**BUILDING FACE MASK DETECTION SYSTEM**

**Lê Tấn Lộc, Huỳnh Thanh Phong,**

**Trần Thị Tứ, Linh, Phạm Duy Minh**

(Sinh viên năm 2, Khoa *Công Nghệ Thông Tin*)

Instructor*: ThS. Lương Trần Ngọc Khiết*

**Abstract**: Convolutional Neural Networks (CNN) is the most popular neural network model being used for image classification problem. It help us to build systems to detect object and give back the result with accuracy . The goal of this project is focus on how to detect people who are wear face mask and who are not to stop the spread of coronavirus

**Key word**: Convolutional Neural Network, Deep Learning, Face Mask Detection, MobileNetV2, COVID-19

# An Introduction to face mask detection

## Introduction

Our world right now is face on a realy serious situation when the new virus call COVID 19 keep spreading to fast. From that we can see how important to control the crowd, control people who are wear face mask and who are not to against the virus. And CNN model is a way to solve this problem

The idea to build model to detect people wear face mask with hight accuracy is realy need right now and many researcher are realy care this idea. Especially for countrys are loosing control the virus

## Object Detection

About Computer Science when it allow us to make computer or machine do some task that human can do like create eye for computer to see and detect object on photos or video like normal human can do

Thank to improve in AI (Artificial Intelligence) and research method to neural network. That can help us to solve this problem easily

The important thing to improve the sight of computer is the data that help us to train and improve the sight of computer . It have important roll in detection problem

Through the sight of computer, it will detect some feature on the face in the photo and compare in database to know what the face look like. From that it can detect identify that person just by see the face

Define detection problem:

* Step 1: Input an image.
* Step 2: Analysis the image to know faces location
* Step 3: Check if face area have a face mask or not
* Step 4: Give the result

|  |  |  |
| --- | --- | --- |
| A picture containing person, person, posing  Description automatically generated | A person wearing glasses  Description automatically generated with medium confidence | 🡺Without mask |
| A picture containing person  Description automatically generated | 🡺With mask |

Image ‑ Simulate detection problem.

From that we can use this system on some device like camera to check people use mask in real time

## Reason choosing project

Convolutional neural networks has been use in may problem about identify object. But if we want to use CNN for our own problem we have to build model with a big number of layers and each layer have their own analysis problem

Suggest build model using CNN to see a person have a face mask or not. Especially right now when COVID 19 keep spearding and it very easy to lost the track of the virus if we control the crowd not carefully. This is the big issue that it will more bigger in many country that realy need this system

# Research purposes, research subject, scope of the research, methodology

## Research purposes

In this research, we focus on how to solve the face mask detection to prevent the disease. The main purpose of this project is to analysis, improve for detect people who wear the mask. And build the system with hight effective and hight accuracy to support in control the crowd

## Research subject

Project has chosen research subject to research:

* Detection problem.
  + Face detection
  + Mask detection
* Image dataset.
  + Statistical data
  + Analysis data
* Convolutional neural networks.
  + Analysis features in image
  + Detect features in image

## Scope of the research

Scope of this project is detect people wear the face mask through device like camera in real time.

This system can be use in supermarket, convenience store, school … Subject has to achieve a condition is has more 2/3 face in the frame so the system can detect.

## Methodology

All research method include is:

* Simulate and build the model base on CNN for detect people who wear the mask and who not in public place
* Setting and experiment on computer.
* Analysis, evaluate and conclude.

# Research result

## Dataset

* **Condition for input**:

Gather the data have to achieve some condition: have more 2/3 of face, have enough the light and the quality of image is 360p or above



Image ‑ Some correct data.

