

# Determining the Degree of Malignancy on Digital Mammograms by Artificial Intelligence Deep Learning

<sup>1\*</sup>*Sang-Bock Lee*, <sup>2</sup>*Hwunjae Lee*, <sup>3</sup>*V. R. Singh*

Received : 25 June 2020 / Accepted : 15 October 2020 / Published online 28 December 2020

©The Author(s) 2020

**Abstract** In this paper, we propose a method for determining degree of malignancy on digital mammograms using artificial intelligence deep learning. Digital mammography is a technique that uses a low-energy X-ray of approximately 30 KVp to examine the breast. The goal of digital mammography is to detect breast cancer in an early stage by identifying characteristic lesions such as microcalcifications, masses, and architectural distortions. Frequently, microcalcifications appear in clusters that increase ease of detection. In general, larger, round, and oval-shaped calcifications with uniform size have a higher probability of being benign; smaller, irregular, polymorphic, and branching calcifications with heterogeneous size and morphology have a higher probability of being malignant. The experimental images for this study were selected by searching for "mammogram" in the NIH database. The images were converted into JPEG format of 256 X 256 pixels and saved. The stored images were segmented, and edge detection was performed. Most of the lesion area was low frequency, but the edge area was high frequency. DCT was performed to extract the features of the two parts. Similarity was determined based on DCT values entered into the neural network. These were the findings of the study:

- 1) There were 6 types of images representing malignant tumors.
- 2) There were 2 types of images showing benign tumors.
- 3) There were two types of images demonstrating tumors that could worsen into malignancy.

Medical images like those used in this study are interpreted by a radiologist in consideration of pathological factors. Since discrimination of medical images by AI is limited to image information, interpretation by a radiologist is necessary. To improve the discrimination ability of medical images by AI, extracting accurate features of these images is necessary, as is inputting clinical information and accurately setting targets. Study of learning algorithms for neural networks should be continued. We believe that this study concerning recognition of cancer on digital breast images by AI deep learning will be useful to the radiomics (radiology and genomics) research field.

**Keywords:** Breast cancer, Morphology, Feature extraction, Neural networks

---

<sup>1\*</sup>*Sang-Bock Lee* (✉) **corresponding author**  
(62271)Department of Radiology, Nambu University,  
Nambudae-ro, Gwangsan-gu, Gwangju, Korea

<sup>2</sup>*Hwunjae Lee*

<sup>1</sup>Department of Radiology, College of Medicine, Yonsei University, 50-1, Yonsei-ro, Seodaemun-gu, Seoul,

<sup>3</sup>*V. R. Singh*

Director, PDM University, India  
e-mail: vr-singh@ieee.org

## I. Introduction

Research on the principles of radiation discovery and generation in various spectra laid the foundation for modern physics, leading to quantum physics and the theory of relativity<sup>[2]</sup>.

In the late 20th century, radiology was combined with computer technology to develop advanced imaging equipment such as CT, MRI, and PET. These imaging techniques are used to evaluate the anatomy and physiological characteristics of the human body<sup>[3]</sup>. The 4th Industrial Revolution was introduced at the 2016 World Economic Forum and is explained by a variety of new technologies that integrate the physical, biological, and digital worlds based on “big” data and affect all sectors such as economics and industry<sup>[4]</sup>. With the 4th Industrial Revolution, the medical environment is changing from patient diagnosis and treatment to personalized precision medicine, and the center of change is artificial intelligence (AI) technology<sup>[5]</sup>. AI technology aims to be used in the entire process of 'before, during, and after' patient examination using diagnostic imaging equipment, and this will be the basis for precise medical imaging. Breast cancer is the world's leading cancer and cause of death among women. Methods of examining breast cancer include digital breast imaging, MRI, and ultrasound. All the information from these methods are stored and managed as “big” data on a computer<sup>[6]</sup>.

In this paper, we propose a method to recognize cancer in digital breast images through in-depth artificial intelligence. The proposed method will contribute to development of personalized precision medicine.

## II. Materials and Methods

### 1. Digital mammogram

Advances in computer technology and development of new digital imaging detectors have enabled digital mammography, which uses low-energy X-ray of approximately 30 KVp to examine the breast. The goal of digital mammography is to detect breast

cancer at an early stage by identifying characteristic masses or micro-calcifications<sup>[7]</sup>. Like all X-ray images, mammograms are imaged using ionizing radiation, and MO generally uses lower energy X-rays than those for bones (K shell X-ray energy of 17.5 and 19.6 keV) and Rh (20.2 and 22.7 keV)<sup>[8]</sup>. The image is analyzed for detection of abnormalities. Ultrasound, ductogram, PET, and MRI-based breast imaging play a secondary role. Ultrasound is commonly used for new evaluations of masses found during mammography or palpation of masses not visible on X-ray mammograms. MRI evaluates suspected disease and is useful in enabling surgical method changes such as those associated with breast-conserving mastectomy in preoperative evaluation of patients with confirmed breast cancer<sup>[9]</sup>. Recently, the tomosynthesis technique has been used as a digital mammography method<sup>[10]</sup>.

