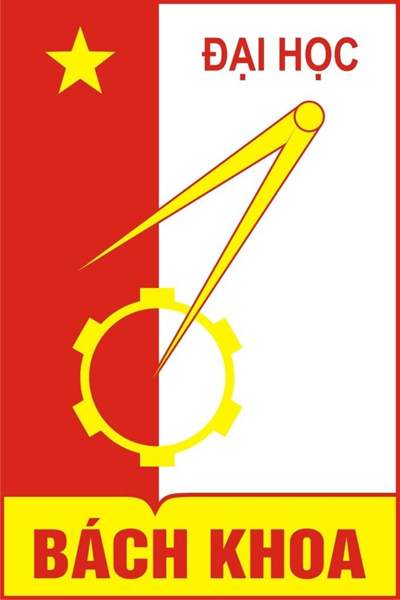
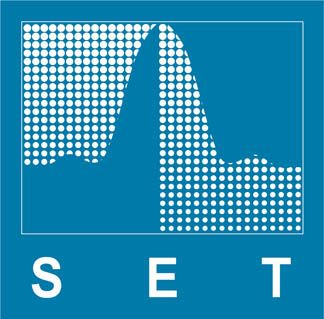
TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI

Viện Điện tử - Viễn thông

Bộ môn CNĐT & Kỹ thuật Y sinh

--------------------------------

**XỬ LÝ ẢNH Y TẾ**

ĐỀ TÀI: PHÂN LOẠI ẢNH UNG THƯ DẠ DÀY

SỬ DỤNG MÔ HÌNH TÚI TỪ TRỰC QUAN

*Giảng viên hướng dẫn :* **TS. Nguyễn Việt Dũng**

*Nhóm 10:  
Sinh viên thực hiện MSSV*

**Trịnh Anh Dương 20130759**

**Chử Văn Thành 20133502**

**Trần Hồng Giang 20131120**

**Nguyễn Đức Thiện 20133742**

***Hà Nội, 06/2017***

**MỤC LỤC**

# LỜI MỞ ĐẦU

Ngày này ung thư trong đường tiêu hóa là một trong những căn bệnh nguy hiểm trên toàn thế giới, tuy nhiên loại ung thư này có ít triệu chứng ở giai đoạn đầu. Chính vì vậy, việc cung cấp các công cụ tốt hơn để chẩn đoán có thể giúp việc phát hiện sớm và điều trị là vô cùng cần thiết. Các công nghệ xử lý hình ảnh y tế để chẩn đoán ung thư đã và đang được phát triển trong thế kỷ qua, đặc biệt là nội soi. Chúng cho phép thu được hình ảnh của các mô trên đường tiêu hóa với độ phân giải tốt. Các kỹ thuật để phân tích những hình ảnh này đã được phát triển và đã được sử dụng bởi hệ thống Chẩn đoán bằng máy tính (Computer Aided Diagnosis -CAD).

Trong hệ tiêu hóa, các kỹ thuật của hệ thống CAD đã cho phép các bác sĩ phân tích các hình ảnh nội soi. Tuy nhiên, sự nhận biết ung thư trong đường tiêu hóa là một vấn đề hết sức phức tạp, trong đó chỉ có các bác sĩ được đào tạo có tỷ lệ thành công cao. Một số giải pháp nhận dạng mẫu bao gồm vấn đề này đã được phát hiện và ứng dụng trong quá khứ. Trong đề tài này, chúng em đã tạo một Bag of visual Words (BoW), từ BoW và dữ liệu ảnh tạo nên một bộ phân loại SVM tuyến tính 2 lớp.

Chúng em đã thực hiện qua bộ dữ liệu hình ảnh mà chúng em đã điều chỉnh tất cả các thông số có liên quan. Phương pháp của chúng em trình bày hiệu suất tối ưu của nó khi điều chỉnh: hình ảnh chuyển đổi sang màu xám, giảm kích thước 50%. Phương pháp thực hiện tốt hơn khi sử dụng “Máy dò tính tăng SURF” (*Speeded up robust features*), mang lại hiệu quả tốt hơn khi dùng phương pháp chia lưới lên ảnh. Chúng em đã đánh giá phương pháp và đưa ra được sai số trung bình khoảng 6 ± 1%.

Cuối cùng, chúng em xin chân thành cảm ơn thầy giáo **TS. Nguyễn Việt Dũng** đã có những bài giảng bổ ích trong quá trình học tập và hướng dẫn để chúng em có thể hoàn thành đề tài, chúng em xin chúc thầy hạnh phúc và thành đạt trong cuộc sống.

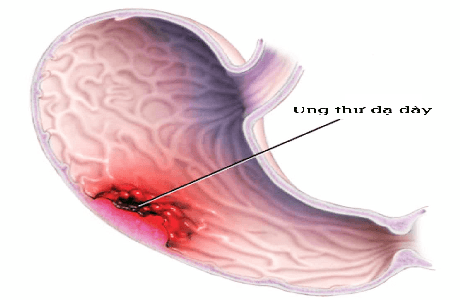
# CƠ SỞ LÝ THUYẾT

**Chương 1: Tổng quan đề tài**

* 1. Ung thư dạ dày

Ung thư dạ dày là một trong những bệnh ung thư nghiêm trọng nhất, do tỷ lệ tử vong cao và sự phổ biến của nó. Ung thư có thể phát triển ở bất cứ phần nào của dạ dày, có thể lan ra khắp dạ dày và đến các cơ quan khác của cơ thể, đặc biệt là thực quản, phổi, hạch bạch huyết và gan. Ung thư dạ dày mỗi năm có thể gây ra 800.000 ca tử vong trên khắp thế giới.

