

Article

Comprehensive Serum Glycopeptide Spectra Analysis Combined with Artificial Intelligence (CSGSA-AI) to Diagnose Early-Stage Ovarian Cancer

Kazuhiro Tanabe ^{1,2,*}, Masae Ikeda ³, Masaru Hayashi ³, Koji Matsuo ^{4,5}, Miwa Yasaka ³, Hiroko Machida ³, Masako Shida ³, Tomoko Katahira ¹, Tadashi Imanishi ⁶, Takeshi Hirasawa ³, Kenji Sato ³, Hiroshi Yoshida ³ and Mikio Mikami ^{3,*}

¹ Medical Solution Promotion Department, Medical Solution Segment, LSI Medience Corporation, Tokyo 1748555, Japan; katahira.tomoko@ma.medience.co.jp

² Research Supporting Department, Kyushu Pro Search Limited Liability Partnership, Fukuoka 8190388, Japan

³ Department of Obstetrics and Gynecology, Tokai University School of Medicine, Kanagawa 2591193, Japan; ikedam@tokai-u.jp (M.I.); masaru.hayashi@tokai.ac.jp (M.H.); yasakamiwa@tokai.ac.jp (M.Y.); hiroko.machida@tokai.ac.jp (H.M.); shida@is.icc.u-tokai.ac.jp (M.S.); hira@is.icc.u-tokai.ac.jp (T.H.); igymichael@gmail.com (K.S.); hiroshiyoshidamd@gmail.com (H.Y.)

⁴ Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, University of Southern California, Los Angeles, CA 90033, USA; koji.matsuo@med.usc.edu

⁵ Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA 90033, USA

⁶ Department of Molecular Life Science, Division of Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, Kanagawa 2591193, Japan; imanishi@tokai-u.jp

* Correspondence: tanabe.kazuhiro@mp.medience.co.jp (K.T.); mmikami@is.icc.u-tokai.ac.jp (M.M.); Tel.: +81-70-5597-7284 (K.T.); +81-463-93-1121 (M.M.)

† These authors contributed equally.

Received: 21 July 2020; Accepted: 18 August 2020; Published: 21 August 2020



Abstract: Ovarian cancer is a leading cause of deaths among gynecological cancers, and a method to detect early-stage epithelial ovarian cancer (EOC) is urgently needed. We aimed to develop an artificial intelligence (AI)-based comprehensive serum glycopeptide spectra analysis (CSGSA-AI) method in combination with convolutional neural network (CNN) to detect aberrant glycans in serum samples of patients with EOC. We converted serum glycopeptide expression patterns into two-dimensional (2D) barcodes to let CNN learn and distinguish between EOC and non-EOC. CNN was trained using 60% samples and validated using 40% samples. We observed that principal component analysis-based alignment of glycopeptides to generate 2D barcodes significantly increased the diagnostic accuracy (88%) of the method. When CNN was trained with 2D barcodes colored on the basis of serum levels of CA125 and HE4, a diagnostic accuracy of 95% was achieved. We believe that this simple and low-cost method will increase the detection of EOC.

Keywords: AI; deep learning; ovarian cancer; screening; early stage; convolutional neural network; glycopeptide; mass spectrometry; glycomics; AlexNet

1. Introduction

Ovarian cancer is one of the most common malignancies worldwide. Although the 5-year average annual death rate due to ovarian cancer decreased by 2.3% during 2011–2015 [1], epithelial ovarian cancer (EOC) remains the leading cause of death among gynecological cancers [2,3]. According to the National Cancer Institute [4], among Caucasians, the 5-year survival rate has improved from 37% to 45% from 1975 to 2006, whereas that among African-Americans has worsened from 43% to 37% during the same

period. Diagnostics for early-stage EOC are urgently needed because more than 50% of symptomatic women are diagnosed at an advanced stage and the 5-year survival rate of patients with advanced stage EOC is less than 30% [5]. Imaging technologies, such as positron emission tomography/computed tomography (PET/CT) with fluorodeoxyglucose, are useful in detecting early-stage EOC that is less than 3 cm in size; however, it is not widely employed for screening because of examinee's economical and physical burdens [6,7]. On the contrary, a convenient blood test using tumor markers, such as cancer antigen 125 (CA125), is partially used for screening of EOC [8,9]. However, the sensitivity and specificity of the CA125 test are not sufficient to detect early-stage cancer, and it does not alleviate the mortality rate [10,11]. Recent advances in the detection of early-stage EOC, such as micro RNA or cell-free DNA detection, have shown significant improvement in sensitivity and specificity [12,13]. However, these technologies are only partially successful in detecting early-stage EOC, because they cannot distinguish EOC from benign ovarian tumor, which has a much higher prevalence than that of EOC. To establish a reliable and robust blood test for early-stage EOC detection, we developed a glycopeptide profiling technology using mass spectrometry, which detects aberrant serum sugar chains in patient sera [14]. In our previous studies, we have shown that C4-binding protein with fully sialylated sugar chains is specifically increased in the sera of EOC patients [15–17]. Based on this, we developed a comprehensive serum glycopeptide spectra analysis (CSGSA) technology by combining glycopeptide profiling with orthogonal partial least squares discriminant analysis (OPLS-DA) [18]. This approach has remarkably improved the diagnostic potential of this test to distinguish between EOC and non-EOC.

Recently, artificial intelligence (AI), in particular image recognition using convolutional neural network (CNN), has emerged to distinguish objects in images [19], and now contributes on pedestrian recognition [20], surveillance videos [21] or financial forecasting [22]. CNN has multiple process layers (steps) and optimizes parameters assigned in each layer to achieve correct recognition of target objects using a lot of training samples [23,24]. CNN can not only classify images more accurately than conventional neural networks or machine-learning systems [25] but also transfer the acquired knowledges to other studies to reduce training effort [26]. The “transfer learning” technology has enabled clinical scientists to utilize AI in studies with a limited number of cases [25,27]. Oh et al. [25] used CNN to learn 695 brain images of Alzheimer's disease patients acquired by magnetic resonance imaging (MRI) to distinguish it from normal brain and achieved 87% accuracy. Since preparing a large number of clinical specimens is difficult, this transfer learning technology is meaningful to accelerate clinical studies using AI [28]. Moreover, since the development of CNN is outstanding among general deep learning technologies, some scientists have attempted to convert numerical (non-image) data into images and let CNN learn the images. Sharma et al. [29] have developed a tool named DeepInsight, which converts non-image numerical data, such as RNA-seq, into well-organized images that allow CNN to learn and distinguish phenotypes.