### 2. Breast cancer basic morphology

Breast cancer has some characteristic lesions such as microcalcifications, masses, and architectural distortions. Breast asymmetry can also be a breast cancer indicator. Microcalcifications are small-sized lesions, typically in the range of 0.05 to 1 mm, and relatively difficult to detect. Microcalcifications are bright on imaging and have various sizes, shapes, and distributions. In some cases, these lesions have low contrast due to a reduced intensity difference between suspicious areas and the surroundings. Another reason for difficult detection of these lesions is the proximity to the surrounding tissues. In dense tissues, suspicious areas are nearly invisible as a result of tissue superimposition. Some anatomic structures such as fibrous strands, breast borders, or hypertrophied lobules are similar to microcalcifications on mammographic imaging<sup>[11]</sup>. Frequently, microcalcifications appear in clusters, resulting in easier detection<sup>[12]</sup>. There is a high

correlation between presence of microcalcifications and breast cancer, particularly when the microcalcifications are clustered. Therefore, accurate detection of microcalcifications is essential to early detection of the majority of breast cancers<sup>[13]</sup>. In general, larger, round, and oval-shaped calcifications with uniform size have a higher probability of being benign; smaller, irregular, polymorphic, and branching calcifications with heterogeneous size and morphology have a higher probability of being malignant<sup>[14]</sup>. Masses appear as dense regions of

different sizes and properties and can be circular, oval, lobular, or irregular and spiculated. Mammographic masses can be described as follows:

- circumscribed, which are well-defined with distinctly demarcated borders;
- obscured, which are hidden by superimposed or adjacent tissue;
- micro-lobulated, which have undulating circular borders;
- ill-defined, which have poorly defined scattered borders;
- spiculated/stellate, which are radiating thin lines.

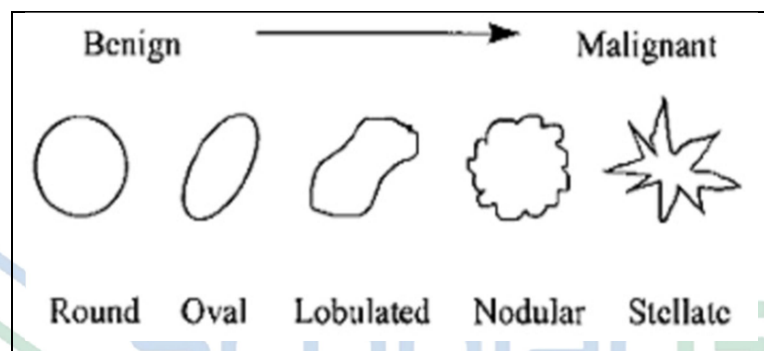


Figure 1. Morphologic spectrum of mammographic masses<sup>[15]</sup>

### 3. Image segmentation and edge detection

In digital image processing, image segmentation is the process of dividing a digital image into multiple segments to simplify the representation of the image and convert it to more meaningful and easier to analyze data. More precisely, image segmentation is the process of assigning labels to all pixels in an image so that pixels with the same label share certain characteristics<sup>[16]</sup>. The result of image segmentation is a set of segments that represents the entire image or a set of outlines extracted from the image. All pixels in a segment are similar in some way such as color, intensity, or texture. Adjacent areas differ significantly with respect to the chosen characteristics. Image segmentation is commonly used to find objects and boundaries in an image. The

simplest image segmentation method is called the threshold method<sup>[17]</sup>. This method converts a grayscale image to a binary image based on threshold, the choice of which is key to detection<sup>[17]</sup>. Several popular methods are used in industry, including the maximum entropy method, balanced histogram threshold, Otsu method, and k-means clustering<sup>[17]</sup>. Edge detection involves a variety of mathematical methods that aim to identify points in a digital image in which image brightness is discontinuous. The point at which the image brightness changes rapidly consists of curved segments, commonly called edges<sup>[18]</sup>. Identification of a discontinuity in a one-dimensional signal is called step detection, and identification of signal discontinuity over time is called change detection<sup>[19]</sup>. Edge detection is the primary tool in image processing, especially feature

detection and feature extraction. Edge thinning is a technique used to remove unwanted spurious points on the edges of an image. This technique is used after filtering the “noise” from the image, detecting the edges using the edge operator, and smoothing the edges to an appropriate threshold<sup>[19]</sup>. This process creates an edge element that is 1 pixel thick by removing all unwanted points and applying them

carefully.

The advantages of edge detection are:

1. Sharp and thin edges increase the efficiency of object recognition.
2. Use of Hough transformation to detect lines and ellipses results in high-quality thinning.
3. When an edge is a boundary of a segment and becomes thinner, image parameters are easily obtainable without the need for algebra<sup>[20]</sup>.

Table 1. Segmentation and edge detection function

```
% Segmentation and edge detection function
I = imread('Exam-image');
subplot(1,3,1);imshow(I); title('Original Image');
[~,threshold] = edge(I,'sobel'); fudgeFactor = 0.5;
BWs = (I >= 127); % threshold 127
se90 = strel('line',3,90); se0 = strel('line',3,0);
BWsdil = imdilate(BWs,[se90 se0]);
BWdfill = imfill(BWsdil,'holes');
BWnobord = imclearborder(BWdfill,4);
seD = strel('diamond',1);
BWfinal = imerode(BWnobord,seD);
BWfinal = imerode(BWfinal,seD);
subplot(1,3,2); imshow(BWfinal);
title('Segmented Image');
subplot(1,3,3); imshow(edge(BWfinal));
title('Edge detection');
```

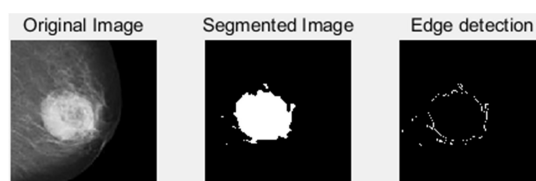


Figure 2. Segmentation and edge detection of an image

#### 4. Feature extraction by DCT

Feature extraction by DCT is an orthogonal transformation coding method using a discrete cosine function as a coefficient for transforming an image

signal on a time axis onto a frequency axis. In 1988, ITU-T was adopted as the video compression technology of the H.261 encoding method for conference calls. ITU-T was subsequently adopted



by MPEG, the international standard for video compression, and is now the mainstream of high-efficiency encoding and compression of video<sup>[21]</sup>. Discrete cosine transformation (DCT) decomposes the image signal on the time axis and transforms it into a small, low-frequency domain with multiple signal powers. Since the power of the image signal is concentrated in the low-frequency region, quantizing with proper allocation produces the entire bit<sup>[22]</sup>. The result can be compressed with a small amount of data.