Việc phát hiện ra ung thư dạ dày tương tự như với những loại ung thư khác, và nó không còn là một hoạt động y tế độc lập. Thật vậy, sự kết hợp giữa bí quyết chẩn đoán trong y tế với những tiến bộ có liên quan trong kỹ thuật ung thư, mới lạ và sáng tạo cho một đặc tính khách quan của mô. Kỹ thuật đó được tạo ra và phát triển bởi các kỹ sư y sinh. Mối quan hệ giữa y học và kỹ thuật đã trở nên gần gũi hơn trong vài năm qua, cho phép bác sĩ đi từ triệu chứng đơn giản (phân tích các dấu hiệu và triệu chứng bằng mắt thường), đến sử dụng các công cụ cung cấp sự đánh giá khách quan và tin cậy cho bệnh nhân, (ví dụ, nội soi, hoặc gần đây hơn là hệ thống chẩn đoán được hỗ trợ bởi máy tính CAD). Những giải pháp này mặc dù sáng tạo nhưng vẫn chưa đủ linh hoạt để giải quyết được các biến thể không xác định hoặc vô hình của hình thái khối u. Điểm bất lợi này thúc đẩy sự phát triển của một hệ thống phân tích hình ảnh ung thư dạ dày gắn với chiến lược nhận dạng ung thư, cho ra đánh giá đáng tin cậy trong trường hợp lâm sàng mới.



**Hình 1.1 Ung thư dạ dày**

* 1. Kỹ thuật nội soi
     1. Khái niệm

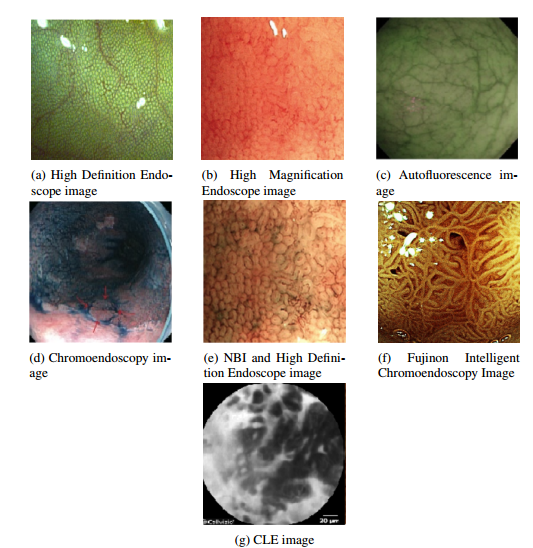
Nội soi (Endoscopy) là một kỹ thuật [y học](https://vi.wikipedia.org/wiki/Y_h%E1%BB%8Dc) hiện đại được ứng dụng trong việc [khám bệnh](https://vi.wikipedia.org/w/index.php?title=Kh%C3%A1m_b%E1%BB%87nh&action=edit&redlink=1), chẩn đoán bệnh bằng việc sử dụng các loại dụng cụ chuyên biệt để quan sát trực tiếp bên trong các cơ quan của [cơ thể](https://vi.wikipedia.org/w/index.php?title=C%C6%A1_th%E1%BB%83&action=edit&redlink=1). Với kỹ thuật nội soi, người ta có thể [quay phim](https://vi.wikipedia.org/w/index.php?title=Quay_phim&action=edit&redlink=1), chụp hình bên trong các cơ quan, lấy dị vật, [sinh thiết](https://vi.wikipedia.org/w/index.php?title=Sinh_thi%E1%BA%BFt&action=edit&redlink=1) và thậm chí là thực hiện hiện phẫu thuật nội soi. Nội soi hiện nay được sử dụng trong hầu hết các chuyên khoa: [tai mũi họng](https://vi.wikipedia.org/w/index.php?title=Tai_m%C5%A9i_h%E1%BB%8Dng&action=edit&redlink=1), [tiêu hóa](https://vi.wikipedia.org/wiki/Ti%C3%AAu_h%C3%B3a) ([thực quản](https://vi.wikipedia.org/wiki/Th%E1%BB%B1c_qu%E1%BA%A3n), [dạ dày](https://vi.wikipedia.org/wiki/D%E1%BA%A1_d%C3%A0y), ruột già…) sản, [ngoại, tiết niệu](https://vi.wikipedia.org/w/index.php?title=Ngo%E1%BA%A1i,_ti%E1%BA%BFt_ni%E1%BB%87u&action=edit&redlink=1), xương khớp, [thần kinh](https://vi.wikipedia.org/w/index.php?title=Th%E1%BA%A7n_kinh&action=edit&redlink=1), thẩm mỹ…