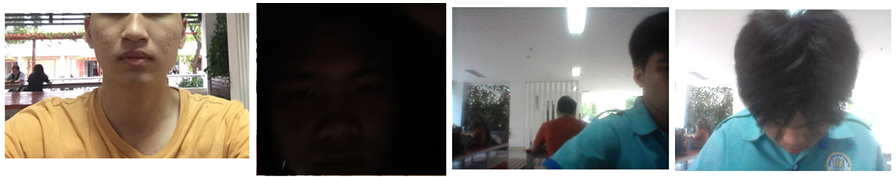


Image ‑ Some incorrect data

* Gather the data: attached labels on the images, organize into folder tree
* Dataset gather from 2 source are Sriman\_Mitra and our own data has been made by each member. Because the result we want for this problem is ‘Yes’ or ‘No’ so we have seperated into 2 folder is with\_mask ; without\_mask and contained in folder call dataset

|  |  |
| --- | --- |
|  |  |

* Attached labels with condition: Attribute\_Status(nếu có)\_Index :
* ***Attribute:*** will be ‘*mask*’ if that image has a mask on or ‘*notmask*’ if it not have the mask .
* ***Status:*** will be ‘multi’ if image has many face in it.
* ***Index:*** is interger with 4 digit.

|  |  |
| --- | --- |
| Folder with\_mask | Folder without\_mask |
| mask\_0003.jpg | notmask\_0003.jpg |
|  |  |
|  |  |
| Image ‑ Data with subject has mask | Image ‑ Data with subject has no mask |

* Divided into Training, test sets:

|  |  |
| --- | --- |
| Table ‑ Statisical amount of image in training and testing. | Table ‑ Parameters in the model. |
| |  |  |  | | --- | --- | --- | |  | Training (80%) | Testing (20%) | | With mask | 1647 | 412 | | Without mask | 1598 | 399 | | |  |  | | --- | --- | | **Parameter** | **Value** | | INIT\_LR |  | | EPOCHS | 20 | | BS | 32 | |

* **Convolutional Neural Networks**

Convolutional Neural Networks (CNN) is the big topic in machine learning and AI, CNN is a network that connect all layers together from the input layer to output layer and layers between 2 of them is called hidden layers

CNN is amount of Convolution layers stack up together and use nonlinear activation like ReLU and tanh to caculate weight from nodes. Each layer will find some features and will through next layer will find some other features. So all input neuron in the next layer will take the ouput from the previous layer has been filter. Each layer use these filter different to find feature and usually each layer has 100 or 1000 filter to find more feature

Beside those layer, there are another layers for another work like pooling or subsmapling layer use to reduce or more precisely use for throw away some infromation is useless.

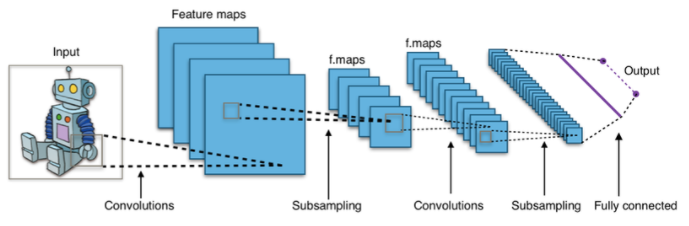


Image ‑ Structure of Convolution Neural Network - CNN.

* **Libary.**

Main libary:

* ***TensorFlow*** is an end-to-end open source platform for machine learning It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications. And Tensorflow is backend for library Keras to work.
* ***Keras*** is an API designed for human beings, not machines run on backend is Tensorflow. In this project Keras focus on handling image and frame from camera into array numpy so that it can be fit input shape of model
* ***OpenCV-Python*** is library handling image and read model of face detection, draw box around the face and write result on the image or frame from camera in real time. After that it will show on the screen

Libary support :

* ***Numpy*** provide formular, array and matrix.
* ***Matplotlib*** is graph libary.

## Model of face mask detection

### Model

* **Structure**

Project use model call MobileNetV2

MobileNetV2 is improvement version of model MobileNetV1 use for build apps that use computer vision run on phone.

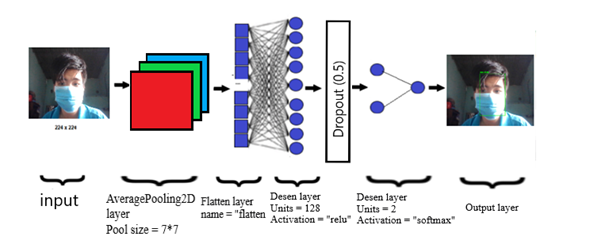


Image ‑ Mô hình huấn luyện của đề tài.

