

# ỨNG DỤNG SÓNG SIÊU ÂM VÀ MẠNG HỌC SÂU ĐỂ NHẬN BIẾT SỰ TÍCH LŨY KIM LOẠI NẶNG TRONG KHOAI LANG

Nguyễn Tiến Hiên\*, Lê Văn Dũng, Nguyễn Trọng Kương

*Khoa Công nghệ thông tin, Học viện Nông nghiệp Việt Nam*

\*Tác giả liên hệ: nguyentienhien@vnua.edu.vn

Ngày nhận bài: 23.03.2020

Ngày chấp nhận đăng: 02.11.2020

## TÓM TẮT

Thực phẩm nhiễm kim loại nặng gây ra nhiều hậu quả nghiêm trọng cho sức khỏe con người, luôn được các cơ sở giám sát chất lượng an toàn thực phẩm (ATTP) kiểm tra bằng các qui trình và phương pháp cẩn thận, thậm chí tốn kém từ việc lấy mẫu đến ước lượng thành phần chất gây hại được tích trong đó. Trong nghiên cứu này, chúng tôi hướng đến sử dụng sóng siêu âm để đánh giá mức độ liên quan với sự tích lũy kim loại nặng trong khoai lang. Sóng siêu âm an toàn với thực phẩm và vì thế không ảnh hưởng đến chất lượng thực phẩm của mẫu kiểm tra. Cùng với dữ liệu thu được, chúng tôi sử dụng mạng học sâu như một bộ phân lớp hiệu quả và đang được quan tâm của nhiều nhà nghiên cứu hiện nay cho việc nhận biết sự khác nhau giữa các mẫu khoai trước và sau khi cho qua dung dịch chì sunfat. Với 31 bộ dữ liệu siêu âm về các mẫu khoai thu được và sử dụng hai mạng nơron (NN) và mạng Deep Boltzmann Machine (DBM) để nhận dạng hai nhóm mẫu khoai lang trước và sau khi cho qua nhiễm chì sunfat. Kết quả cho thấy việc nhận dạng giữa 2 nhóm mẫu chưa cho nhiễm và cho nhiễm chì sunfat của mạng NN là 62% với dữ liệu huấn luyện và 55% với dữ liệu kiểm tra, và của mạng DBM là 68% với dữ liệu huấn luyện và 65% với kiểm tra.

Từ khóa: An toàn thực phẩm, kim loại nặng, trí tuệ nhân tạo, học máy, mạng nơron, máy Boltzmann, học sâu.

## Application of Ultrasound and Deep Networks in Recognizing the Presence of Heavy Metals Contaminated in Sweet Potatoes

### ABSTRACT

Food contaminated with heavy metals causes serious consequences for human health, it is always the prime concern of any food safety control systems, even required through costly processes from sample collecting and evaluating contaminated components in the food samples. This study aimed to use ultrasound coupling with deep networks to assess the presence of heavy metals in sweet potatoes, while ultrasound is safe. To classify the acquired ultrasound data sets, we used deep networks that presently become a powerful tool and attract many researchers in order to recognize the data associating with the presence of lead sulfate in samples of sweet potatoes. For the 31 ultrasonic data sets of sweet potato samples acquired, the application of Neuron Network (NN) and Deep Boltzmann machine (DBM) as our target deep networks yielded the results showing that the accuracies of the NN was 62% for training set and 55% for testing set, and of DBM was 68% for training set and 65% for testing set, respectively.

Keywords: Food safety, Artificial intelligence, machine learning, neural network, Boltzmann machine, deep learning.

### 1. ĐẶT VẤN ĐỀ

Nhóm bệnh gây ra bởi thực phẩm không an toàn đang là một gánh nặng cho xã hội và kinh tế của mỗi quốc gia (WHO, 2017; Young & cs., 2016), và các nước trên thế giới đang ngày càng quan tâm chú trọng đến vấn đề này. Nâng cao

mức ATTP vì thế là vấn đề chính yếu cho sự ổn định an ninh, xã hội và sự phát triển kinh tế của mỗi nước. Do tình trạng ô nhiễm môi trường ngày càng nghiêm trọng hơn làm cho thực phẩm có thể nhiễm bị độc ở bất cứ khâu nào từ sản xuất đến tiêu dùng. Thêm vào đó, sự phát triển mạnh mẽ của hệ thống vận tải toàn cầu mà thực

phẩm có thể được vận chuyển trên thị trường và giữa các quốc gia một cách nhanh chóng, vì thế việc kiểm tra phát hiện thực phẩm có nguy cơ gây hại đến sức khỏe là một việc quan trọng và cần được giám sát chặt chẽ ở các khâu của chuỗi cung ứng thực phẩm.