Thiết bị nội soi là một thiết bị y tế có khả năng thu được hai hình ảnh đồng thời và thậm chí các mẫu mô từ đường tiêu hóa. Mặc dù có rất nhiều cấu hình nhưng các loại nội soi đều có cùng một số đặc điểm:

* Bao gồm một ống mềm dẻo
* Có khả năng chiếu sáng và thu thập hình ảnh được kết hợp
* Có một kênh phụ, có thể chụp lấy mẫu sinh thiết.
  + 1. Các kỹ thuật nội soi

**Bảng 1.1 Các kỹ thuật nội soi**

|  |  |  |
| --- | --- | --- |
| CÔNG NGHỆ NỘI SOI | ỨNG DỤNG | MỤC ĐÍCH |
| Tiêu chuẩn và HD | Phát hiện những thay đổi tinh vi trong niêm mạc dễ dàng hơn | Tăng bề mặt nội soi và giảm các thành phần lạ trong ảnh |
| HME ( Phóng đại cao) | Xác định hiện tượng tân mạch | Quan sát tốt chi tiết bề mặt nội soi và mạch máu. |
| AFI ( Hình ảnh phát huỳnh quang tự động) | Xác định hoặc phát hiện sớm ung thư | Mô tả mô và phát hiện ung thư |
| NBI ( Chụp hình với biên độ hẹp) | Xác định hiện tượng tân mạch kết đặc biệt kết hợp với nội soi phóng đại. | Tăng cường mô hình vi mô |
| I-scan và FICE | Xác định hiện tượng tân mạch kết đặc biệt kết hợp với nội soi phóng đại | Phát hiện, chẩn đoán và điều trị bệnh dạ dày ruột dễ dàng hơn |
| CLE (Chụp quang tuyến Laser) | Xác định hiện tượng tân mạch | Cung cấp hình ảnh cực nhỏ có độ phân giải cao |



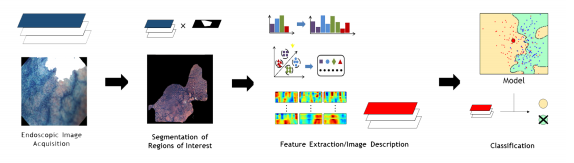
**Hình 1.2 Hình ảnh minh họa 7 công nghệ nội soi**

1. Hình ảnh nội soi có độ phân giải cao của bề mặt niêm mạc bình thường
2. Phóng đại cao của niêm mạc ruột bình thường
3. AFI trong ruột kết bình thường
4. Hình ảnh Chromoendoscope
5. Hình ảnh NBI và HD
6. Hình ảnh nội soi màu nâu nhạt của một tổn thương niêm mạc
7. Hình ảnh mô bình thường sử dụng đầu dò CLE
   1. Thị giác máy tính trong nội soi dạ dày

Việc thu thập lại hình ảnh của đường tiêu hóa bằng cách sử dụng đầu dò nội soi là một tiến trình dễ bị nhiễu. Bởi nó được thực hiện trong điều kiện không được kiểm soát, một số vấn đề liên quan đến chiếu sáng, xoay đầu dò, đổ bóng hoặc bị kẹt. Thị giác máy tính (CV) trong nội soi dạ dày phải đối phó với những bất ổn này cũng như với các mẫu hình ảnh yếu, cản trở khả năng nhận biết tổn thương.

Các công trình được phát triển trước đây trong lĩnh vực nghiên cứu này bao gồm một loạt các kỹ thuật mở rộng của CV. Trong lĩnh vực này nó sẽ sử dụng một thuật toán có khả năng trích xuất mạnh mẽ các đặc trưng liên quan. Mặc dù cần phải giải quyết một vấn đề y tế, nhưng nó cũng là một mô hình nhận dạng đầy thách thức. Một giải pháp khả thi có thể bao gồm các thuật toán xử lý ảnh mà chưa được thử trước đây trong lĩnh vực này.

Hình cho thấy tiến trình chung cho việc thu và xử lý hình ảnh nội soi trong một hệ thống CAD. Đặc biệt nhấn mạnh đến tính năng trích xuất và mô tả hình ảnh.



**Hình 1.3 Quy trình thu và xử lý ảnh nội soi trong hệ thống CAD**

**Chương 2: Tổng quan về đề tài nghiên cứu tương tự**

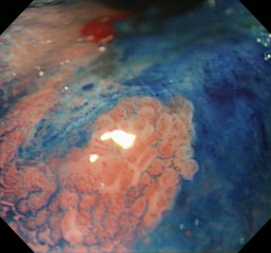
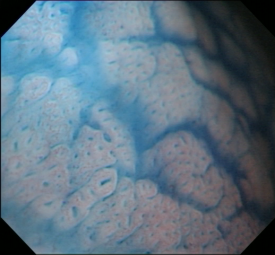
**Đề tài**: Phân loại ảnh ung thư dạ dày dựa vào mô hình túi từ trực quan.