* AveragePooling2D layer: is layer that gather some important feature from the image size 224 x 224 x 3pixel reduce to 32 x 32 x 3 pixel (Pool size = 7 \* 7 that mean it will gather important features and the width, height will reduce 7 times).
* Flatten layer: is layer that flatten the result from AveragePooling2D layer into a hot vector.
* Desen layer: is layer that neuron connect to neuron from previous layer.
  + Units: It defines the output shape
  + Activation is parameter to call activate function like ReLU, tanh, Softmax…

|  |  |
| --- | --- |
|  | . |
| Image ‑ Graph for function ReLU | Image ‑ Graph for function Softmax |

* The output layer is responsible for producing the final result.

### Result.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Chart, line chart  Description automatically generated | |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | Precision | Recall | F1-score | Support | | With\_mask | 0.98 | 0.99 | 0.99 | 412 | | Without\_mask | 0.99 | 0.98 | 0.99 | 399 | | Accuracy |  |  | 0.99 |  | |
| Image 3‑9 Training and test Loss and Accuracy. | Table 3‑3 Result after training. |

This system has been train on many situation and can detect one or multiple people on one image or frame from camera run on real time.

The way system detect people have mask using camera run on real time:

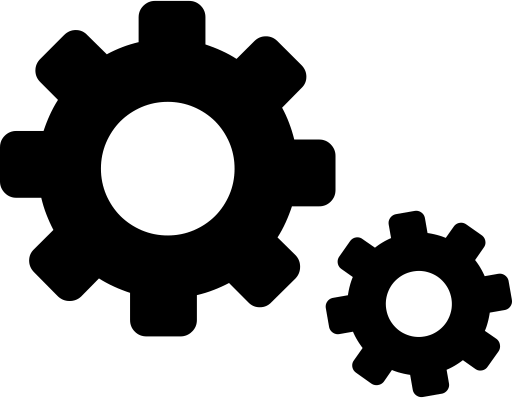
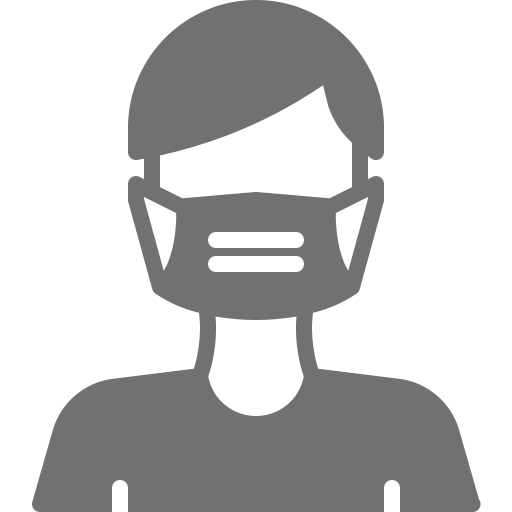
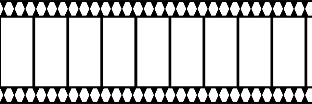


Image ‑ Simulate the way system detect people wear mask using camera

* Connect the system to camera and run on real time
* The system will catch every frame from camera
* Handling frame so that it can fit with input shape of system
* The system will give the frame back as result that after go through detect system and draw red box if no mask and green box if face has worn mask and accuracy of system with notification will appeard on the above the box

Some image output from system

|  |  |
| --- | --- |
| Image ‑ Subject no mask | Image ‑ Subject wear mask. |
| Image ‑ Subject no mask and subject has mask in one image. | Image ‑ Can be detect with multiple subject in same time. |

# Conclusion and improvement

## Conclusion

We has explained detect subject wear a mask. It is problem that have important roll for right now and for computer vision.Thesis has focus on research theory and application for detection problem

Improve accuary of detect subject wear a mask by gather more data set, try out in public, memorize mistake and analysis speed to help system more faster.