Thực phẩm nhiễm độc bởi các kim loại nặng thuộc nhóm cần quan tâm nhất cho việc bảo vệ sức khỏe như đã được chỉ ra trong các báo cáo (WHO, 2017; Rahman & cs., 2014), bởi vì kim loại nặng tích lũy trong thực phẩm có tác động nghiêm trọng đến sức khỏe con người. Ước lượng mức độ nhiễm kim loại nặng tích lũy trong thực phẩm trước khi sử dụng là một yêu cầu quan trọng cho vấn đề ATTP. Thực phẩm có thể bị nhiễm kim loại từ các nguồn như đất, nước, không khí cho đến các khâu chế biến và chuẩn bị trước khi ăn (Toth & cs., 2016; Gan & cs., 2017). Thông thường, phân tích thành phần hóa chất sẽ chỉ ra các nguyên tố kim loại nặng có chứa trong thực phẩm. Tuy nhiên, vai trò đảm nhiệm phân tích chỉ được thực hiện ở các cơ quan chức năng đặt ở một số vùng trung tâm ở Việt Nam.

Trong báo cáo gần đây của Ngân hàng Thế giới về quản lý nguy cơ ATTP ở Việt Nam (WB, 2017) đã chỉ ra rằng vấn đề ATTP nằm trong mối quan tâm rất lớn của cả người tiêu dùng và người làm chính sách an toàn. Sẽ không thể giải quyết bài toán ATTP nếu không có sự kiểm soát chặt chẽ và một hệ thống giám sát an toàn chuỗi thực phẩm đầy đủ và hiệu quả. Báo cáo cũng nhấn mạnh những công việc chính cần làm để nâng mức ATTP. Đặc biệt, cần thiết lập một chương trình đồng bộ từ bộ phận tư nhân đến nhà nước cho việc nâng cao hệ thống giám sát ATTP ở Việt Nam.

Ở địa phương, theo báo cáo về ATTP và chương trình hành động trong năm 2018 của Tỉnh Quảng Ninh (Quangninh, 2017), thực phẩm nhiễm kim loại nặng nằm trong danh sách những nhiệm vụ quan tâm giám sát để đảm bảo ATTP trong phạm vi tỉnh trong giai đoạn 2016-2020. Hệ thống giám sát ATTP hiện tại của tỉnh cần thiết lập mạnh hơn nữa để kiểm soát sự phân phối thực phẩm không an toàn đồng thời nâng cao nhận thức của người tiêu

dùng về vấn đề giữ gìn vệ sinh ATTP. Về công việc hàng ngày, hệ thống giám sát ATTP của tỉnh thu thập những mẫu thực phẩm và gửi mẫu đi phân tích các thành phần gây hại cho sức khỏe ở trung tâm kiểm định ATTP của vùng theo chỉ định của Nhà nước. Rõ ràng công việc này cần thời gian vận chuyển mẫu mà không thể thông báo ngay kết quả chất lượng mẫu tại nơi lấy mẫu được, việc đó không thể đáp ứng ngay được với yêu cầu của thị trường phân phối thực phẩm cần xử lý nhanh.

Thêm nữa là, tình trạng ô nhiễm kim loại có nhiều nguy cơ tiềm ẩn trên địa bàn tỉnh Quảng Ninh. Quảng Ninh hiện đang có rất nhiều mỏ than, thậm chí cả các công trường mỏ đang khai thác lộ thiên. Sự khai thác mỏ là nguyên nhân chủ yếu dẫn đến ô nhiễm nguồn nước, đất và không khí. Với nền công nghiệp khai thác mỏ phủ rộng trên địa bàn tỉnh, các vấn đề hệ quả chẳng hạn như thực phẩm nhiễm độc đang là mối quan tâm rất lớn. Cho đến nay, ô nhiễm kim loại trên địa bàn tỉnh thu hút nhiều quan tâm. Nhiều nghiên cứu như đánh giá mức nhiễm kim loại và nguy cơ cho sức khỏe dân cư trên địa bàn tỉnh (Ha & cs., 2018) hoặc nghiên cứu đánh giá sự ảnh hưởng của ô nhiễm kim loại đến tăng trưởng của lúa gạo trồng gần khu vực mỏ Quang Ninh (Marquez & cs., 2018) đã được tiến hành.

Nghiên cứu đánh giá về vấn đề này, ở trong nước có nhiều nghiên cứu quan tâm đến ô nhiễm kim loại nặng và nguy cơ đến sức khỏe người dân như nghiên cứu của Ha & cs. (2018). Nghiên cứu này đã chỉ ra vấn đề ô nhiễm không khí có chứa kim loại nặng gây ra những nguy cơ ảnh hưởng đến sức khỏe người dân. Hay nghiên cứu của Marquez & cs. (2018) còn xem xét sự tích lũy của kim loại nặng ảnh hưởng đến sự sinh trưởng của cây lúa gạo. Vấn đề ô nhiễm kim loại nặng còn nằm trong mối quan tâm rất lớn của các cơ quan quản lý, điều đó thể hiện ở các báo cáo hàng năm (WB, 2017; Quangninh, 2017) ở cả trung ương và địa phương.