* 1. Mô hình túi từ trực quan

Mô hình túi từ trực quan (Bag of words model) là một trong những phương pháp phổ biến nhất dành cho phân loại object (đặc biệt trong xử lý ảnh). Mô hình này học một bộ từ vựng từ của tất cả các văn bản, rồi mô hình hóa các văn bản bằng cách đếm số lần xuất hiện của mỗi từ trong văn bản đó. Từ đó mã hóa mỗi văn bản thành các vector.

* 1. Thực hiện
     1. Dataset:

Sử dụng tập dữ liệu hình ảnh nội soi “*Chromoendoscopy*” gồm 176 ảnh, mỗi ảnh có kích thước 518x481, trong đó có 56 ảnh mô bình thường, và 120 ảnh mô bất thường (96 ảnh mắc chứng metaplasia và 24 ảnh thuộc triệu chứng dysplasia) cùng với vùng ảnh quan tâm (Region Of Interest) tương ứng với mỗi ảnh, được xác định bởi bác sĩ.



a) Mô bình thường b) Mô bị nhiễm dysplasia c) Mô bị nhiễm metaplasia

**Hình 2.1 Hình ảnh mô tả các loại mô khác nhau từ dạ dày đại diện cho bộ dữ liệu**

* + 1. Các bước thực hiện

Bước 1: Tiền xử lý

* Loại bỏ nhiễu và tách lấy vùng cần quan tâm ROI (ảnh xám) để thu được tập data mới chứa 120 ảnh.
* Tạo tập data mới ngẫu nhiên có 112 ảnh chia đều cho 2 lớp bình thường và bất thường.
* Chia tập data mới thành 2 tập ảnh *trainingSets* và *testingSets* với tỷ lệ 7:3.

Bước 2: Tạo túi từ trực quan (Túi từ tính năng – BoW)

* Trích xuất tính năng của từng ảnh trong *trainingSets.*
* Tối ưu hóa các tính năng.
* Tạo ra túi từ trực quan bằng phương pháp phân cụm K-means.

Bước 3: Đào tạo “Bộ phân loại”

sử dụng hàm “*trainImageCategoryClassifier*” trong Matlab.

Bước 4: Kiểm định “Bộ phân loại” và tính tỷ lệ lỗi

* 1. Các tham số
* Kích thước hình ảnh
* tỷ lệ *N* = *trainingSets : testingSets*
* Kích thước Vocabulary (số cụm K trong thuật toán K-means)
* Hệ số *StrongestFeatures*
* Tham số *PointSelection*
* Tham số *GridStep*
* Độ lệch chuẩn của bộ lọc Gaussian
  1. Kết quả

**Bảng 2.1 Kết quả kiểm định tập *trainingSets* và *testingSets* ở chế độ tối ưu**

|  |  |
| --- | --- |
| *trainingSets* | *testingSets* |
| PREDICTED  KNOWN | Abnormal\_50 Normal\_50  -------------------------------------------  Abnormal\_50 | 0.92 0.08  Normal\_50 | 0.06 0.94  \* Độ chính xác trung bình: 93% | PREDICTED  KNOWN | Abnormal\_50 Normal\_50  -------------------------------------------  Abnormal\_50 | 0.94 0.06  Normal\_50 | 0.03 0.97  \* Độ chính xác trung bình: 95% |

* 1. Đánh giá

**Chương 3: Cơ sở lý thuyết**

* 1. Giới thiệu về Machine Learning

Machine Learning là một lĩnh vực của [trí tuệ nhân tạo](https://vi.wikipedia.org/wiki/Tr%C3%AD_tu%E1%BB%87_nh%C3%A2n_t%E1%BA%A1o) liên quan đến việc nghiên cứu và xây dựng các kĩ thuật cho phép các hệ thống "học" tự động từ dữ liệu để giải quyết những vấn đề cụ thể. Ví dụ như các máy có thể "học" cách phân loại [thư điện tử](https://vi.wikipedia.org/wiki/Th%C6%B0_%C4%91i%E1%BB%87n_t%E1%BB%AD) xem có phải [thư rác (spam)](https://vi.wikipedia.org/wiki/Th%C6%B0_r%C3%A1c) hay không và tự động xếp thư vào thư mục tương ứng. Học máy rất gần với [suy diễn thống kê](https://vi.wikipedia.org/w/index.php?title=Suy_di%E1%BB%85n_th%E1%BB%91ng_k%C3%AA&action=edit&redlink=1) (statistical inference) tuy có khác nhau về thuật ngữ.

Học máy có liên quan lớn đến [thống kê](https://vi.wikipedia.org/wiki/Khoa_h%E1%BB%8Dc_Th%E1%BB%91ng_k%C3%AA), vì cả hai lĩnh vực đều nghiên cứu việc phân tích dữ liệu, nhưng khác với thống kê, học máy tập trung vào sự phức tạp của các giải thuật trong việc thực thi tính toán. Nhiều bài toán suy luận được xếp vào loại bài toán [NP-khó](https://vi.wikipedia.org/wiki/NP-kh%C3%B3) (nondeterministic polynomial time), vì thế một phần của học máy là nghiên cứu sự phát triển các giải thuật suy luận xấp xỉ mà có thể xử lý được.