AlexNet is a pretrained CNN architecture developed by Krizhevsky [30], who won the first prize in the ImageNet Large Scale Visual Recognition Challenge (LSRVC) contest in 2012. AlexNet has 25 layers, which include five convolutional layers interleaved with max pooling layers and local response normalization layers. Each layer consists of sets of feature maps, which are the responses of a filter on the output of the preceding layer. In this study, we aimed to use AlexNet, the latest CNN-based technology, as a discrimination tool for CSGSA to identify early-stage EOC. To facilitate CNN training, we converted numerical data of glycopeptide expression to 2D barcode images and let CNN learn and distinguish early-stage EOC.

2. Results

2.1. The Training Options

The training options, namely initial learning rate (speed of convergence), maximum number of epochs (a full pass of the data), and minimum batch size (a size of data subunit), were chosen in such

a way that the calculation is fast and accurate. When the initial learning rate is too low or too high, training takes a long time or does not reach the correct solution. If the number of epochs is too small, calculation stops before converging, and if the minimum batch size is too big (i.e., number of batches is too small), calculating cost through the network becomes expensive, i.e., the efficacy of optimization becomes worse. When the initial learning rate was set at 0.0001, maximum number of epochs at 30, and minimum batch size at 5, the training converged within acceptable period (10 to 20 min) without diverging, and the accuracy and loss curves reached a plateau before calculation stopped, indicating that the maximum number of epochs was adequate (Figure S1). Thus, we performed the following experiments with these values.

Two optimizers—adaptive moment estimation (adam) and stochastic gradient descent with momentum (sgdm)—were evaluated regarding how they correctly distinguish EOC from non-EOC. Each model was trained using 60% of data and evaluated using the remaining 40% of data. This process was repeated 10 times by randomly reselecting training data sets, and ROC-AUCs (area under curve of receiver operating characteristic) were obtained from predicted values of test samples. As a result, there was no significant difference between the two optimizers (adam: 0.878, sgdm: 0.881). Therefore, we used sgdm for the following experiments.

2.2. Learning Efficacy of Pretrained CNN Models

The diagnostic performance of two AlexNet models, non-pretrained and pretrained (transferred), was compared. The purpose of this evaluation was to clarify the effectiveness of pretraining, even if it was not related to actual training (determination of EOC). The non-pretrained AlexNet model was established using AlexNet framework involving fresh 25 layers without any training. The number of weights, bias, slide, and padding were reproduced according to an original AlexNet framework (Table 1 [30]). The transferred AlexNet model was based on a pretrained AlexNet model that was already trained by more than one million images from the ImageNet database [31]. This model was able to categorize 1000 objects, such as keyboard, mouse, pencil, and many animals. Only the last three layers, namely fully connected layer, softmax layer, and classification layer, were replaced with new fresh layers.

Table 1. AlexNet framework.

Number	Layer Name	Role	Description
1	data	Image Input	227 × 227 × 3 images with “zero centering” normalization
2	conv1	Convolution	96 × 11, 11 × 3 convolutions with stride [4 4] and padding [0 0 0 0]
3	relu1	ReLU	ReLU
4	norm1	Cross Channel Normalization	cross channel normalization with 5 channels per element
5	pool1	Max Pooling	3 × 3 max pooling with stride [2 2] and padding [0 0 0 0]
6	conv2	Grouped Convolution	2 groups of 128 (5 × 5 × 48) convolutions with stride [1 1] and padding [2 2 2 2]
7	relu2'	ReLU	ReLU
8	norm2	Cross Channel Normalization	cross channel normalization with 5 channels per element
9	pool2	Max Pooling	3 × 3 max pooling with stride [2 2] and padding [0 0 0 0]
10	conv3	Convolution	384 3 × 3 × 256 convolutions with stride [1 1] and padding [1 1 1 1]
11	relu3	ReLU	ReLU
12	conv4	Grouped Convolution	2 groups of 192 3 × 3 × 192 convolutions with stride [1 1] and padding [1 1 1 1]
13	relu4	ReLU	ReLU
14	conv5	Grouped Convolution	2 groups of 128 3 × 3 × 192 convolutions with stride [1 1] and padding [1 1 1 1]
15	relu5	ReLU	ReLU
16	pool5	Max Pooling	3 × 3 max pooling with stride [2 2] and padding [0 0 0 0]
17	fc6	Fully Connected	4096 fully connected layer
18	relu6	ReLU	ReLU
19	drop6	Dropout	50% dropout
20	fc7	Fully Connected	4096 fully connected layer
21	relu7	ReLU	ReLU
22	drop7	Dropout	50% dropout
23	fc8	Fully Connected	1000 fully connected layer
24	prob	Softmax	softmax
25	output	Classification Output	crossentropyex with “EOC” and “Non-EOC”

Rectified Linear Unit.

The layers used in the two models are illustrated in Figure 1a. The gray layers indicate that they were already pretrained and acquired image features, and the parameters are not changed during additional training. On the contrary, the white layers indicate that their parameters were not pretrained

and were optimized during training. To evaluate the diagnostic performance of the two models, we calculated ROC-AUCs using the repeated holdout validation method as described next. First, the two models were trained by 60% of the 351 samples and evaluated by the remaining 40% samples, and this process was repeated 10 times by randomly reselecting 60% training samples. ROC-AUCs were calculated by predicted values obtained from the test samples. Wilcoxon rank sum test was deployed to assess the statistical difference in AUC values as the two groups were independent. As a result, the average ROC-AUC of the transferred AlexNet model reached 0.881, whereas that of the non-pretrained AlexNet model was 0.853 (Figure 1b). Further, the p -value of the Wilcoxon rank sum test between the two models was 0.020 (Figure 1c). This suggests that even an irrelevant experience, for example, recognizing keyboard, mouse, and pencil, increases learning efficacy for recognizing 2D barcodes of EOC glycopeptide patterns. Additionally, it indicates that the latent features of images are partially common among real objects and artificially generated 2D barcodes.