Though optimized transformation includes the Karen-Lube transformation (KLT), for video signals, DCT can achieve high coding efficiency approximating that of the KLT<sup>[23]</sup>. In this study, the DCT was extracted as follows. Since energy is concentrated in the low-frequency region, many kinds of feature parameters of DCT coefficients are extracted there. Table 2 shows the block location of the extraction amplitude characteristic parameters.

Table 2. DCT function

```

% DCT Function
in_img=imread('Exam_Image');
fileout=fopen('Out-File','w');
imshow(in_img);
dct_img=rgb2gray(in_img); dct_img=dct2(dct_img);
figure; imshow(log(abs(dct_img)));
disp('starting extraction step for summations of each blocks.....');
disp('...'); sum=zeros(8,8);
for h=0:7
    for v=0:7
        hap=0;
        for hp=1:16
            for vp=1:16
                hap=hap+abs(dct_img(h*16 + hp , v*16 + vp));
            end
        end
        sum(v+1, h+1)=hap;
    end
end
disp('normalization about summaiton of each blocks');
disp('...'); sumline=zeros(1,32);
line=0;
for m=3:8
    k=0;
    for n=1:10
        m=m-1;
        k=k+1;
        if m==0 break
        end
        line=line+1;
        sumline(1,line)=sum(m,k);
    end
    if n==32 break
    end
end
Max = -10000;
Min = 10000;
for su=(1:32)
    if ( sumline(su) > Max )
        Max = sumline(su);
    end;
    if ( sumline(su) < Min )
        Min = sumline(su);
    end;
end;
end;

```

```

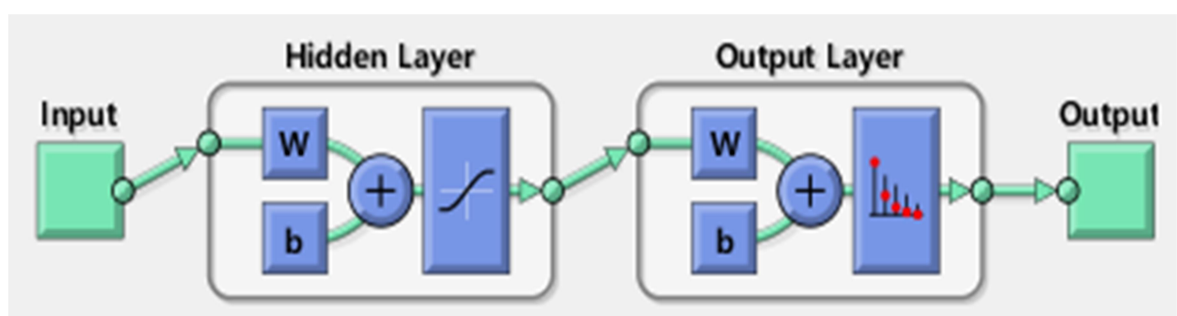
for su=(1:32)
    sumline(su) = ((sumline(su)- Min) / (Max -Min)) - 0.5;
end;
disp('saving normaliziton data for summation of each blocks');
disp('...'); for su=(1:32)
    fprintf(fileout,'%10.4f',sumline(su));
end;
fclose(fileout);

```

## 5. Neural network

Artificial intelligence (AI) is generally limited to a specific task. For example, most AI technologies are applied to a specific task such as automatic language translation, image recognition, speech recognition, and development and analysis of artificial neural network structures can be said to be artificial intelligence in a narrow sense<sup>[24]</sup>. Artificial intelligence is being applied to various fields with development and power of technology related to the 4<sup>th</sup> Industrial Revolution. In particular, use of AI in the medical field has spread rapidly. To date, representative uses of AI in the medical field are related to medical services in various fields ranging from secondary analysis of medical data to predicting and diagnosing diseases and analysis of medical

images<sup>[25]</sup>. In general, AI-based medical devices contain software that can improve performance by analyzing medical data and medical devices including the same. Artificial intelligence-based medical systems can be viewed as competitive with healthcare professionals, but use of AI in the medical field should be complementary to, rather than competitive with, the work of medical personnel<sup>[25]</sup>. Use of AI-based medical systems is rapidly expanding, particularly in the 4 areas of robotic surgical systems, virtual nursing systems, medical diagnostic systems, and medical imaging systems<sup>[26]</sup>. In particular, AI-based medical imaging analysis performs the most important functions in medical diagnosis and, along with pathology-related data, is a key area of AI-based medical technology<sup>[26]</sup>.