But our system in project that has some limitation like can’t detect subject if not enough light, instead a face mask we can block our mouth, nose by something else and system still count as have a face mask. Eventhough there are still some mistake like that, but we are continuing to fix those mistake in no time

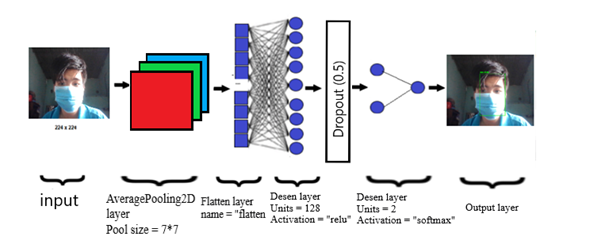


Image ‑ Mô hình minh họa thiết kê model.

## Hướng phát triển

Because we were limited by time, so we only can research in normal size dataset. After this we will try training model with more bigger dataset.

After this diseas, the system will change the purpose. Because we don’t have to wear a mask everytime anymore.

So our team we change purpose to connect this system with another system is indentify detection in real time for indentify of people who still wear a mask. So with this system we can roll up in company, school or army ...

# REFERNCES

1. Mạng nơ ron tích chập cheatsheet: Đại học Stanford

(<https://stanford.edu/~shervine/l/vi/teaching/cs-230/cheatsheet-convolutional-neural-networks>)

1. S. A. Sanjaya and S. Adi Rakhmawan, "Face Mask Detection Using MobileNetV2 in The Era of COVID-19 Pandemic," 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), Sakheer, Bahrain, 2020, pp. 1-5,
2. Valueva, M.V.; Nagornov, N.N.; Lyakhov, P.A.; Valuev, G.V.; Chervyakov, N.I. (2020). "Application of the residue number system to reduce hardware costs of the convolutional neural network implementation". Mathematics and Computers in Simulation. Elsevier BV. 177: 232–243. doi: 10.1016/j.matcom.2020.04.031. ISSN 0378-4754. Convolutional neural networks are a promising tool for solving the problem of pattern recognition.
3. Zhang, Wei (1988). "Shift-invariant pattern recognition neural network and its optical architecture". Proceedings of Annual Conference of the Japan Society of Applied Physics.
4. COVID-19 pandemic – Wikipedia

(<https://en.wikipedia.org/wiki/COVID-19_pandemic>)

1. Lưu ý về đeo khẩu trang ngừa lây nhiễm COVID-19 - Bộ Y tế - Trang tin về dịch bệnh viêm đường hô hấp cấp COVID-19

(<https://ncov.moh.gov.vn/en/web/guest/-/luu-y-ve-eo-khau-trang-ngua-lay-nhiem-covid-19>)

1. H. A. Rowley, S. Baluja and T. Kanade, "Neural network-based face detection," in IEEE Transactions on Pattern Analysis and Machine Intelligence.
2. Ševo and A. Avramović, "Convolutional Neural Network Based Automatic Object Detection on Aerial Images," in IEEE Geoscience and Remote Sensing Letters.

**Giáo viên hướng dẫn:** ThS. Lương Trần Ngọc Khiết.  
(Số điện thoại: 0387469465 - Email: [khietltn@hcmue.edu.vn](mailto:khietltn@hcmue.edu.vn) )

Nhóm sinh viên thực hiện:

* Lê Tấn Lộc

(Số điện thoại: 0785607050 - Email: [ah09program@gmail.com](mailto:ah09program@gmail.com))

* Huỳnh Thanh Phong

(Số điện thoại: 0931493972 - Email: [huynhthanhphong0931@gmail.com](mailto:huynhthanhphong0931@gmail.com))

* Phạm Duy Minh

(Số điện thoại: 0333454717 - Email: [duyminh081201@gmail.com](mailto:duyminh081201@gmail.com))

* Trần Thị Tứ Linh

(Số điện thoại: 0353544089 - Email: [tranthitulinh1305@gmail.com](mailto:tranthitulinh1305@gmail.com))