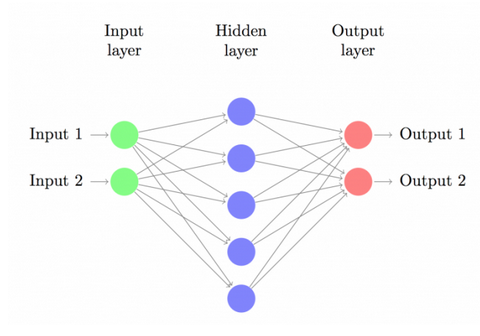
Học máy hiện nay được áp dụng rộng rãi bao gồm [máy truy tìm dữ liệu](https://vi.wikipedia.org/wiki/M%C3%A1y_truy_t%C3%ACm_d%E1%BB%AF_li%E1%BB%87u), [chẩn đoán y khoa](https://vi.wikipedia.org/wiki/Ch%E1%BA%A9n_%C4%91o%C3%A1n), phát hiện [thẻ tín dụng giả](https://vi.wikipedia.org/w/index.php?title=Th%E1%BA%BB_t%C3%ADn_d%E1%BB%A5ng_gi%E1%BA%A3&action=edit&redlink=1), phân tích [thị trường chứng khoán](https://vi.wikipedia.org/wiki/Th%E1%BB%8B_tr%C6%B0%E1%BB%9Dng_ch%E1%BB%A9ng_kho%C3%A1n), phân loại các [chuỗi DNA](https://vi.wikipedia.org/w/index.php?title=Chu%E1%BB%97i_DNA&action=edit&redlink=1), [nhận dạng tiếng nói](https://vi.wikipedia.org/wiki/Nh%E1%BA%ADn_d%E1%BA%A1ng_ti%E1%BA%BFng_n%C3%B3i) và [chữ viết](https://vi.wikipedia.org/w/index.php?title=Nh%E1%BA%ADn_d%E1%BA%A1ng_ch%E1%BB%AF_vi%E1%BA%BFt_tay&action=edit&redlink=1), [dịch tự động](https://vi.wikipedia.org/wiki/D%E1%BB%8Bch_t%E1%BB%B1_%C4%91%E1%BB%99ng), [chơi trò chơi](https://vi.wikipedia.org/w/index.php?title=Tr%C3%B2_ch%C6%A1i_chi%E1%BA%BFn_l%C6%B0%E1%BB%A3c&action=edit&redlink=1) và [cử động rô-bốt](https://vi.wikipedia.org/w/index.php?title=C%E1%BB%AD_%C4%91%E1%BB%99ng_r%C3%B4-b%E1%BB%91t&action=edit&redlink=1) (robot locomotion).

* 1. Convolutional neural network

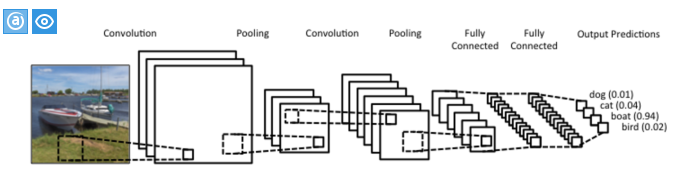
Mạng noron tích chập- CNN (convolutional neural network) là một trong những mô hình Deep Learning tiên tiến giúp cho chúng ta xây dựng được những hệ thống thông minh với độ chính xác cao như hiện nay như hệ thống xử lý ảnh lớn như Facebook, Google hay Amazon đã đưa vào sản phẩm của mình những chức năng thông minh như nhận diện khuôn mặt người dùng, phát triển xe hơi tự lái hay drone giao hàng tự động.

Mạng- là thuật ngữ để chỉ sự dùng một cấu trúc có các Node được liên kết với nhau bằng các cạnh, có một hoặc nhiều lớp kết nối với nhau, có đầu vào và đầu ra. Mạng CNN là một tập hợp các lớp Convolution chồng lên nhau và sử dụng các hàm nonlinear activation như ReLU và tanh để kích hoạt các trong số trong các Node. Mỗi một lớp sau khi đi qua hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo.

Trong mô hình mạng truyền ngược (feedforward neural network) thì mỗi noron đầu vào (input node) cho mỗi noron đầu ra trong các lớp tiếp theo. Mô hình này gọi là mạng kết nối đầy đủ (fully connected layer) hay mạng toàn vẹn (affine layer). Còn trong mô hình CNNs thì ngược lại. Các layer liên kết được với nhau thông qua cơ chế convolution. Layer tiếp theo là kết quả convolution từ layer trước đó, nhờ vậy mà ta có được các kết nối cục bộ. Như vậy mỗi noron ở lớp kế tiếp sinh ra từ kết quả của filter áp đặt lên một vùng ảnh cục bộ của noron trước đó.