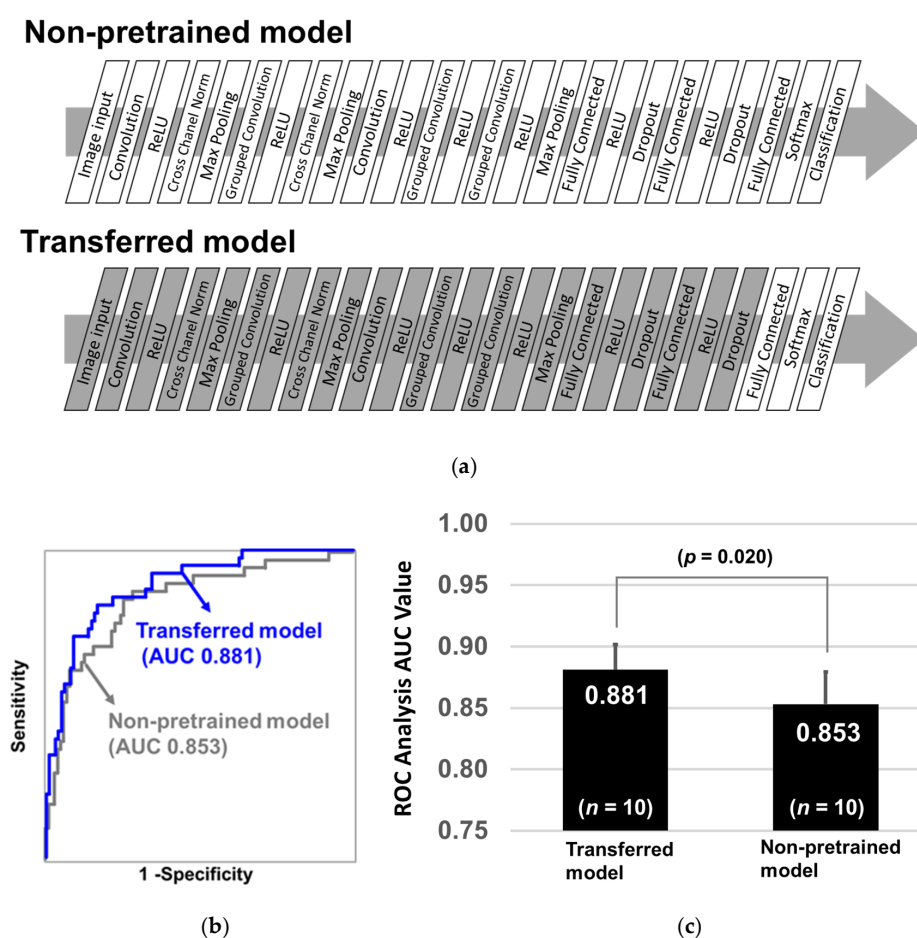
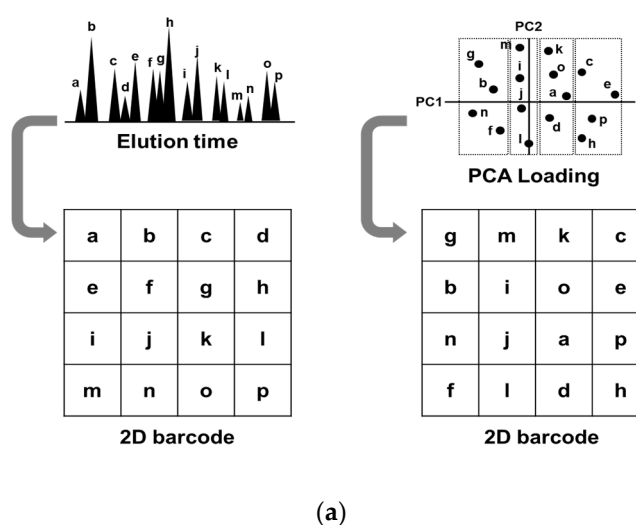


Figure 1. Diagnostic performance of non-pretrained and transferred AlexNet models. (a) Diagnostic performance (AUC (area under curve) of ROC (receiver operating characteristic)) of the non-pretrained and transferred AlexNet convolutional neural network (CNN) models was compared. The non-pretrained model was established by an original AlexNet framework with 25 fresh layers whose parameters were not optimized (Non-pretrained model). The composition of the transferred AlexNet model was the same as that of the non-pretrained model, but the parameters were already trained with the help of more than one million images from the ImageNet database; further, only last three layers were initialized (Transferred model). (b) White layer: non-trained, Gray layer: pretrained. After the two models were trained with 60% of the 351 samples, they were evaluated by the remaining 40% samples. ROC curves were shown with blue line (transferred) and gray line (non-pretrained). (c) This process was repeated 10 times and the p -values of the Wilcoxon rank sum test were calculated between two groups.

2.3. Glycopeptide Alignment to Generate 2D Barcodes

To understand how glycopeptide alignment to generate a 2D barcode affects diagnostic performance, we prepared two sets of 2D barcodes and compared their diagnostic performances. One set was aligned by the order of liquid chromatography elution time (Rt-based barcode), and the other was aligned by the order of principal component analysis (PCA) loading values (PCA-based barcode, Figure 2a). Considering that CNN recognizes spatial relative distances among object parts, for example, CNN recognizes a human face by identifying relative distances between mouth, nose, and eyes, CNN would recognize EOC features more precisely if the glycopeptides with similar expression pattern are closely located. Since PCA classifies glycopeptides based on their expression similarity, we used PCA to determine the positions of glycopeptides in a 2D barcode. PCA-based barcode was generated as follows: first, PCA of 1712 glycopeptides from 351 samples was performed, and the loading values of the first and second components for 1712 glycopeptides were obtained. Then, 1712 glycopeptides were sorted based on the first loading values, followed by dividing them into 41 groups. Further, they were sorted based on the second loading values within each group. The X coordinate in a 2D barcode indicated the number of groups and the Y coordinate indicated the order in each group. We used Rt-based barcode as a control for PCA-based barcode, assuming that Rt-based barcode is estimated as random glycopeptide alignment, because there is no relation between elution order and expression similarity. Rt-based barcode was generated by aligning glycopeptides from left-top to right-down based on the order of elution time. The method of how to change black and white contrast is shown in Table S1. Since a 41×42 matrix has 1722 cells, and 10 blanks are generated on the bottom right corner with the expression values of 1712 glycopeptides, the blank cells were replaced with white. The typical 2D barcodes of EOC and non-EOC generated by Rt-based alignment and PCA-based alignment are shown in Figure 2b. It can be observed that black dots (strong expression) are more converged in the PCA-based barcode than in the Rt-based barcode. Training and testing were performed using the transferred AlexNet model, and the diagnostic performance was evaluated with the repeated holdout validation method. As a result, CNN was able to recognize early-stage EOC more precisely using PCA-based barcodes (AUC: 0.881) than using Rt-based barcodes (AUC: 0.852) (Figure 2c). The p -value of the Wilcoxon rank sum test between the two models was 0.024 (Figure 2d). Overall, this exercise indicated that CNN could elicit EOC features more precisely when it is trained by PCA-based barcodes.



(a)
Figure 2. Cont.

