cố xảy ra.



Hình 7 Dùng AI để nhận diện khuôn mặt

Ai có thể dựa trên những thay đổi bất thường của dòng hình ảnh được ghi lại qua camera để cảnh báo xâm nhập trái phép, hiệu cảnh báo khi có người lạ đột nhập vào vùng quan sát của camera để người dùng có thể nhận biết và can thiệp sớm nhất có thể.

2.4. Ứng dụng AI để xử lý ảnh viễn thám:

Trong viễn thám, việc quan trọng nhất là phân biệt được các đối tượng có trong ảnh. Việc làm này nếu được làm một cách thủ công do con người thì sẽ rất mất thời gian và công sức, bên cạnh đó vẫn tồn tại sai sót do hạn chế về mắt người.

Machine Learning đã giúp máy tính thực thi những việc tưởng chừng như không thể vào: phân loại cả ngàn vật thể khác nhau trong các bức ảnh, tự tạo chú thích cho ảnh, bắt chước giọng nói và chữ viết của con người, giao tiếp với con người, hay thậm chí cả sáng tác văn hay âm nhạc. Vì vậy, việc ứng dụng công nghệ này vào viễn thám là một việc làm tất yếu.



Hình 8 Dùng AI phát hiện các chi tiết cần quan tâm

3. Thực nghiệm:

Dùng thư viện Face detection with OpenCV and deep learning để nhận diện khuôn mặt và thư viện Face Emotion Detection with Deep Learning để phân loại cảm xúc của người trong ảnh đã được nhận diện.

3.1. Giới thiệu về thư viện:

OpenCV's deep learning face detector: dựa trên Single Shot Detector (SSD) với mạng cơ sở ResNet (không giống như các SSD OpenCV khác đã thấy thường sử dụng MobileNet làm mạng cơ sở). Mạng cơ sở được huấn luyện sử dụng Caffe Deep Learning framework. Nếu muốn tìm hiểu sâu hơn cách thức hoạt động có thể tham khảo qua đường link <https://www.pyimagesearch.com/2018/02/26/face-detection-with-opencv-and-deep-learning/> và <https://caffe.berkeleyvision.org/>

Face Emotion Detection with Deep Learning:
https://colab.research.google.com/drive/1SBHyy-4iuslEdtcLX26AbzW_-a1CclFm

3.2. Ứng dụng:

Nhận biết cảm xúc dự đoán khách hàng tiềm năng: cảm xúc **vui, buồn, bất ngờ, háo hức...** trong quá trình giới thiệu sản phẩm, dịch vụ nhằm phân loại khách hàng tiềm năng.

Nhận biết cảm xúc từ đó dự đoán hành vi: Ngân hàng, hoặc các nơi giao dịch tiền tệ, nhận diện và phân loại cảm xúc dự đoán hành vi gây nguy hiểm để tiến hành theo dõi.

Và còn rất nhiều ứng dụng khác được phát triển từ việc phân loại cảm xúc. Qua đó phần thực nghiệm sẽ mô tả quá trình phân loại cảm xúc các gương mặt đã được nhận diện trong 1 bức ảnh đơn giản.

3.3.Thực hiện:

Quá trình thực hiện qua các bước sau:

1. Import ảnh vào ứng dụng Colaboratory in Google:

```
#image = cv2.imread(image_file, cv2.IMREAD_UNCHANGED)
```

```
image = cv2.imread(image_file)
```

```
# resize it to have a maximum width of 400 pixels
```

```
image = imutils.resize(image, width=400)
```

```
(h, w) = image.shape[:2]
```

```
print(w,h)
```

```
cv2_imshow(image)
```



Hình 9 Thực nghiệm : Ảnh gốc

2. Thêm 2 thư viện đã giới thiệu ở trên:

```
!wget -N
```

```
https://raw.githubusercontent.com/opencv/opencv/master/samples/dnn/face\_detector/deploy.prototxt
```

```
!wget -N
```

```
https://raw.githubusercontent.com/opencv/opencv\_3rdparty/dnn\_samples\_face\_detector\_20170830/res10\_300x300\_ssd\_iter\_140000.caffemodel
```

```

prototxt = 'deploy.prototxt'
model = 'res10_300x300_ssd_iter_140000.caffemodel'
net = cv2.dnn.readNetFromCaffe(prototxt, model)