**Hình 3.1 mô hình mạng truyền ngược**



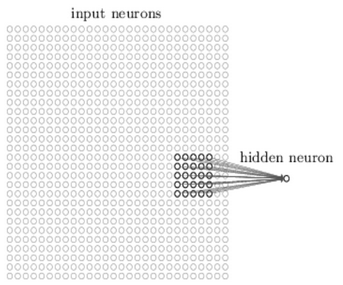
**Hình 3.2 Mô hình mạng CNN**

Mỗi một lớp được sử dụng các filter khác nhau. Thông thường có hàng trăm nghìn filter như vậy và kết hợp kết quả của chúng lại. Ngoài ra có một số layer khác như pooling/subsampling layer dùng để chắt lọc lại các thông tin hữu ích hơn (loại bỏ các thông tin nhiễu). Trong quá trình huấn luyện mạng (training) CNN tự động học các giá trị qua các lớp filter dựa vào cách thức mà bạn thực hiện. Ví dụ trong tác vụ phân lớp ảnh, CNNs sẽ cố gắng tìm ra thông số tối ưu cho các filter tương ứng theo thứ tự raw pixel > edges > shapes > facial > high-level features. Layer cuối cùng được dùng để phân lớp ảnh.

Trong mô hình CNN có 2 khía cạnh cần quan tâm là Location Invariance (tính bất biến) and Compositionality (tính kết hợp). Với cùng một đối tượng, nếu được chiếu theo các góc độ khác nhau (translation, rotation, scaling) thì độ chính xác của thuật toán sẽ bị ảnh hưởng đáng kể. Pooling layer sẽ cho bạn tính bất biến đối với phép dịch chuyển (translation), phép quay (rotation) và phép co giãn (scaling). Tính kết hợp cục bộ cho ta các cấp độ biểu diễn thông tin từ mức độ thấp đến mức độ cao và trừu tượng hơn thông qua convolution từ các filter. Đó là lý do tại sao CNNs cho ra mô hình với độ chính xác rất cao. Cũng giống như cách con người nhận biết các vật thể trong tự nhiên.

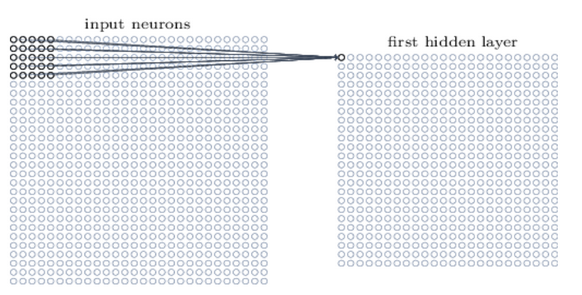
Mạng tích chập sử dụng 3 ý tưởng cơ bản: các trường tiếp nhận cục bộ (local receptive field), trọng số chia sẻ (shared weights) và tổng hợp (pooling).

**Trường tiếp nhận cục bộ (local receptive field):** đầu vào của mạng CNN là một ảnh. Ví dụ như ảnh có kích thước 28x28 thì tương ứng đầu vào là một ma trận có 28x28 và giá trị mỗi điểm ảnh là một ô trong ma trận. Trong mô hình mạng ANN truyền thống thì chúng ta sẽ kết nối các noron đầu vào vào tầng ảnh. Tuy nhiên trong CNN chúng ta không làm như vậy mà chúng ta chỉ kết nối trong một vùng nhỏ của các noron đầu vào như một filter có kích thước 5x5 tương ứng 25 điểm ảnh đầu vào. Mỗi một kết nối sẽ học một trọng số và mỗi noron ẩn sẽ học một bias.

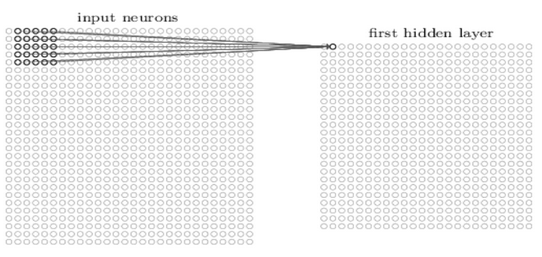


Tiếp tục filter sẽ trượt trên tấm ảnh tương ứng với mỗi vùng trượt sẽ có một noron ẩn trong lớp ẩn đầu tiên.

* Bước 1: Tạo ra noron ẩn đầu tiên trong lớp ẩn 1

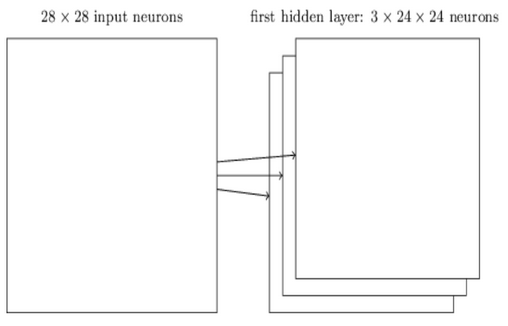


* Bước 2: Dịch filter qua bên phải một cột sẽ tạo được noron ẩn thứ 2.



Như vậy kết thúc nếu ta có ảnh đầu vào là 28x28 và filter có kích thước là 5x5 thì ta sẽ có bao nhiêu nút trong lớp ẩn. Kết quả là 28-5+1 = 24 noron trong lớp ẩn đầu tiên do filter trượt đến dòng cuối chỉ có thể thực hiện 23 lần trước khi chạm đường biên. Đối với bài toán nhận dạng ảnh người ta thường gọi ma trận lớp đầu vào là feature map, trọng số xác định các đặc trương là shared weight và độ lệch xác định một feature map là shared bias.