Sóng siêu âm là kỹ thuật được ứng dụng trong nhiều thiết bị bảo quản thực phẩm vì sự an toàn của nó với sức khỏe con người. Trong nghiên cứu của Danh & cs. (2017), nhóm nghiên cứu đã

sử dụng sóng siêu âm trong việc rửa rau củ quả và vệ sinh ATTP. Liên quan đến sử dụng sóng siêu âm trích xuất thành phần thực phẩm, Ly & cs. (2018) đã ứng dụng sóng siêu âm để xác định thành phần protein trong gạo. Trong kiểm soát ATTP, sóng siêu âm đã được dùng để phân tích tương quan giữa các tính chất lý sinh của thực phẩm trong các nghiên cứu của Young & cs. (2016), Awad & cs. (2012) và Sheshrao & cs. (2018). Tuy nhiên, dữ liệu thu được để phân tích là rất phức tạp và cần có phương pháp phân tích hiệu quả để tìm ra mối tương quan giữa thành phần hóa chất trong thức ăn với các mẫu sóng siêu âm quan sát được. Điều này đã được nghiên cứu và đánh giá trong các nghiên cứu trên.

Gần đây, các phương pháp của học máy và mạng học sâu có nhiều cải tiến cùng với công nghệ hỗ trợ tính toán đã có nhiều ứng dụng hiệu quả trong việc phân tích dữ liệu lớn và nó đang cuốn hút nhiều quan tâm (Bengio, 2016). Trong các nghiên cứu của Kuong & cs. (2017, 2018a, 2018b), đã sử dụng phương pháp của học máy để nhận dạng các mẫu sóng siêu âm trong y tế. Với nhiều phương pháp hiệu quả trong khoa học máy tính được nghiên cứu và chứng minh gần đây giải quyết cho bài toán có dữ liệu lớn, điều đó hướng chúng tôi đến sử dụng sóng siêu âm và phương pháp học sâu cho bài toán nhận biết mức độ tích lũy thành phần kim loại nặng trong khoai lang trong phòng thí nghiệm.

## 2. PHƯƠNG PHÁP NGHIÊN CỨU

### 2.1. Thu thập dữ liệu

Trong nghiên cứu gần đây của Roro & cs. (2019), đã xem xét mức độ tích lũy của kim loại nặng tích trong khoai lang ở vùng mà hệ thống thủy lợi có sự ảnh hưởng của công nghiệp dệt may và những hệ quả đến sức khỏe của người dân xung quanh. Nghiên cứu đó cho thấy sự tồn dư kim loại nặng gây hại như Cờ rôôm (Cr), Đồng (Cu), Chì (Pb), Kẽm (Zn) trong khoai lang sau thu hoạch. Tương tự, ở Việt Nam, sự phát triển của các khu công nghiệp, nhà máy xí nghiệp dệt may trong những năm gần đây đã được đánh giá trong các nghiên cứu của Ghaly & cs. (2014) và Duy & cs. (2019) về tình hình ô nhiễm từ nền công nghiệp dệt may. Điều đó khiến chúng tôi đi

đến xem xét nghiên cứu bước đầu những mẫu khoai lang với sự tích lũy kim loại gây hại này.