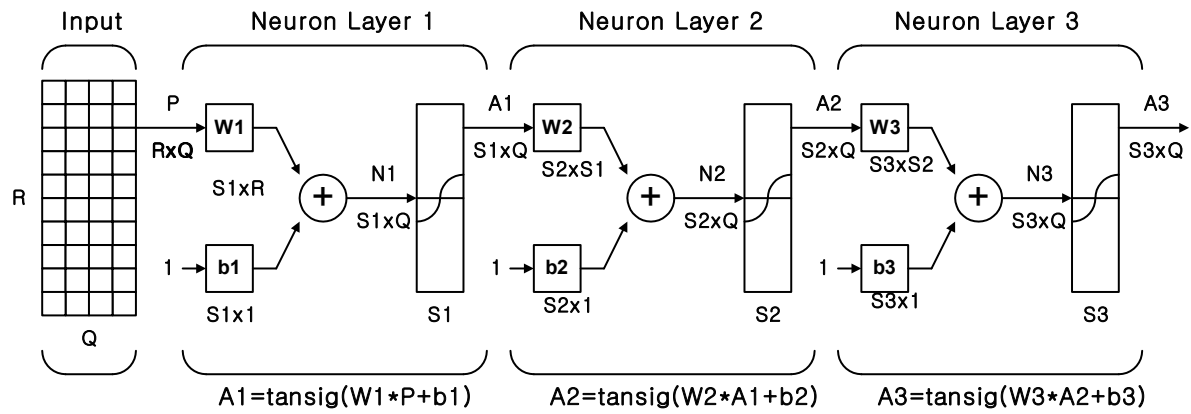


Figure 3. Neural network diagram

### III. Experiment

The experimental images for this study were selected by searching for "mammogram" in the NIH database. The images selected for the experiment were converted into a JPEG format of 256 X 256

pixels and saved. The stored experimental images were segmented, and edge detection was performed. Most of the segments are low frequency, but the edges are of high frequency. DCT was performed to extract the features of the two parts, and similarity was determined based on DCT values input into the neural network. Figure 4 shows the experimental procedure.

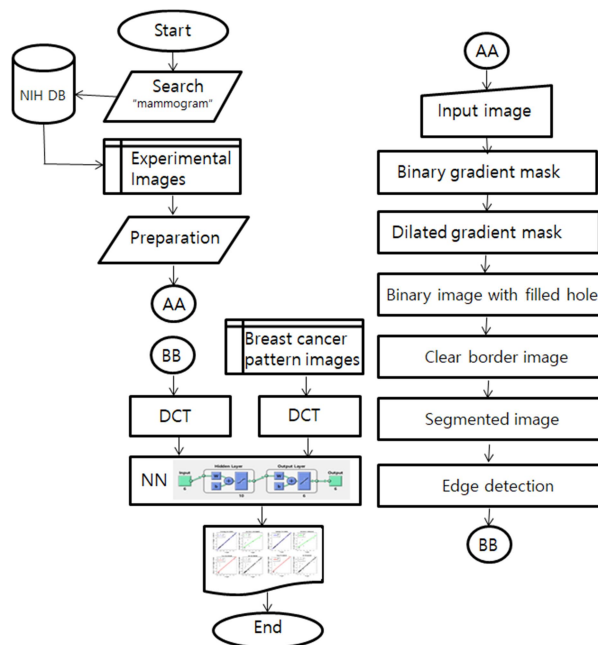
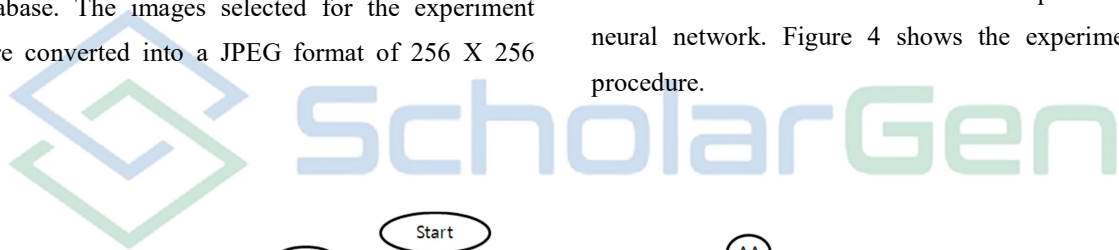


Figure 4. Flowchart of the experimental procedure

### 1. Image segmentation and edge detection

In the breast images for the experiment, the cancer area was segmented using the basic morphology, and the edge was detected. Segmentation is based on the large difference in contrast of a cancer area compared to the surroundings. Contrast change was detected by calculating the image slope. To create a binary mask including the segmented region, a gradient image was calculated, and a threshold value of 127 or more was applied. The edge was detected using Matlab's function `edge()` and `sobel()` operators. Next, the image was expanded. In the binary gradient mask, a line representing the high contrast of the image is displayed. Difficulty arises in recognizing the outline of the desired object with these lines. Compared

with the original image, there is a gap between the lines surrounding the object of the tilt mask. This line spacing disappears when the image is expanded by the `sobel()` operation using a linear structure element. Two orthogonal linear structural elements were created using the `strel()` function. The next step was filling the inner gap. The expanded gradient mask has a well-arranged cell outline, but holes exist inside the cell. To fill these holes, the `imfill()` function was used. The next step was to remove the objects in contact with the border. All objects in contact with the border of the image were removed using the `imclearborder()` function. To remove the diagonal connection, the connectivity of the `imclearborder()` function was set to 4. Finally, the image was eroded twice with a diamond structure element, created using the `strel()` function, for smoothing.