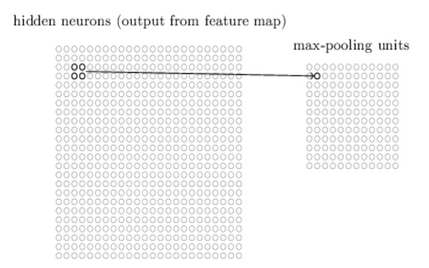
Như vậy đơn giản nhất là qua các bước trên chúng ta chỉ có 1 feature map. Tuy nhiên trong nhận dạng ảnh chúng ta cần nhiều hơn một feature map.



Như trong hình trên, một mạng có một ảnh đầu vào sẽ xây dựng 3 feature map. Như vậy để làm gì? Một shared weght và shared bias sẽ tạo một feature map. Như vậy ta tìm được 3 loại đặc trưng khác nhau với mỗi đặc trưng được phát hiện trên toàn bộ ảnh. Trong thực tế các bài toán sử dụng tìm được số lượng feature khác nhau như mạng LeNet sử dụng 6 bản đồ đặc trưng khác nhau.

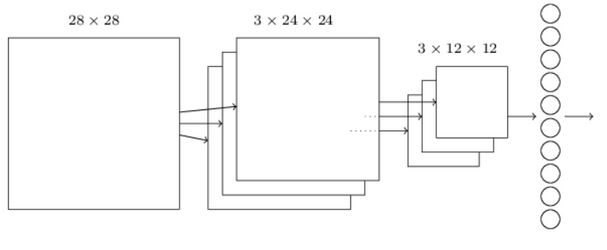
**Trọng số chia sẻ (shared weights):** Mối quan hệ giữa số lượng Feature map với số lượng tham số là gì? Chúng ta thấy mỗi fearture map cần 25 = 5x5 shared weight và 1 shared bias. Như vậy mỗi feature map cần 5x5+1 = 26 tham số. Như vậy nếu có 10 feature map thì có 10x26 = 260 tham số. Chúng ta xét lại nếu lớp đầu tiên có kết nối đầu đủ nghĩa là chúng ta có 28x28=784 neuron đầu vào như vậy ta chỉ có 30 noron ẩn. Như vậy ta cần 28x28x30 shared weight và 30 shared bias. Tổng số tham số là 28x28x30+30 tham số lớn hơn nhiều so với CNN. Ví dụ vừa rồi chỉ mô tả để thấy được sự ước lượng số lượng tham số chứ chúng ta không so sánh được trực tiếp vì 2 mô hình khác nhau. Nhưng nếu mô hình có số lượng tham số ít hơn thì nó sẽ chạy nhanh hơn.

**Tổng hợp (pooling):** lớp pooling thường được sử dụng ngay sau lớp convolutional để đơn giản hóa thông tin đầu ra để giảm bớt số lượng noron. Thủ tục pooling phổ biến là Max pooling, thủ tục này chọn giá trị lớn nhất trong vùng đầu vào 2x2.



Như vậy qua lớp Max Pooling thì số lượng noron giảm đi phân nửa. Trong một mạng CNN có nhiều Feature Map nên mỗi Feature Map chúng ta sẽ cho một Max Pooling khác nhau. Có thể thấy rằng Max Pooling là cách hỏi xem trong các đặc trưng này thì đặc trưng nào là đặc trưng nhất. Ngoài Max Pooling còn có L2 Pooling.

Cuối cùng ta đặt tất cả các lớp lại với nhau thành một CNN với đầu ra gồm các noron với số lượng tùy bài toán. Như bài toán nhận dạng chữ số viết tay thì đầu ra là 10.



Lớp cuối cùng của các kết nối trong mạng là một lớp đầy đủ kết nối. Đó là, lớp này nối mọi noron từ lớp max pooled tới mọi noron của tầng ra.

* 1. Multiclass SVM

Support vector machine (SVM) ban đầu được thiết kế cho phân loại nhị phân. Để mở rộng cho bài toán đa lớp, một số mô hình phân loại đã được đề xuất. Tuy nhiên, số lượng các biến là sản phẩm của số lượng mẫu theo số lớp, tạo ra độ phức tạp tính toán lớn.

SVM ban đầu tách các lớp nhị phân (k=2) với tiêu chí số dư được tối đa hóa. Tuy nhiên, vấn đề thực tế thường yêu cầu phân biệt cho hơn hai loại. Do đó, nhận dạng mẫu đa lớp có nhiều ứng dụng mở rộng bao gồm nhận dạng ký tự quang học (OCR), phát hiện xâm nhập, nhận dạng giọng nói và thông tin sinh học. Trong thực tế, vấn đề phân loại đa lớp (k>2) thường được phân tách thành một loạt các vấn đề nhị phân sao cho tiêu chuẩn SVM có thể được áp dụng trực tiếp. Hai sơ đồ đại diện là hướng tiếp cận một-nhiều (one-versus-rest, 1VR) và một-một (one-versus-one, 1V1). Cả hai hướng này đều là trường hợp đặc biệt của mã đầu ra sửa lỗi (Error Correcting Output Codes, ECOC) khi chia vấn đề đa lớp thành một tập các vấn đề nhị phân đã được định trước. Vấn đề chính của hướng tiếp cận này là xây dựng một ma trận ECOC tốt. Một đề xuất khác là giải quyết trực tiếp vấn đề đa lớp trong một quá trình tối ưu hóa duy nhất. Loại mô hình này kết hợp vấn đề tối ưu hóa nhiều lớp nhị phân vào một hàm khách quan duy nhất và đồng thời phân loại nhiều lớp. Tuy nhiên, độ phức tạp tính toán lớn hơn là cần thiết cho kích thước của vấn đề phương trình bậc hai (Quadratic Programming, QP).