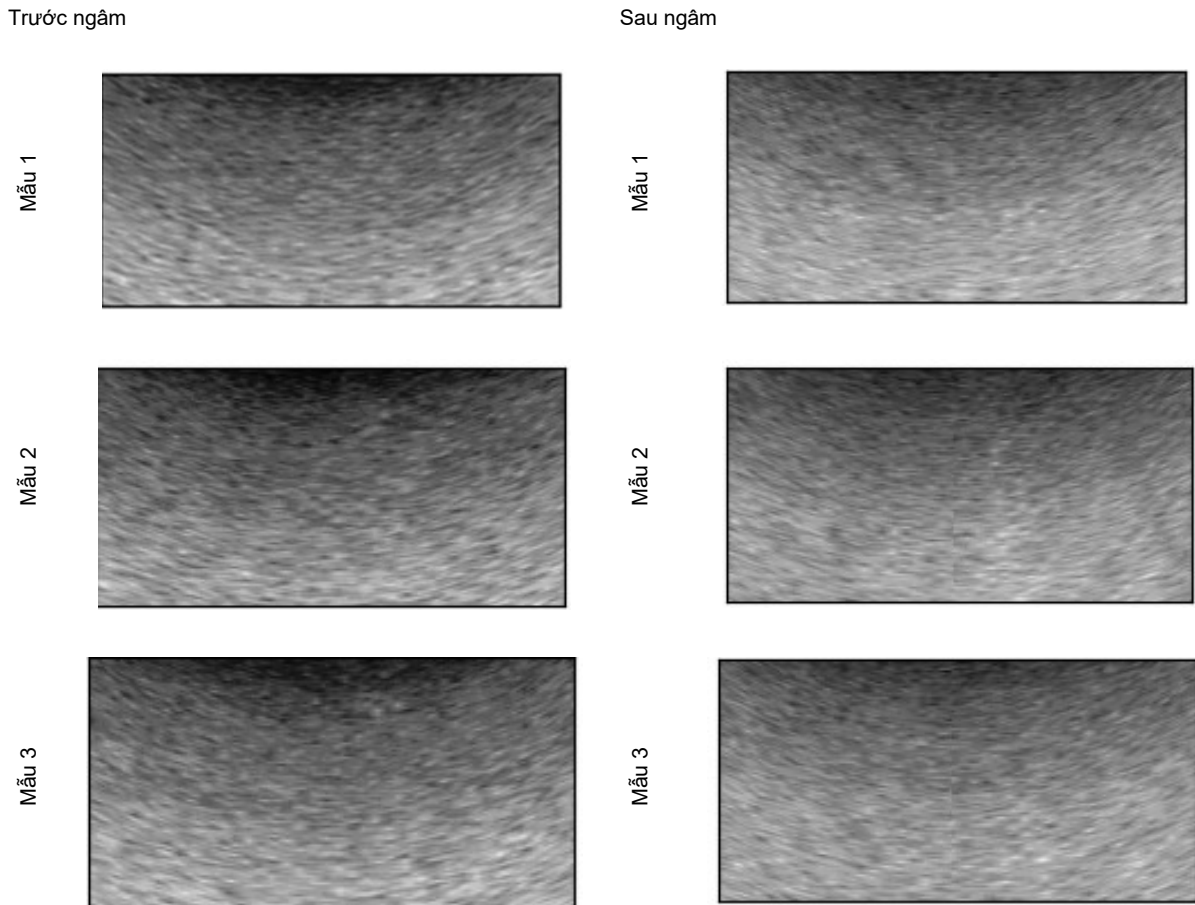
Trong nghiên cứu này, chúng tôi tiến hành lấy dữ liệu sóng siêu âm truyền qua các mô của khoai lang bởi máy siêu âm Chison Eco. Thiết bị siêu âm sẽ phát ra sóng siêu âm và nhận tín hiệu phản hồi lại từ các mô khoai. Trên mỗi mẫu, dữ liệu chúng tôi thu được là các sóng siêu âm phản hồi lại như vậy. Mục đích của chúng tôi là nhận biết sự khác biệt của các mẫu sóng siêu âm này giữa các mẫu khoai lang trước và sau khi cho hấp thụ chì sunfat (ngâm mẫu khoai lang trong dung dịch chì sunfat).

Để xác định nồng độ của dung dịch chì sunfat cho phù hợp, chúng tôi dựa theo nghiên cứu các nghiên cứu của Thanh & cs. (2019) và Huong & cs. (2007), các nghiên cứu đó đánh giá mức độ ô nhiễm của các thành phần hóa chất gây hại trong nguồn nước ở một số khu vực Hà Nội. Theo kết quả đó, hàm lượng dung dịch của chì (Pb) trong nguồn nước sông được đánh giá ở quanh khu vực Hà Nội lớn hơn mức 1 mg/l, thậm chí có thời điểm đến hơn 6 mg/l. Trong thí nghiệm này chúng tôi lấy dung dịch chì sunfat với nồng độ 4 mg/l và cho các mẫu khoai ngâm trong dung dịch đó. Sau khoảng thời 24 giờ lại được chúng tôi đưa ra thu dữ liệu sóng siêu âm để xem xét khả năng khác biệt của các mẫu sóng trước và sau khi cho hấp thụ với dung dịch chì sunfat.

Dữ liệu các sóng siêu âm trên mỗi mẫu sau khi tiền xử lý loại bỏ các vùng gần chịu cộng hưởng của thiết bị, dữ liệu được biểu diễn cho quan sát được thể hiện ở dạng ảnh B-mode. Hình 1 biểu diễn ảnh B-mode của 3 trường hợp trong số các mẫu chúng tôi xem xét.

### 2.2. Mạng học sâu

Với mục đích sử dụng mạng học sâu như là bộ phân lớp mục tiêu cho việc nhận biết các mẫu siêu âm. Trong phần này, chúng tôi trình bày về mạng nơron truyền thẳng (Hagan & cs, 1996) được xem như mạng học sâu cơ bản dành cho việc học có giám sát đã được phát triển từ lâu. Tiếp đến là mạng DRBM (Hilton, 2012; Hugo & cs., 2012; Lecun & cs., 2015) được phát triển gần đây, nó thu hút sự quan tâm của nhiều nhà nghiên cứu trong việc học giám sát và học không có giám sát.