Table 3. Edge detection of breast cancer basic morphology











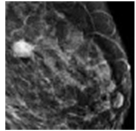
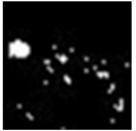

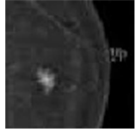


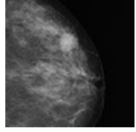


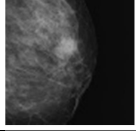


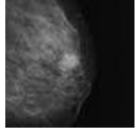


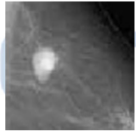


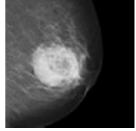


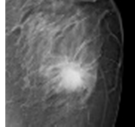


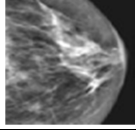

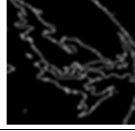
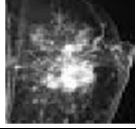

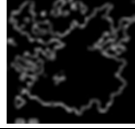
Type	Basic morphology	Edge detection
Round		
Oval		
Lobulated		
Nodular		
Stellate		

Table 4. Segmentation and edge detection of experimental images

	Image	Segment	Edge
Img_1			
Img_2			
Img_3			
Img_4			
Img_5			
Img_6			
Img_7			
Img_8			
Img_9			
Img_10			

**2. Perform DCT**

The features of the experimental images and

breast cancer basic morphological images were extracted from the edge detection image using Matlab's dct() function.

Table 5. Extracted-feature values of basic morphological images of breast cancer

Round	0.363	0.5	0.039	0.01	0.024	-0.02	0.082	0.056	-0.049	-0.12	0.074	0.21	-0.184	-0.1	-0.15	0.006	-0.07	-0.17	-0.15	0.166	-0.19	-0.17	-0.12	-0.18	-0.2	0.072	-0.5	-0.5	-0.5	-0.5	
Oval	0.5	0.227	-0.17	0.08	-0.174	-0.16	-0.18	-0.24	-0.31	-0.1	-0.06	-0.2	-0.25	-0.275	-0.2	-0.03	-0.1	-0.19	-0.26	-0.23	-0.19	-0.19	-0.07	-0.01	-0.26	-0.28	-0.18	-0.5	-0.5	-0.5	-0.5
Lobulated	0.5	0.492	-0.05	0.07	-0.044	0.015	0.052	0.203	-0.047	-0.19	0.022	0.22	-0.03	-0.115	-0.1	-0.15	-0.04	-0.06	-0.2	-0.17	-0.06	-0.22	-0.24	-0.2	-0.25	-0.21	-0.21	-0.5	-0.5	-0.5	-0.5
Nodular	0.5	0.444	0.252	0.07	0.161	0.111	0.053	0.005	0.135	0.179	0.002	-0	-0.02	0.053	0.15	-0.03	-0.08	-0.03	-0.13	0.219	0.229	-0.07	-0.08	-0	-0.03	-0.05	0.24	-0.5	-0.5	-0.5	-0.5
stellate	0.5	0.374	-0.04	0.04	-0.082	0.007	-0.13	-0.1	-0.057	-0.14	-0.08	-0.1	-0.11	-0.099	-0.1	-0.16	-0.14	-0.21	-0.19	-0.12	3E-04	-0.15	-0.17	-0.24	-0.19	-0.13	-0.12	-0.5	-0.5	-0.5	-0.5

Table 6. Extracted-feature values of experimental images

img_1	0.4	0.5	0.313	0.195	0.2216	0.1699	0.0646	0.188	0.116	-0.012	-0.017	-0.025	-0.061	-0.0344	-0.104	-0.1507	-0.092	-0.096	-0.169	-0.127	-0.286	-0.219	-0.2	-0.146	-0.195	-0.1824	-0.24	-0.5	-0.5	-0.5	-0.5
img_2	0.295	0.4237	0.369	0.161	0.1678	0.5	0.082	-0.024	0.203	0.133	0.0636	-0.058	-0.017	-0.0236	0.1134	0.0715	-0.031	-0.137	-0.078	-0.079	-0.064	-0.084	-0.08	-0.181	-0.188	-0.1018	-0.22	-0.5	-0.5	-0.5	-0.5
img_3	0.37	0.5	0.177	0.326	0.2006	0.1807	0.3051	0.156	0.159	0.0019	0.1257	0.0248	0.0127	-0.015	-0.174	-0.041	0.0159	-0.05	-0.096	-0.092	-0.286	-0.117	-0.12	-0.093	-0.097	-0.1648	-0.24	-0.5	-0.5	-0.5	-0.5
img_4	0.5	0.3899	0.234	0.28	0.0284	0.049	0.0617	0.054	-0.1	-0.168	-0.08	-0.076	-0.076	-0.2008	-0.329	-0.2391	-0.18	-0.182	-0.22	-0.302	-0.348	-0.4	-0.33	-0.266	-0.303	-0.3406	-0.43	-0.5	-0.5	-0.5	-0.5
img_5	0.5	0.4383	0.345	0.447	0.348	0.1193	0.2834	0.402	0.173	-0.078	0.1241	0.1311	0.095	-0.0817	-0.274	-0.0811	-0.099	-0.02	-0.029	-0.248	-0.385	-0.304	-0.14	-0.079	-0.135	-0.2804	-0.36	-0.5	-0.5	-0.5	-0.5
img_6	0.343	0.5	0.399	0.291	0.4593	0.2666	0.1895	0.188	0.311	0.0303	0.1298	0.3581	0.1709	0.254	0.027	0.0425	0.1899	0.2976	0.1251	0.233	-0.063	-0.056	0.024	0.1082	0.1463	0.0408	0.146	-0.5	-0.5	-0.5	-0.5
img_7	0.366	0.5	0.18	0.396	0.2299	0.0622	0.1309	0.19	0.054	0.0131	0.0013	0.075	-0.056	-0.0031	-0.038	-0.11	-0.013	-0.012	-0.071	-0.047	-0.118	-0.131	-0.08	-0.062	-0.086	-0.1098	-0.08	-0.5	-0.5	-0.5	-0.5
img_8	0.378	0.5	0.18	0.344	0.1778	0.0571	0.2143	0.309	0.078	0.0381	0.1736	0.2377	0.0654	-0.0424	-0.065	0.0495	0.1468	0.1256	0.0127	-3E-04	-0.036	-0.021	0.129	0.0666	0.0043	0.0218	-0.06	-0.5	-0.5	-0.5	-0.5
img_9	0.41	0.5	0.222	0.348	0.1995	0.1416	0.2325	0.251	0.191	0.0503	0.1228	0.2107	0.1475	0.1472	-0.077	0.0229	0.1317	0.0759	0.1361	0.0646	-0.052	-0.103	-0.03	-0.012	0.014	0.0743	0.039	-0.5	-0.5	-0.5	-0.5
img_10	0.423	0.5	0.097	0.321	0.3954	0.0838	0.0984	0.191	0.262	0.3112	0.0093	0.022	0.1533	0.1659	0.0838	0.0116	0.0237	0.0037	0.1661	0.4413	0.0849	0.0204	-0.06	-0.043	0.0116	0.1062	0.304	-0.5	-0.5	-0.5	-0.5