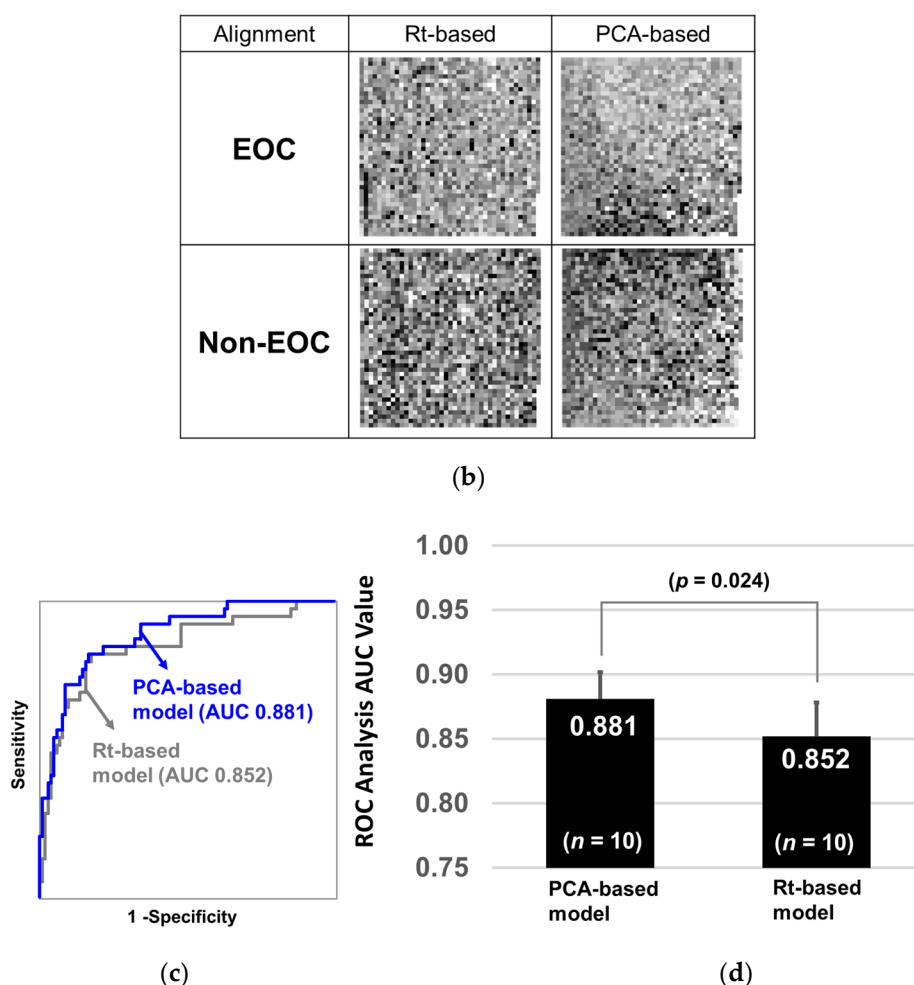


Figure 2. Alignment of glycopeptides to generate 2D barcode. (a) The Rt-based barcode was generated by aligning the glycopeptides from left-top to right-down by the order of elution time. The PCA-based barcode was developed by aligning the glycopeptides by the order of PCA loading values. (b) Typical Rt-based and PCA-based 2D-barcode images of the samples from patients with epithelial ovarian cancer (EOC) and of those with non-EOC. (c) After the two models were trained with 60% of the 351 barcodes, they were evaluated by the remaining 40% samples. The ROC curves are shown with a blue line (PCA-based) or with a gray line (Rt-based). (d) This process was repeated 10 times and the p -values of the Wilcoxon rank sum test were calculated between the two groups.

2.4. Training with 2D Barcodes Colored by CA125 and HE4

To further enhance the learning efficacy and diagnostic performance of CNN, we added CA125 and HE4 information into the 2D barcode by changing the color (multicolored model). When the level of CA125 in the serum increases, the red level of the barcode decreases, and when the serum level of HE4 increases, the green level of the barcode decreases (Figure 3a). Table S2 shows how to color the barcode. Ten blanks generated on the bottom right corner were filled with the colors defined by both CA125 and HE4. The barcodes before and after coloring on the basis of CA125 and HE4 levels are shown in Figure 3b,c. It can be observed that the 2D barcodes of EOC tend to be blue and those of non-EOC tend to light purple depending on CA125 and HE4 levels. Training and testing were performed using the transferred and PCA-based AlexNet model and the diagnostic performance was evaluated with the repeated holdout validation method. As a result, CNN classified the multicolored 2D barcodes by AUC of 0.954 (Figure 3d), and the p -value of the Wilcoxon rank sum test was less than 10^{-3} between multicolored and non-multicolored models (without CA125 and HE4 information,

Figure 3e). Since ROC-AUCs of CA125 and HE4 were 0.86 and 0.87, respectively, the combination successfully increased the diagnostic performance of CNN.