Hình 1. Ảnh B-mode của dữ liệu sóng siêu âm từ các mẫu

### 2.2.1. Mạng nơ ron

Hình 2 mô tả cấu trúc mạng nơ ron nhiều tầng, ở đó  $x$  là dữ liệu đầu vào (input) có nhãn là  $y$  nằm ở tầng trên cùng. Giữa  $x$  và  $y$  được kết nối thông qua nhiều tầng  $h$  (từ  $h_1$  đến  $h_n$ ) còn gọi là tầng ẩn. Kết nối giữa các tầng được thể hiện bởi các trọng số kết nối  $w^k = w_{ij}^k$ , ở đó  $w_{ij}^k$  là trọng số kết từ nối đơn vị nơ ron  $i$  ở tầng  $k - 1$  đến nơ ron  $j$  ở tầng  $k$  (với  $k = 0$  là tầng input,  $k = n + 1$  là tầng nhãn  $y$ ).

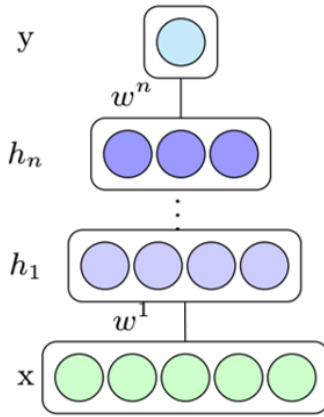
Việc huấn luyện cho mạng nơ ron dựa trên nguyên lý cực tiểu hóa hàm chi phí bởi công thức  $F(f^m(x) - y) \rightarrow \min$ , ở đó  $F$  là hàm chi phí xác định sự sai lệch giữa giá trị nhãn đầu ra của mạng với giá trị nhãn của dữ liệu,  $f$  là hàm kích hoạt. Chi tiết về mạng và thuật toán huấn luyện của mạng nơ ron được trình bày cụ thể bởi Hagan & cs. (1996).

### 2.2.2. Mạng restricted Boltzmann machine (RBM) chuẩn

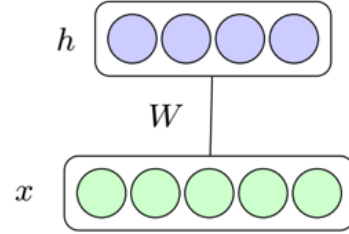
Mạng restricted Boltzmann machine (RBM) là một kiểu mạng nơ ron học phân bố xác suất của dữ liệu đầu vào, ở đó về kiến trúc, nó sử dụng các biến trong tầng ẩn  $h = (h_1, h_2, \dots, h_H)$  để học phân phối của các biến biểu diễn dữ liệu cần học hay dữ liệu input  $x = (x_1, x_2, \dots, x_N)$ . Mỗi đơn vị  $x_i$  có sự kết nối với trọng số  $w_{ij}$  tới mỗi đơn vị  $h_j$ . Không có sự kết nối giữa các đơn vị trong cùng tầng ẩn hay cùng tầng dữ liệu. Các trọng số  $b_i$  và  $c_j$  phản ánh mức độ tác động của mỗi đơn vị  $x_i$  và  $h_j$  tương ứng trong mạng. Mạng RBM học thông qua việc điều chỉnh hàm năng lượng xác định bởi công thức (1):

$$E(x, h) = -\sum_{i,j} w_{ij} x_i h_j - \sum_i b_i x_i - \sum_j c_j h_j \quad (1)$$

Phân phối đồng thời  $P(x, h)$  của  $x$  và  $h$  được xác định bởi phương trình (2):



Hình 2. Mạng neuron



Hình 3. Mạng RBM

$$P(x, h) = \frac{\exp(-E(x, h))}{Z} \quad (2)$$

ở đó  $Z$  là hằng số chuẩn hóa. Xác suất có điều kiện cho các đơn vị  $h_j$  và  $x_i$  được xác định dựa theo phân phối Boltzmann bởi (3) và (4)

$$P(h_j | x) = \text{sigm} \left( \sum_i w_{ij} x_i + c_j \right) \quad (3)$$

và

$$P(x_i | h) = \text{sigm} \left( \sum_j w_{ij} h_j + b_i \right) \quad (4)$$

ở đó  $\text{sigm}(x) = \frac{1}{1 + e^{-x}}$  là hàm sigmoid.

Mạng RBM chuẩn được mô tả như ở hình 3.

Mạng RBM chuẩn được trang bị thuật toán CD-k (Hilton, 2012), nó cho phép mạng có thể học không giám sát (unsupervised learning) phân phối đồng thời giữa các đơn vị tầng ẩn và tầng input. Ở một nghĩa nào đó thì tầng ẩn tham gia vào học phân phối của input và đóng vai trò làm rút ngắn số chiều của tầng input.