### 3. Perform similarity determination

Experimental image data of *Img\_1* to *Img\_10* were input into the neural network to determine which type of breast cancer basic morphology was similar to the experimental image. Neural networks were trained by inputting breast cancer basic morphology data of “Round, Oval, Lobulated, Nodular, and Stellate” into each set

of experimental image data.

Input examination image ‘*Img\_1*’ is a 1 x 32 matrix, representing static data of 32 samples of 1 element. Target breast cancer basic morphology ‘Round’ is a 1 x 32 matrix, representing static data of 32 samples 1 element. The experiment was conducted as below. Figures 5 to 8 show performance, training state, error histogram, and fit of the method.



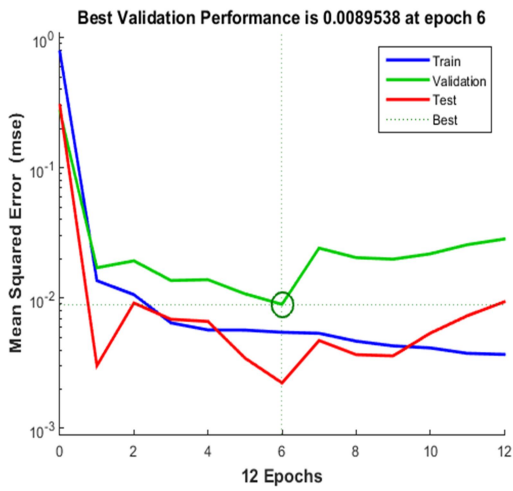


Figure 5. Performance

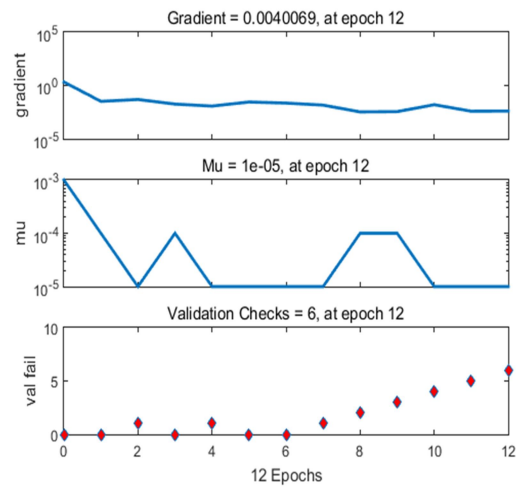


Figure 6. Training state



Figure 7. Error histogram

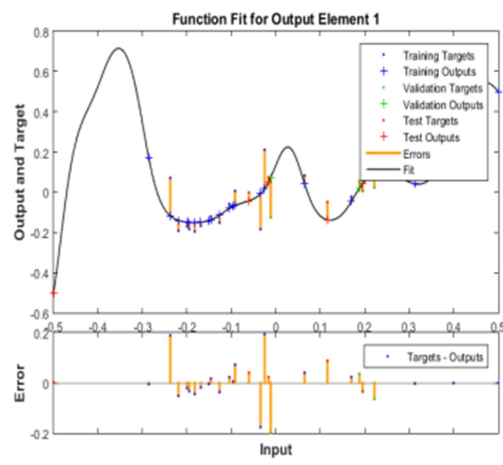


Figure 8. Fit

Figure 9 shows regression graphs of experimental image data when 'Img\_1' was the neural network input and breast cancer basic

morphology 'Round' was the neural network target. The figure shows the degree of regression in R-value.