```

3. Khuôn mặt đã được nhận diện



Hình 10 Thực nghiệm: Thử nghiệm khả năng nhận diện

Chưa nhận diện được khuôn mặt đeo mask

4. Chi tiết các khuôn mặt đã được nhận diện



Hình 11 Thực nghiệm: Ảnh khuôn mặt sau khi nhận diện

5. Dự đoán cảm xúc cho các khuôn mặt đã được nhận diện

```

from keras import models
emotion_dict= {'Angry': 0, 'Sad': 5, 'Neutral': 4, 'Disgust': 1, 'Surprise': 6, 'Fear': 2,
'Happy': 3}
model = models.load_model(path_Test + "model_v6_23.hdf5")

for face_image in FaceList:
    ShowImage([face_image], 1, 5)

```

```

face_image = cv2.resize(face_image, (48,48))
face_image = cv2.cvtColor(face_image, cv2.COLOR_BGR2GRAY)
face_image = np.reshape(face_image, [1, face_image.shape[0], face_image.shape
[1], 1])

```

```

predicted_class = np.argmax(model.predict(face_image))
label_map = dict((v,k) for k,v in emotion_dict.items())
predicted_label = label_map[predicted_class]
print("Emotion :", predicted_label)

```



Emotion : Happy



Emotion : Surprise

Hình 12 Thực nghiệm: Nhận diện tâm trạng của người trong ảnh

Mã nguồn tham khảo trong file đính kèm.

3.4.Nhận xét, đánh giá:

Quá trình nhận diện khuôn mặt diễn ra khá nhanh với khả năng suy đoán cảm xúc của khuôn mặt được nhận diện. Tuy vậy, để quá trình nhận diện được diễn ra tốt nhất cần nhiều điều kiện khác nhau như: ảnh cần có chất lượng tốt, ánh sáng đủ, vật thể cần nhận diện không bị che khuất hoặc chỉ bị che khuất một phần nhỏ, phong nền càng đơn giản càng tốt (ví dụ màu đơn sắc) ...

Trong quá trình thực nghiệm trong bài tập lớn này, quá trình nhận diện khuôn mặt còn gặp nhiều trở ngại vì chất lượng ảnh kém, góc nhìn chưa được tốt, bị các vật thể khác che 1 phần khuôn mặt, quá nhiều đối tượng nhiễu trong ảnh. Quá trình nhận diện cảm xúc cũng chưa được tốt, so với cảm xúc thật lúc chụp ảnh chỉ chính xác khoảng 50%.

Hiện nay mặt dù các ứng dụng ứng dụng trí tuệ nhân tạo để khai phá dữ liệu hiện đang phát triển rất nhanh, tuy nhiên chúng vẫn còn bị một số hạn chế như trong bài tập này. Việc xử dụng AI hoàn toàn cho các công việc cần độ chính xác cao cũng nên cân nhắc, nhất là khi các dữ liệu nền bị nhiễu mạnh. Tuy vậy, ứng dụng các công nghệ tiên tiến của khai phá dữ liệu và kết hợp với AI để thực hiện các công việc trợ giúp cho con người là hoàn toàn khả thi.

Hiện nay đã có nhiều nơi sử dụng dữ liệu lớn và AI để giúp công việc hiện tại dễ dàng hơn. Tại Việt Nam, nhiều cơ sở đã xử dụng nhận dạng hình ảnh để giúp các bác sỹ tìm các điểm bất thường nhanh hơn, tạo cơ sở dữ liệu người bệnh để tra cứu và gợi ý các bệnh/đơn thuốc cho bệnh nhân dựa trên các dữ liệu trước đó của chính bệnh nhân đó kết hợp với thông tin dịch tễ tại địa phương...

Với sự phát triển không ngừng của khoa học hiện đại, các hạn chế mà khai phá dữ liệu hiện tại đang gặp phải đang dần giảm bớt (máy ảnh chất lượng cao hơn, vi xử lý mạnh hơn, dữ liệu “sạch” hơn...). Tin rằng một ngày không xa, việc ứng dụng khai phá dữ liệu để thực hiện các công việc tự động hoàn toàn sẽ trở thành hiện thực.