### 2.2.3. Mạng restricted Boltzmann machine phân lớp

Mạng RBM phân lớp (classification restricted Boltzmann machine- classRBM) là một trường hợp mở rộng của RBM bằng cách thêm các đơn vị mã hóa cho nhãn tương ứng với các input (Hugo & cs., 2012). Cụ thể, nếu các input  $x$  có nhãn là  $k$  trong số  $K$  lớp của dữ liệu

thì lớp cho nhãn gồm  $K$  đơn vị xác định bởi qui tắc “one-hot”, hay đơn vị thứ  $k$  có giá trị bằng 1 còn lại bằng 0. Khi đó tương tự như mạng RBM chuẩn, hàm năng lượng được cho bởi các phương trình (5) dưới đây:

$$E(x, h, y_k) = -\sum_{i,j} w_{ij} x_i h_j - \sum_i b_i x_i - \sum_j c_j h_j - \sum_j U_{kj} h_j - d_k \quad (5)$$

ở đó  $U_{kj}$ ,  $d_k$  là các trọng số kết nối với các đơn vị ẩn và trọng số của đơn vị nhãn tương ứng. Không có kết nối giữa các đơn vị nhãn với các đơn vị input. Phân phối đồng thời của các đơn vị được xác định bởi:

$$P(x, h, y) = \frac{\exp(-E(x, h, y))}{Z} \quad (6)$$

ở đó  $Z$  là hằng số chuẩn hóa. Các xác suất có điều kiện được xác định bởi:

$$P(h_j | x, y_k) = \text{sigm} \left( \sum_i w_{ij} x_i + U_{kj} + c_j \right) \quad (7)$$

và

$$P(x_i | h) = \text{sigm} \left( \sum_j w_{ij} h_j + b_i \right) \quad (8)$$

$$P(y_k | h) = \frac{\exp(\sum_j U_{kj} h_j + d_k)}{\sum_l \exp(\sum_j U_{lj} h_j + d_l)} \quad (9)$$

Xác suất hậu nghiệm cho việc xác định phân lớp là:

$$P(y_k | h) = \frac{\exp[d_k + \sum_j f(\sum_i w_{ij} x_i + U_{kj} + c_j)]}{\sum_{y_l} \exp[d_l + \sum_j f(\sum_i w_{ij} x_i + U_{lj} + c_j)]} \quad (10)$$

ở đó  $f(x) = \log(1 + \exp(x))$  là hàm softplus. Mô hình classRBM được minh họa ở hình 4.

Như vậy, khi trang bị thêm tầng nhãn  $y$  thì mạng classRBM phục vụ cho việc học có giám sát (supervise learning). ClassRBM đã được chứng tỏ khả năng huấn luyện hiệu quả với các thuật toán được trang bị như đã được trình bày bởi Hugo & cs. (2012).

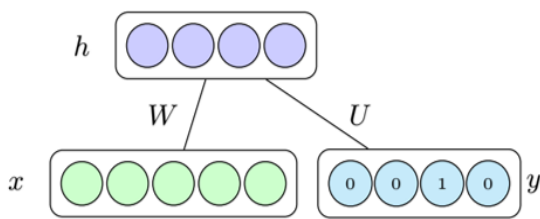
**2.2.4. Mạng deep Boltzmann machine và học sâu**

Mạng deep Boltzmann machine (DBM) là sự xếp chồng của nhiều RBMs (Lecun & cs., 2015). Với thuật toán hiệu quả CD-k, nó cho phép tầng ẩn  $h$  tham gia vào học phân phối của input, đồng thời tầng ẩn lại tham gia như là một input cho tầng ẩn tiếp theo. Đó là cơ sở đẩy mạnh sự phát triển mạng học sâu. Trong nghiên cứu này chúng tôi sử dụng mạng DBM với 2 tầng ẩn ở đó tầng ẩn thứ hai có sự tham gia của mạng classRBM, nghĩa là, việc huấn luyện ở mạng thứ 2 là học có giám sát kết hợp với nhãn để nhận diện các đoạn mẫu âm. Mô hình mạng DBM được sử dụng trong nghiên cứu này được mô tả ở hình 5.

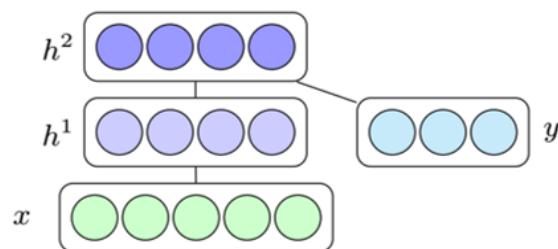
**2.3. Kết quả phân lớp**

Trong nghiên cứu này, chúng tôi thu thập được 31 mẫu dữ liệu siêu âm gồm dữ liệu của các mẫu khoai chưa cho hấp thụ qua dung dịch chì sunfat và sau khi ngâm qua dung dịch chì sunfat với thời gian 24 giờ và 48 giờ. Tất cả các mẫu sau khi ngâm được chúng tôi xếp vào một lớp làm đối chứng với các mẫu chưa qua dung dịch. Trên mỗi ảnh dữ liệu thô, chúng tôi lấy ảnh có kích thước  $200 \times 200$  ở giữa tâm làm ảnh thuộc tính cho đánh giá phân lớp.