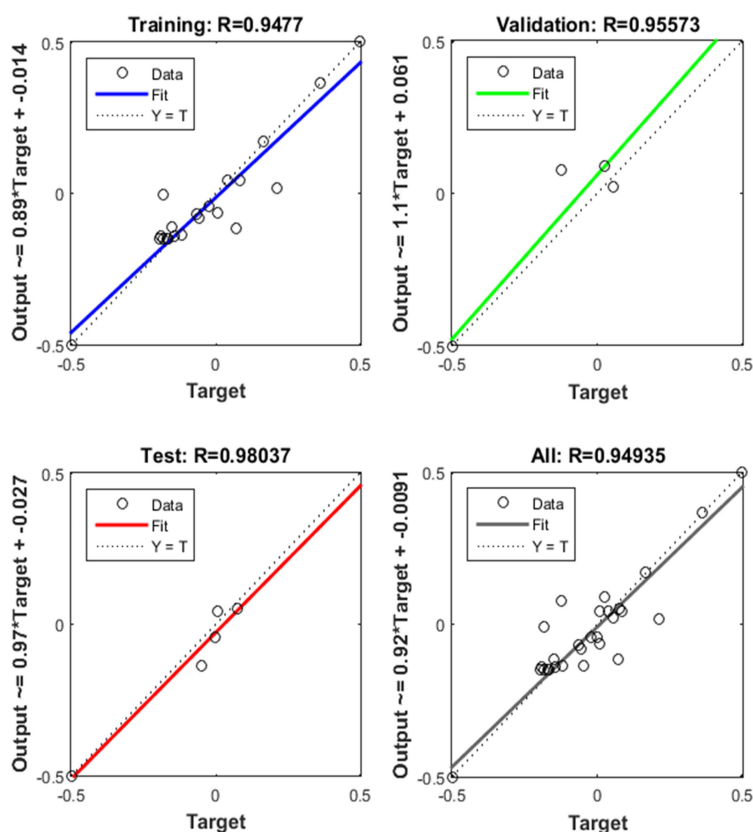


Figure 9. Regression

#### IV. Results and discussion

To determine the degree of malignancy in breast imaging, the similarity to the previously studied breast cancer basic morphology was examined. The results are shown in Table 7.

1. Img\_1 showed the highest R value of 0.9874, which indicates similarity to Stellate.
2. Img\_2 showed the highest R value of 0.9246, indicating similarity to Nodular.
3. Img\_3 showed the highest R value of 0.978183, which indicates similarity to Stellate.
4. Img\_4 showed the highest R value of 0.96967, indicating similarity to Stellate.
5. Img\_4 showed the highest R value of 0.94653,

indicating similarity to Stellate.

6. Img\_4 showed the highest R value of 0.92613, indicating similarity to Round.

7. Img\_4 showed the highest R value of 0.90947, indicating similarity to Nodular.
















8. Img\_4 showed the highest R value of 0.97085, indicating similarity to Stellate.

9. Img\_4 showed the highest R value of 0.9343, indicating similarity to Round.

10. Img\_4 showed the highest R value of 0.97475, indicating similarity to Stellate.

Images similar to malignant cancer were Img\_1, Img\_3, Img\_4, Img\_5, Img\_8, and Img\_10. The images similar to the nodular shape were Img\_2 and Img\_7. The image similar to benign was Img\_6.

Table 7. Results of examination

Basic morphology Exam_img		Benign----->malignant				
		Round	Oval	Lobulated	Nodular	Stellate
Image	edge					
Img_1		0.87409	0.90754	0.88755	0.9119	0.9874
Img_2		0.80258	0.028656	0.77009	0.9246	0.90448
Img_3		0.93114	0.86929	0.90165	0.88674	0.978183
Img_4		0.82084	0.90451	0.95427	0.86761	0.96967
Img_5		0.83403	0.93647	0.93425	0.89497	0.94653
Img_6		0.92613	0.35626	0.92289	0.84943	0.85255
Img_7		0.79847	0.75773	0.89233	0.90947	0.85966
Img_8		0.89469	0.85434	0.91494	0.92007	0.97035
Img_9		0.9343	0.93303	0.84103	0.85129	0.91545
Img_10		0.87204	0.68936	0.82704	0.94892	0.97475

### V. Conclusion

In this paper, a method to determine breast cancer basic morphology associated with image shapes is presented. For similarity discrimination, the experimental breast cancer image was segmented based on the cancer region, and the edge of the segmented cancer region was detected. Features were extracted from the detected edge image of the experimental image by DCT. The breast cancer basic morphology image was also edge detected, and features were extracted by DCT. AI learning for similarity determination consisted of the input layer, hidden layer, and output layer; and iterative learning occurred through the error backpropagation method. The neural network was trained by sequentially

inputting the feature values of the experimental image and the target data of the neural network as the breast cancer basic morphology images (Round, Oval, Lobulated, Nodular, and Stellate).

The neural network training results of 10 experimental images are as follows.

- 1) There were 6 images representing malignant tumors.
- 2) There were 2 images showing benign tumors.
- 3) There were two images that could worsen into malignant tumors.

The experimental images selected for the experiment were from the breast cancer image database. Medical images are interpreted by a

radiologist in consideration of pathological factors other than imaging. Since the discrimination of medical images by AI to date is limited to image information, a radiologist's input is necessary. To improve the discrimination ability of medical images by AI, extracting accurate features of medical images and inputting clinical information are necessary in addition to accurately setting targets. The study of learning algorithms for neural networking should be continued. We believe that this study concerning recognition of cancer on digital breast images by AI deep learning will be useful to the radiomics (radiology and genomics) research field.

### Competing interests

The authors declare that there is no conflict of interest regarding the publication of this paper

### [Acknowledgment]

This study was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (NRF-2020R1I1A1A01060851), and in part by research funds from Nambu University, 2020.