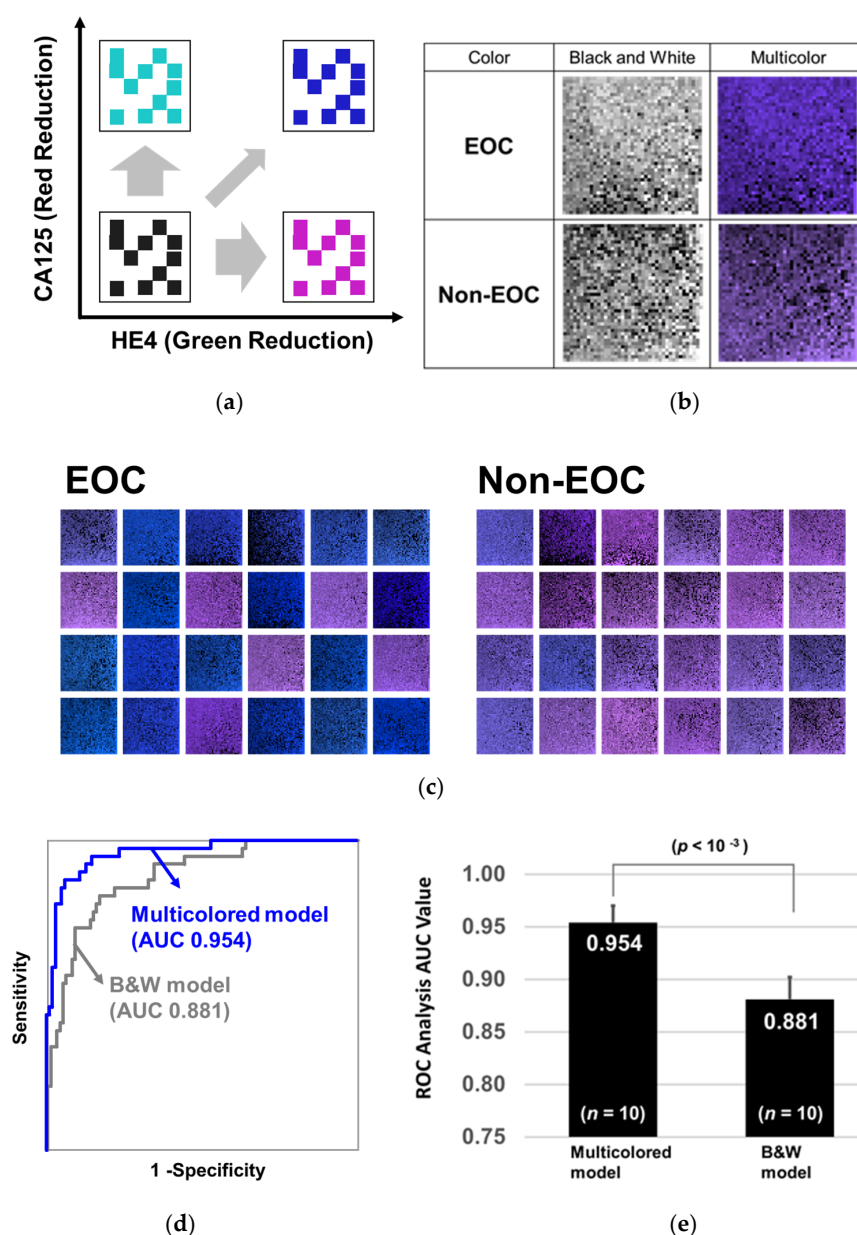


Figure 3. Coloring barcodes on the basis of serum levels of CA125 and HE4. (a) The non-multicolored barcodes were colored on the basis of the serum levels of CA125 and HE4. When the level of CA125 increases, the level of red in the barcode decreases, and when the level of HE4 increases, the level of green decreases. (b) Comparison of patients with EOC and those with non-EOC before and after the coloring of 2D-barcode images. (c) Typical 2D-barcode images of patients with EOC (left) and of those with non-EOC (right). (d) After the two models were trained with 60% of the 351 barcodes, they were evaluated by the residual 40% samples. The ROC curves are shown with a blue line (Multicolor) or with a gray line (Black and white). (e) This process was repeated 10 times and the p -values of the Wilcoxon rank sum test were calculated between the two groups.

2.5. CSGSA-AI Scheme and its Diagnostic Performance

The scheme of CSGSA combined with AI (CSGSA-AI) is shown in Figure 4. Serum glycoproteins obtained from examinees were digested into peptides or glycopeptides using a protease (trypsin).

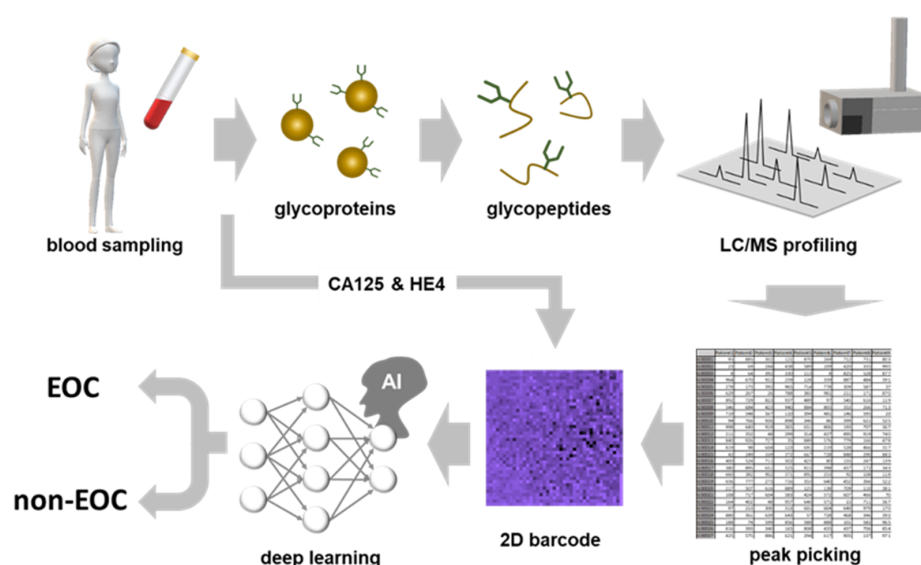


Figure 4. The whole scheme of artificial intelligence-based comprehensive serum glycopeptide spectra analysis (CSGSA-AI). The glycopeptides obtained by trypsin digestion of serum proteins were enriched and analyzed by the QTOF-type mass spectrometer. The expression levels of target glycopeptides were converted into 2D barcodes, and they were colored on the basis of serum CA125 and HE4 levels. After the CNN was trained by the supervised cases, it distinguished the barcodes generated from samples being screened into EOC or non-EOC.

After glycopeptides were enriched by ultrafiltration, they were analyzed using liquid chromatography quadruple time-of-flight mass spectrometry (LC-QTOF-MS) to obtain the glycopeptide expression pattern. This pattern was converted into a 2D barcode (Table S2). When the 2D barcode was generated, glycopeptides were aligned by the order of PCA loadings, and the color was determined based on the levels of CA125 and HE4. The CNN model, which was previously trained by supervised samples and acquired EOC discrimination ability, could now determine unknown 2D barcodes generated from asymptomatic participants and distinguished between EOC or non-EOC. It took two days to complete diagnosis, and 50 samples could be analyzed within one day with one mass spectrometer. The diagnostic performance of CSGSA-AI is shown in Table 2. The sensitivity, specificity, positive predictive value, and negative predictive value were 79%, 96%, 89%, and 92%, respectively. Considering that the EOC group consisted of only stage I patients, this performance is acceptable for the screening of early-stage EOC.

Table 2. Sensitivity and specificity of CSGSA-AI.

CSGSA-AI (Test) (Cut Off = 0.5)						
		Condition		Total	PPV and NPV	
		EOC Stage1	Non		PPV	NPV
CSGSA-AI	Pos	31	4	35	PPV	89%
	Neg	8	98	106	NPV	92%
Total		39	102	141		
Sensitivity and Specificity		Sens	Spec	Accuracy		
		79%	96%	91%		