Hai mạng phân lớp chúng tôi sử dụng là nơron và mạng DBM, trong đó chúng tôi để cấu hình ai mạng cùng có 2 tầng ẩn có kích thước là 300 và 100. Nghĩa là ở tầng ẩn thứ nhất có 300 nơron tham gia và tầng ẩn thứ 2 có 100 nơron tham gia. Với mạng nơron, chúng tôi sử dụng thuật toán backpropagation (Hagan & cs., 1996). Với mạng DBM thì giữa tầng input  $x$  và tầng ẩn  $h^1$  là mạng RBM được huấn luyện bởi thuật toán CD-1 như giới thiệu bởi Hilton (2012). Giữa tầng ẩn  $h^1$  và tầng ẩn  $h^2$  có sự tham gia của lớp nhãn  $y$  hay là mạng classRBM và được huấn luyện bởi thuật toán học có giám sát của classRBM được trình bày bởi Hugo & cs. (2012). Ở đây các tham số được khởi tạo ban đầu rất nhỏ trong khoảng  $\pm 10^{-5}$ .



Hình 4. Mạng classRBM



Hình 5. Mạng DBM

**Bảng 1. Kết quả độ chính xác trong huấn luyện và kiểm tra của DBM và NN**

	DBM	NN
Độ chính xác với dữ liệu huấn luyện	68%	65%
Độ chính xác với dữ liệu kiểm tra	62%	55%

Chúng tôi sử dụng phương pháp 5-fold cho việc phân chia dữ liệu huấn luyện và dữ liệu kiểm tra. Để đánh giá, so sánh kết quả huấn luyện và kiểm tra của hai mạng, chúng tôi sử dụng độ chính xác nghĩa là tỉ lệ phần trăm giữa số mẫu được nhận dạng đúng trên tổng số mẫu đầu vào. Kết quả biểu diễn khả năng học của cả hai bộ phân lớp mục tiêu trên được thể hiện ở bảng 1.

#### 4. KẾT LUẬN

Nghiên cứu này nhằm sử dụng phương pháp học máy trong việc nhận biết các mẫu sóng siêu âm trong việc phát hiện các dư lượng kim loại nặng tích lũy trong khoai lang. Nó phản ánh tính cần thiết trong điều kiện các vấn đề ATTP đang được quan tâm của xã hội cùng với ô nhiễm môi trường. Mặc dù dữ liệu cho nghiên cứu còn hạn chế, tuy nhiên kết quả của nghiên cứu cho thấy khả năng ứng dụng các phương pháp của mạng học sâu với dữ liệu sóng siêu âm trong việc trợ giúp giám sát sự tích lũy kim loại gây hại trong rau củ.

Sử dụng mạng DBM trong nghiên cứu này cũng đã chứng tỏ được khả năng nhận dạng tốt hơn so với mạng nơron truyền thống. Cho dù vậy, việc so sánh với các phương pháp phân lớp khác cũng là cần thiết cho sự phát triển tiếp theo trong việc đi đến cài đặt phương pháp phân lớp hiệu quả nhằm nâng cao độ chính xác của nhận dạng các mẫu.

#### LỜI CẢM ƠN

Nghiên cứu này được thực hiện trong khuôn khổ của đề tài “Sử dụng sóng siêu âm nhận biết sự tích lũy kim loại nặng trong khoai lang”, mã số: T2019-10-56, cấp bởi Học viện Nông nghiệp Việt Nam. Chúng tôi cũng xin cảm ơn sự trợ giúp của TS. Sử Thanh Long quản lý trung tâm Gaia đã tạo điều kiện cho nhóm nghiên cứu chúng tôi trong việc sử dụng thiết bị, thiết kế và thu thập dữ liệu sóng siêu âm cho nghiên cứu này.

#### TÀI LIỆU THAM KHẢO

Awad T.S., Moharram H.A., Shaltout O.E., Asker D. & Youssef M.M. (2012). Applications of ultrasound

in analysis, processing and quality control of food: A review. *Food research international*. 48(2): 410-427.