### [References]

- [1] Simone Natale, "**THE INVISIBLE MADE VISIBLE**", Journal of Media History, Volume 17, Issue 4, pp. 345~358(2011).
- [2] Stephen T. Thornton, Andrew Rex, "**Modern Physics for Scientists and Engineers**", Publisher, Physics and Astronomy: Charles Hartford, PP. 272~298(2013)
- [3] Bruce H. Hasegawa, Koji Iwata et al, "**Dual-Modality Imaging of Function and Physiology**",

Academic Radiology, Volume 9, Issue 11, pp. 1305~1321(2002)

- [4] Klaus Schwab, "**The Fourth Industrial Revolution**", Crown Business, pp. 6~21(2016)
- [5] Chayakrit Krittanawong, HongJu Zhang, et al, "**Artificial Intelligence in Precision Cardiovascular Medicine**", JACC, Volume 69, Issue 21, pp. 2655~2664(2017)
- [6] Christiane Katharina Kuhl, Rita K. Schmutzler, et al, "**Breast MR Imaging Screening in 192 Women Proved or Suspected to Be Carriers of a Breast Cancer Susceptibility Gene: Preliminary Results**", Radiology, Vol. 215, No. 1, pp. 267~279 (2000)
- [7] Clarisse Dromain & Fabienne Thibault et al, "**Dual-energy contrast-enhanced digital mammography: initial clinical results**", Eur Radiol., Volume 21, pp. 565~574(2011)
- [8] Azhar Taha Adam Mohamed, "**Evaluation of Radiation Dose to Patients During Mammography Examination**", Sudan Academy of Science(SAS) Atomic Energy Console(2018)
- [9] Virginia Gonzalez, Kerstin Sandelin, et al, "**Preoperative MRI of the Breast (POMB) Influences Primary Treatment in Breast Cancer: A Prospective, Randomized, Multicenter Study**", World Journal of Surgery, Volume 38, pp. 1685~1693(2014)
- [10] Mark A. Helvie, "**Digital Mammography Imaging: Breast Tomosynthesis and Advanced Applications**", Radiol Clin North Am., Vol. 48, No. 5, pp. 917~929(2010)
- [11] Deepa Sankar, Tessamma Thomas, "**Fractal Features based on Differential Box Counting**

**Method for the Categorization of Digital Mammograms**", IJCISIM, Vol.2, pp. 11~19(2010)

[12] Weijie Chen, Maryellen L. Giger, et al. "**Computerized interpretation of breast MRI: Investigation of enhancement variance dynamics**". Med. Phys., Vol. 31, No. 5, pp. 1076~1082(2004)

[13] H. Li, K. J. R. Liu, S-C.B. Lo, "**Fractal Modeling and Segmentation for the Enhancement of Microcalcification in Digital Mammograms**", ISR Technical Research Report, pp 1~33 (2007)

[14] Arnau Oliver, Xavier Llad'o, et al. "**False Positive Reduction in Mammographic Mass Detection Using Local Binary Patterns**", MICCAI 2007: Medical Image Computing and Computer-Assisted Intervention – MICCAI, pp. 286-293(2007)

[15] L.M. Bruce ; R.R. Adhami, "**Classifying mammographic mass shapes using the wavelet transform modulus-maxima method**", IEEE Transactions on Medical Imaging, Volume 18, Issue 12, pp. 1170~1177(1999)

[16] Dzung L, Phamy, Chenyang Xu , Jerry L. Prince, "**Current Methods in Medical Image Segmentation**", Annual Review of Biomedical Engineering, Volume 2, pp. 315~337(2000)

[17] K. Bhargavi, "**A Survey on Threshold Based Segmentation Technique in Image Processing**", IJIRD, Vol. 3, Issue 12, pp. 234~239(2014)

[18] Ali Can, Hong Shen, J.N. Turner, H.L. Tanenbaum, B. Roysam, "**Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms**", IEEE Transactions on Information Technology in Biomedicine, Volume 3, Issue 2, pp. 125~138(1999)

[19] Wenshuo Gao, Xiaoguang Zhang, Lei Yang, Huizhong Liu, "**An improved Sobel edge detection**", 2010 3rd International Conference on Computer Science and Information Technology, INSPEC Accession Number: 11520501

[20] Tony Lindeberg, "**Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention**", International Journal of Computer Vision volume 11, pp. 283~318(1993)

[21] Tokumichi Murakami, "**The Development and Standardization of Ultra High Definition Video Technology**", Springer Link Part of the Signals and Communication Technology book series, High-Quality Visual Experience, pp. 81-135(2010)

[22] Bertran Denis, Jean Cote, Rene Laprise, "**Spectral Decomposition of Two-Dimensional Atmospheric Fields on Limited-Area Domains Using the Discrete Cosine Transform (DCT)**", American Meteorological Society, Volume 130, Issue 7, pp. 1812~1829(2002)

[23] Barbara Penna, Tammam Tillo, Enrico Magli, Gabriella Olmo, "**Transform Coding Techniques for Lossy Hyperspectral Data Compression**", IEEE Transactions on Geoscience and Remote Sensing, Volume 45, Issue 5, pp. 1408~1421(2007)

[24] Dennis M. Dimiduk, Elizabeth A. Holm, Stephen R. Niezgod, "**Perspectives on the Impact of Machine Learning, Deep Learning, and Artificial Intelligence on Materials, Processes, and Structures Engineering**", Springer Integrating Materials and Manufacturing Innovation, Volume 7, pp. 157~172(2018)

[25] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, et al. "*A Survey on Deep Learning in Medical Image Analysis*", *Medical Image Analysis*, Volume 42, pp. 60~88(2017)

[26] Werner Horn, "*AI in medicine on its way from knowledge-intensive to data-intensive systems*", *Artificial Intelligence in Medicine*, Volume 23, Issue 1, pp. 5~12(2001)