3. Discussion

In our previous study, we developed a tool that combined CSGSA with OPLS-DA to provide accurate recognition of early-stage EOC. OPLS-DA is based on a liner algebraic, which features robust

and less over-fitting; however, distinguishing complicated cases remains challenging due to its limited separation ability. On the contrary, deep learning powered by CNN is based on a non-linear approach that includes several parameters; therefore, it is advantageous for eliciting features to distinguish specific categories. The biggest challenge to introduce CNN into EOC recognition is to collect a large number of samples to let CNN learn. The number of samples is generally acknowledged as 1000 and more. It is not an easy task to accomplish, because the number of EOC cases is limited compared to those of other cancers. To circumvent this problem, we used a pretrained CNN model, which was already trained with more than 1 million images of 1000 objects, such as strawberry, blueberry, orange, lemon, cosmos, or dahlia, provided by the ImageNet LSVRC contest [31]. Since the underlying features of these objects are partially common, they can be utilized for classifying other kinds of images, such as mass spectrometry data patterns. In this study, we showed that the pretrained CNN model AlexNet can recognize early-stage EOC using PCA-based 2D barcodes at the ROC of 0.881. The ROC further increased to 0.951 when 2D barcodes were generated based on the serum levels of CA125 and HE4. We emphasize that the EOC samples used in this study were all from patients with early-stage (stage I) cancer, whose 5-year survival rates is over 90% [5]. As this image-based diagnostic technique puts much less economical and physical burden on the patients than the classical blood test, this method will contribute to increase in the number of medical checks and decrease in the number of patients who develop advanced stage cancer. Nonetheless, to progress this study further, we have to overcome some issues. First, the AUCs of ROC of training set were almost 100%, indicating that the training has mostly caused overfitting; in other words, the amount of supervised data was not sufficient. Therefore, CNN does not completely elicit the underlying features. Second, the AUC of ROC (0.951) was acceptable but not enough to introduce this technology into actual screening of EOC. Since the morbidity of EOC is 10 to 20 per 100,000 population, the positive predicted value (PPV) becomes less than 1% if the sensitivity of the diagnostics is more than 50%. To improve the PPV, other architectures, such as Resnet50 (<https://mathworks.com/help/deeplearning/ref/resnet50.html>), VGG-16 (<https://mathworks.com/help/deeplearning/ref/vgg16.html>), VGG-19 (<https://mathworks.com/help/deeplearning/ref/vgg19.html>), and Googlenet (<https://mathworks.com/help/deeplearning/ref/googlenet.html>), can be used instead of AlexNet. These architectures can elicit specific features more precisely than AlexNet; however, the number of parameters that should be optimized is much higher for these architectures than that for AlexNet. Therefore, more samples are required to use these architectures, even if pretrained (transferred) models are used.

Deep learning technology is further evolving, and many new technologies have recently emerged. One of the most promising technology is generative adversarial networks [32], which aims to generate new images that have never existed, but everyone can recognize them as real. Generative adversarial networks consist of two functions: “generator” and “discriminator”. Generator randomly generates new images while evolving itself (optimizing its parameters) to create new images that discriminator misclassifies as real. Contrarily, discriminator also evolves itself by optimizing its parameters to distinguish real from artificial images that generator creates. When this process is repeated many times, the generator finally creates new images that everyone recognizes to be real. The fundamental difficulty in clinical studies is to obtain sufficient number of samples; therefore, this technology will overcome the problem to create the required number of samples and lead to more robust and more reproducible analysis. Marouf et al. [33] proposed the conditional single-cell generative adversarial neural network (cscGAN) for the realistic generation of single-cell RNA-seq data, and showed that cscGAN outperforms existing methods for single-cell RNA-seq data generation in quality and holds great promise for the augmentation of other biomedical data types. When this technology is applied to CSGSA-AI, Generative adversarial networks will create new 2D barcodes of EOC or non-EOC, which will supplement the limited number of cases, resulting in increasing the accuracy of diagnosis. Moreover, if the generator plays the role of a “separator” of unsupervised 2D-barcodes obtained from examinees, CSGSA-AI will increase their accuracy of diagnosis while the tests are being conducted.

This indicates that the more the CSGSA-AI conducts tests, the more the diagnostic accuracy increases, which is just what we expect from AI.

Some issues we should address remain before we realize this technology as a clinical test. It is known that the levels of some glycoproteins in serum are influenced by diet, sex, or circadian activities; therefore, the stability of the target glycopeptides should be validated.

4. Materials and Methods

4.1. Patient Samples

A total of 97 serum samples were collected from patients with early-stage EOC at the time of ovarian mass detection prior to the initiation of any treatment (stage I). The EOC diagnosis and stage classification were performed by FIGO guidelines [34]. The average age of these patients was 54.4 (SD \pm 12.9) years. The non-EOC control group ($n = 254$), whose average age was 53.8 (SD \pm 11.4) years, comprised both healthy women ($n = 220$) and women with gynecologic diseases other than ovarian cancer ($n = 34$). Among the 34 patients with gynecologic disease, 20 had uterine fibroid and 14 had ovarian cysts (Table 3). The sera of patients with EOC and those with other gynecologic diseases were obtained from Tokai University Hospital (Kanagawa, Japan). Those of healthy women were obtained from SOIKEN (Osaka, Japan), KAC (Kyoto, Japan), and Sanfco (Tokyo, Japan) with informed consent from each patient or volunteer. Study specific-exclusion criteria are shown in Table S3. All blood samples were collected by venous puncture before surgery or any treatment. The sera were separated from the blood by centrifugation and stored at -80 °C until further examination. Quality control (QC) sample was prepared with pooling approximately 10 EOC and 10 non-EOC sera (total 5 mL), which was dispensed into microtubes 20 μ L each and stored at -80 °C. Two QC samples were analyzed with 20 patient samples to standardize the glycopeptide expression levels of each patient.

Table 3. Study design.

	Class	Cases	Age
EOC Stage I ($n = 97$)	Clear cell carcinoma	41	54.4 \pm 12.9
	Mucinous carcinoma	14	
	Endometrioid carcinoma	28	
	Serous adenocarcinoma	13	
	Unclassified	1	
Non-EOC ($n = 254$)	Healthy	220	53.8 \pm 11.4
	Uterine fibroid	20	
	Ovarian cyst	14	

4.2. Study Approval

Institutional Review Board (IRB) of Tokai University approved the use of patient clinical information and serum/tumor samples (IRB registration number: 09R-082, approved on 17 September 2009).