Hơn nữa, Szedmak đã đề xuất một mô hình đa lớp cho SVM L1-norm. Trong công thức của họ, khung one-versus-rest được sử dụng. Một vấn đề tiềm năng là khi số lượng các lớp lớn, mỗi bộ phân loại nhị phân trở nên không cân bằng. Vấn đề phân loại không cân bằng xảy ra khi có nhiều mẫu của một số lớp hơn những lớp khác. Trong trường hợp này, bộ phân loại chuẩn có xu hướng bị áp đảo bởi các lớp có quy mô lớn và bỏ qua những lớp nhỏ. Thuật toán SVM đang xây dựng một phân tách siêu phẳng với lề tối đa. Từ chỉ có các vecto hỗ trợ được dùng cho bộ phân loại và đa số mẫu xa ranh giới quyết định có thể loại bỏ, SVM có thể chính xác hơn với dữ liệu không cân bằng vừa phải. Tuy nhiên, SVM nhạy cảm cao độ với bộ phân loại không cân bằng vì nó dễ tạo ra một bộ phân loại mà có sự đánh giá thiên vị mạnh hướng tới lớp lớn và sẽ cho độ chính xác thấp cho lớp nhỏ. Có một đề xuất phương pháp loại bỏ những khó khăn của one-versus-rest bằng cách xử lý nhiều lớp cùng nhau, đó là mở rộng phương pháp LS-SVM thành trường hợp đa lớp. Một nhược điểm của LS-SVM là giải pháp của nó được xây dựng từ hầu hết các ví dụ đào tạo. Ý tưởng PSVM (SVM gần đúng) được áp dụng vào vấn đề đa lớp có liên kết chặt chẽ với phương pháp one-versus-rest. Đối với mỗi vấn đề nhỏ được phân tách ra, giải pháp là tương tự như trường hợp nhị phân của nó: phân loại các mẫu mới bằng cách gán chúng vào gần hai mặt phẳng song song được đẩy cách nhau càng xa càng tốt.

# Chương 4: Thực hiện

# KẾT QUẢ VÀ ĐÁNH GIÁ

# TÀI LIỆU THAM KHẢO

[1] S. Francisco, M. T. Coimbra, and R. Gamelas Sousa, “Recognition of Cancer using a Bag-of-Words Random Forest for Gastroenterology,” 2015

[2] D. Boschetto and E. Grisan, “Superpixel-based classification of gastric chromoendoscopy images” , Conference: SPIE Conference on Medical Imaging, February 2017

[3] Farhan Riaz, Fernando Vilarino, Mario Dinis Ribeiro, and Miguel Coimbra, “Identifying Potentially Cancerous Tissues in Chromoendoscopy Images”, Conference: Pattern Recognition and Image Analysis - 5th Iberian Conference, IbPRIA 2011, Las Palmas de Gran Canaria, Spain, June 8-10, 2011. Proceedings

[4] Ribeiro, M.D.: Clinical, Endoscopic and Laboratorial Assessment of Patients with Associated Lesions to Gastric Adenocarcinoma. Faculdade de Medicina, Universidade do Porto, PhD Thesis (2005)

[5] Jing Zhang, Shi-Bin Guo, Zhi-Jun Duan, “Application of magnifying narrow-band imaging endoscopy for diagnosis of early gastric cancer and precancerous lesion”, in BMC Gastroenterology 11(1):135 · December 2011

[6] <https://aidasub-chromogastro.grand-challenge.org>

[7] https://vi.wikipedia.org/wiki/Học\_máy

[8] <https://techtalk.vn/tat-tan-tat-moi-kien-thuc-co-ban-ve-machine-learning.html>

[9] <https://ongxuanhong.wordpress.com/2015/08/25/danh-gia-mo-hinh-model-evaluation/>

[10] <https://www.mathworks.com/help/vision/ug/image-classification-with-bag-of-visual-words.html?s_tid=srchtitle>

[11] <https://www.mathworks.com/help/vision/examples/image-category-classification-using-bag-of-features.html>

[12] [http://machinelearningcoban.com/general/2017/02/06/featureengineering/#](http://machinelearningcoban.com/general/2017/02/06/featureengineering/)

feature-scaling-and-normalization

[13]<https://vi.wikipedia.org/wiki/M%C3%A1y_vect%C6%A1_h%E1%BB%97_tr%E1%BB%A3>

https://vi.wikipedia.org/wiki/Ung\_th%C6%B0\_d%E1%BA%A1\_d%C3%A0y