- Duy N.N., Phu V.D., Lan N.T.K., Duoc N.T., Hien N.Q., Hiep B.N., Han B.N. & Ha B.M. (2019). Treatment of real textile wastewater using electron beam irradiation. *Acta Chemica Iasi*. 27(2): 303-316.
- Gan Y., Wang L., Yang G., Dai J., Wang R. & Wang W. (2017). Multiple factors impact the contents of heavy metals in vegetables in high natural background area of China. *Chemosphere*. 184: 1388-1395.
- Ghaly A.E., Ananthashankar R., Alhattab M. & Ramakrishnan V. (2014). Production, characterization and treatment of textile effluents: a critical review. *Journal of Chemical Engineering and Process Technology*. 5(1): 1000182.
- Hagan M.T., Demuth H.B. & Beale M. (1996). *Neural network design*. Pws, Boston.
- Ha P.T., Chinh P.M., Cuong D.D., Ly L.T.M., Thinh N.V. & Thai P.K. (2018). Elemental concentrations in roadside dust along two national highways in northern Vietnam and the health risk implication. *Archives of environmental contamination and toxicology*. 74(1): 46-55.
- Hilton E.G. (2012). *A practical guide to training restricted Boltzmann machines*. Lecture Notes in Computer Science, Springer Berlin. 7700: 599-619.
- Hugo L., Michael M., Razvan P. & Yoshua B. (2012). Learning algorithms for the classification restricted Boltzmann machine. *Machine Learning Research*. 13(1): 643-669.
- Huong N.T.L., Ohtsubo M., Li L. & Higashi T. (2007). Heavy metal pollution of the To-Lich and Kim-Nguu River in Hanoi city and the industrial source of the pollutants. *Journal-Faculty of Agriculture Kyushu University*. 52(1): 141.
- Kuong N.T., Uchino E. & Suetake N. (2017). IVUS tissue characterization of coronary plaque by classification restricted Boltzmann machine. *Journal of Advanced Computational Intelligence and Intelligent Informatics*. 21(1): 67-73.
- Kuong N.T., Uchino E. & Suetake N. (2018a). Recognition of coronary atherosclerotic plaque tissue on intravascular ultrasound images by using misclassification sensitive training of discriminative restricted Boltzmann machine. *Journal of Biomimetics, Biomaterials and Biomedical Engineering*. 37: 85-93.
- Kuong N.T., Uchino E. & Suetake N. (2018b). Coronary plaque classification with accumulative training of deep Boltzmann machines. *ICIC Express Letters*. 12(9): 881-886.

- Ly H.L., Tran T.M.C., Tran T.T.T., Ton N.M.N. & Le V.V.M. (2018). Application of ultrasound to protein extraction from defatted rice bran. *International Food Research Journal*. 25(2).
- Lecun Y., Yoshua B. & Hinton E.G. (2015). Deep learning. *Nature*. 521(7553): 436-444.
- Marquez J. E., Pourret O., Faucon M.P., Weber S., Hoa H.T.B. & Martinez R. (2018). Effect of cadmium, copper and lead on the growth of rice in the coal mining region of Quang Ninh, Cam Pha (Vietnam). *Sustainability*. 10(6): 1758.
- Quangninh (2017). Hội nghị tổng kết thực hiện Năm cao điểm hành động vệ sinh an toàn thực phẩm trong lĩnh vực nông nghiệp trên địa bàn tỉnh Quảng Ninh năm 2017 và triển khai nhiệm vụ trọng tâm năm 2018. Truy cập từ <https://www.quangninh.gov.vn/So/sonongnghiep/Trang/ChiTietTinTuc.aspx?nid=5904> ngày 25/09/2019.
- Rahman M.A., Rahman M.M., Reichman Lim S.M., R.P. & Naidu R. (2014). Heavy metals in Australian grown and imported rice and vegetables on sale in Australia: health hazard. *Ecotoxicology and environmental safety*. 100: 53-60.
- Roro A.G., Beshir H. M., Terfa M. T., Tesfaye B., Nadew F. & Olango T. M. (2019). Accumulation of heavy metals in soil and sweet potato (*Ipomoea batatas*) irrigated with treated and untreated textile effluents. *Journal of Applied Sciences*. 19(9): 837-847.
- Sheshrao K. & Jai P.P. (2018). An elementary review on principles and applications of modern non-conventional food processing technologies. *International Journal of Current Microbiology and Applied Sciences*. 7(5): 838-849.
- Thanh K.N., Manh H.N., Kim C.T., Ahmed M.B., Huang Y. & Zhou J.L. (2019). Chemical and microbiological risk assessment of urban river water quality in Vietnam. *Environmental geochemistry and health*. 41(6): 2559-2575.
- Trần Hữu Danh, Lương Vinh Quốc Danh, Trần Thanh Quang, Nguyễn Thị Trâm, Huỳnh Minh Trí & Trần Hữu Nghi (2017). Bể rửa ứng dụng sóng siêu âm. *Tạp chí Khoa học, Trường Đại học Cần Thơ*. 52: 46-53.
- Toth G., Hermann T., Silva M.D. & Montanarella L. (2016). Heavy metals in agricultural soils of the European Union with implications for food safety. *Environment international*. 88: 299-309.
- Young I. & Waddell L. (2016). Barriers and facilitators to safe food handling among consumers: A systematic review and thematic synthesis of qualitative research studies. *PloS One*. 11(12).
- WHO World Health Organization (2017). Food safety. Retrieved from <http://www.who.int/en/news-room/fact-sheets/detail/food-safety> on April 29, 2019.
- World Bank (2017). Vietnam food safety risks management: Challenges and Opportunities, World Bank, Hanoi. Retrieved from <https://openknowledge.worldbank.org/handle/10986/26412> on April 29, 2019.