4.3. Sample Preparation

Ten microliters of fetal calf fetuin (Sigma, St. Louis, MO, USA; 2 mg/mL) aqueous solution was added to the serum sample (20 μ L) to check the efficiency of trypsin digestion or the recovery of glycopeptides. Then, acetone containing 10% of trichloroacetic acid (120 μ L, Wako Pure Chemical Industries, Ltd., Osaka, Japan) was added to remove serum albumin. After mixing and centrifuging at 12,000 rpm for 5 min, the supernatant was removed and cooled acetone (400 μ L) was added to wash the precipitate. The mixture was centrifuged at 10,000 \times g for 5 min and the supernatant was removed again. The precipitate was mixed with a denaturing solution that was prepared by mixing

urea (80 µg, Wako Pure Chemical Industries, Ltd., Osaka, Japan), Tris-HCl buffer (pH 8.5, 100 µL), 0.1 M EDTA (10 µL), 1 M Tris (2-carboxyethyl) phosphine hydrochloride (5 µL, Sigma, St. Louis, MO, USA), and water (38 µL), and the proteins were denatured for 10 min at 37 °C. Subsequently, 1 M 2-iodoacetamide (40 µL, Wako Pure Chemical Industries, Ltd., Osaka, Japan) was added to the denaturing solution to protect thiol residues in proteins. The mixture was incubated for 10 min at 37 °C in the dark. The solution was transferred into a 30 kDa ultrafiltration tube (Amicon Ultra 0.5 mL, Millipore Corp., Burlington, MA, USA) and centrifuged at 12,000 rpm for 30 min to remove the denaturing reagents. The denatured proteins trapped on the filter were washed with 0.1 M Tris-HCl buffer (pH 8.5, 400 µL), followed by centrifugation at 12,000 rpm for 40 min. Further, 0.1 M Tris-HCl buffer (pH 8.5, 200 µL), 0.1 µg/µL trypsin (20 µL, Wako Pure Chemical Industries, Ltd., Osaka, Japan), and 0.1 µg/µL lysyl endopeptidase (20 µL, Wako Pure Chemical Industries, Ltd., Osaka, Japan) were added to the ultrafiltration tube and the denatured proteins were digested for 16 h at 37 °C on the filter. After digestion, the solution was centrifuged for 30 min at 12,000 rpm. The filtrate, which contained digested peptides (including glycopeptides), was transferred to a 10 kDa ultrafiltration tube (Amicon Ultra 0.5 mL, Millipore Corp., MA, USA) and centrifuged for 10 min at 12,000 rpm. Most glycopeptides were trapped on the filter membrane, whereas most non-glycosylated peptides passed through the filter [14]. The trapped glycopeptide fraction was washed with 10 mM ammonium acetate in 10% acetonitrile and 90% water (400 µL) and transferred into a 1.5 mL tube, followed by drying up by vacuum centrifugation. The glycopeptides trapped on the filter were recovered and analyzed by liquid chromatography and mass spectrometry (LC-MS).

4.4. Liquid Chromatography and Mass Spectrometry

The LC-MS data were acquired on a liquid chromatography system (Agilent HP1200, Agilent Technologies, Palo Alto, CA, USA) equipped with a C18 column (Inertsil ODS-4, 2 µm, 100 Å, 100 mm × 1.5 mm ID, GL Science, Tokyo, Japan) coupled with an electrospray ionization quadrupole time-of-flight (Q-TOF) mass spectrometer (Agilent 6520, Agilent Technologies, Palo Alto, CA, USA) [35]. Solvent A was 0.1% formic acid in water, and solvent B was 0.1% formic acid in 9.9% water and 90% acetonitrile. The glycopeptides were eluted at a flow rate of 0.15 mL/min at 40 °C with a following gradient program: 0 to 7 min, 15% to 30% solvent B; 7 to 12 min, 30% to 50% solvent B; and 2 min hold at 100% solvent B. The mass spectrometer was operated in negative mode with the capillary voltage of 4000 V. The nebulizing gas pressure was 30 psi and the dry gas flow rate was 8 L/min at 350 °C. The injection volume was 5 µL. LC-MS raw data were converted to CSV-type data using Mass Hunter Export (Agilent Technologies, Palo Alto, CA, USA) and peak positions (retention times and m/z) and peak intensities (areas) were calculated using R (R 3.2.2, R Foundation, Vienna, Austria). The errors of retention time and m/z were corrected by the internal standard (fetal calf fetuin) peaks and all peak intensities were aligned with the tolerance of 0.5 min (retention time) and 0.08 Da (m/z) by Maeker Analysis developed by LSI Medience Corp. [14].

4.5. Generation of 2D Barcodes

In total, 1712 glycopeptides were chosen from more than 30,000 detected peaks by following three steps: (1) removing low reproducibility peaks (CV (coefficient of variation) > 50%), (2) removing low reliability peaks (S/N (signal to noise) < 5), (3) removing isotope, adduct, and fragment ions. Subsequently, residual 1712 glycopeptide peaks were used for CSGSA diagnostics. QC samples were previously prepared by pooling several EOC and non-EOC samples. The expression values of 1712 glycopeptides were obtained by calculating the ratio between sample and QC values. The glycopeptide expression values were allocated to 41 × 42 cells of Excel sheet matrix (Microsoft, redmont, WA, USA) according to the order of elution time (Rt-based) or the order of PCA loadings (PCA-based). To generate Rt-based 2D barcodes, glycopeptide expression values were first aligned in the order of the elution time. Then, the values were allocated into 41 × 42 matrix of Excel sheet from top-left to bottom-right corner. To produce PCA-based 2D barcodes, PCA analysis (SIMCA 13.0.3, Umetrics,

<https://umetrics.com/products/simca>) was first performed for 1712 glycopeptides and 351 examinees, and loading values of first and second components were obtained. Unit variance scaling and zero average centering were performed for 1712 glycopeptides before PCA calculation. Then, the glycopeptide expression values were sorted based on the first loading values in the ascending order and divided into 41 groups with 42 glycopeptides. The glycopeptide expression values were further sorted in different groups based on the second loading values. The coordinates in 2D barcodes were determined in such a way that X coordinate represented the number of groups and Y represented the order in each group. After allocating 1712 glycopeptide expression values into Excel cells, the cell's brightness was determined according to the criterion described in Table S1 or S2. The generated 2D barcodes were converted to PNG format to let CNN learn.

4.6. Training CNN to Distinguish Between EOC and Non-EOC Patterns

MATLAB and its deep learning toolbox (R2018b, Mathworks, Natick, MA, USA), which provides an environment of deep learning and neural network implementation, was used for training and testing of the CNN to discriminate early-stage EOC. AlexNet was obtained from Mathworks add-in programs. We removed last 3 layers of AlexNet and added a new fully connected layer with two neurons, followed by softmax and classification layer, where we classified the data set into early-stage EOC or non-EOC. In this study, batch size was set to 5, max epochs were set to 30, and initial learning rate was set to 0.0001. The training was performed on a desktop having NVIDIA GeForce GTX 960M DDR5 4096 MB GPU, which took around 10 min. The parameters of CNN were optimized using 60% of samples, and its validity was tested using 40% samples.

5. Conclusions

In this study, we showed that the pretrained CNN model AlexNet can recognize early-stage EOC using PCA-based 2D barcodes. The ROC further increased when 2D barcodes were generated based on the serum levels of CA125 and HE4. We observed that principal component analysis-based alignment of glycopeptides to generate 2D barcodes significantly increased the diagnostic accuracy of the method, and it reached 95.1% of ROC-AUC. We believe that this simple and low-cost method will reduce economical and physical burden on the patients and contribute to the improvement of the survival rates of EOC.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-6694/12/9/2373/s1>, Figure S1: Accuracy and loss converging curve, Table S1: The method of creating 2D barcodes from the levels of glycopeptide expression, Table S2: The method of creating 2D barcodes from the levels of glycopeptide expression and serum levels of CA125 and HE4, Table S3: Exclusion criteria of participants.

Author Contributions: Conceptualization, M.M., M.H., and K.T.; development of CSGSA-AI methodology, M.M., K.T. and T.I.; case reports, M.H., M.I., M.Y., H.M., M.S., K.S., H.Y., T.H. and M.M.; formal analysis, M.H., T.K., T.I. and K.T.; writing—original draft preparation, K.T., K.M., and M.M.; writing—review and editing, all authors; funding acquisition, supervision and administration, M.M., M.S., M.I. and K.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by a grant-in-aid for scientific research from the Ministry of Education, Culture, Sports, Science and Technology (No. 17H04340, M.M.; 18K09274, M.S.; and 18K09300, M.I.), AMED (No. 20lm0203004j0003, M.M.), and Ensign Endowment for Gynecologic Cancer Research (K.M.).

Acknowledgments: We thank Nozomi Kojima and Kae Kitagawa for the development of glycopeptide analysis workflow. Further, we would like to thank Sadayo Iijima for supporting this project.

Conflicts of Interest: The authors are employed by LSI Medience Corporation, which provides ovarian cancer diagnostics.

References

1. Ward, E.M.; Sherman, R.L.; Henley, S.J.; Jemal, A.; Siegel, D.A.; Feuer, E.J.; Firth, A.U.; Kohler, B.A.; Scott, S.; Ma, J.; et al. Annual Report to the Nation on the Status of Cancer, Featuring Cancer in Men and Women Age 20–49 Years. *J. Natl. Cancer Inst.* **2019**, *111*, 1279–1297. [[CrossRef](#)] [[PubMed](#)]
2. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2020. *CA Cancer J. Clin.* **2020**, *70*, 7–30. [[CrossRef](#)] [[PubMed](#)]
3. Henley, S.J.; Ward, E.M.; Scott, S.; Ma, J.; Anderson, R.N.; Firth, A.U.; Thomas, C.C.; Islami, F.; Weir, H.K.; Lewis, D.R.; et al. Annual report to the nation on the status of cancer, part I: National cancer statistics. *Cancer* **2020**, *126*, 2225–2249. [[CrossRef](#)] [[PubMed](#)]
4. Jemal, A.; Ward, E.M.; Johnson, C.J.; Cronin, K.A.; Ma, J.; Ryerson, B.; Mariotto, A.; Lake, A.J.; Wilson, R.; Sherman, R.L.; et al. Annual Report to the Nation on the Status of Cancer, 1975–2014, Featuring Survival. *J. Natl. Cancer Inst.* **2017**, *109*. [[CrossRef](#)] [[PubMed](#)]
5. National Cancer Institute. Cancer Stat Facts: Ovarian Cancer. Available online: <https://seer.cancer.gov/statfacts/html/ovary.html> (accessed on 10 May 2020).
6. Khiewvan, B.; Torigian, D.A.; Emamzadehfard, S.; Paydary, K.; Salavati, A.; Houshmand, S.; Werner, T.J.; Alavi, A. An update on the role of PET/CT and PET/MRI in ovarian cancer. *Eur. J. Nucl. Med. Mol. Imaging* **2017**, *44*, 1079–1091. [[CrossRef](#)]
7. Sironi, S.; Messa, C.; Mangili, G.; Zangheri, B.; Aletti, G.; Garavaglia, E.; Vigano, R.; Picchio, M.; Taccagni, G.; Maschio, A.D.; et al. Integrated FDG PET/CT in patients with persistent ovarian cancer: Correlation with histologic findings. *Radiology* **2004**, *233*, 433–440. [[CrossRef](#)]
8. Meany, D.L.; Sokoll, L.J.; Chan, D.W. Early Detection of Cancer: Immunoassays for Plasma Tumor Markers. *Expert Opin. Med. Diagn.* **2009**, *3*, 597–605. [[CrossRef](#)]
9. Dorigo, O.; Berek, J.S. Personalizing CA125 levels for ovarian cancer screening. *Cancer Prev. Res.* **2011**, *4*, 1356–1359. [[CrossRef](#)]
10. Buys, S.S.; Partridge, E.; Black, A.; Johnson, C.C.; Lamerato, L.; Isaacs, C.; Reding, D.J.; Greenlee, R.T.; Yokochi, L.A.; Kessel, B.; et al. Effect of screening on ovarian cancer mortality: The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA* **2011**, *305*, 2295–2303. [[CrossRef](#)]
11. Pinsky, P.F.; Yu, K.; Kramer, B.S.; Black, A.; Buys, S.S.; Partridge, E.; Gohagan, J.; Berg, C.D.; Prorok, P.C. Extended mortality results for ovarian cancer screening in the PLCO trial with median 15 years follow-up. *Gynecol. Oncol.* **2016**, *143*, 270–275. [[CrossRef](#)]
12. Yokoi, A.; Matsuzaki, J.; Yamamoto, Y.; Yoneoka, Y.; Takahashi, K.; Shimizu, H.; Uehara, T.; Ishikawa, M.; Ikeda, S.I.; Sonoda, T.; et al. Integrated extracellular microRNA profiling for ovarian cancer screening. *Nat. Commun.* **2018**, *9*, 4319. [[CrossRef](#)] [[PubMed](#)]
13. Mari, R.; Mamessier, E.; Lambaudie, E.; Provansal, M.; Birnbaum, D.; Bertucci, F.; Sabatier, R. Liquid Biopsies for Ovarian Carcinoma: How Blood Tests May Improve the Clinical Management of a Deadly Disease. *Cancers* **2019**, *11*, 774. [[CrossRef](#)] [[PubMed](#)]
14. Tanabe, K.; Kitagawa, K.; Kojima, N.; Iijima, S. Multifucosylated Alpha-1-acid Glycoprotein as a Novel Marker for Hepatocellular Carcinoma. *J. Proteome Res.* **2016**, *15*, 2935–2944. [[CrossRef](#)] [[PubMed](#)]
15. Mikami, M.; Tanabe, K.; Matsuo, K.; Miyazaki, Y.; Miyazawa, M.; Hayashi, M.; Asai, S.; Ikeda, M.; Shida, M.; Hirasawa, T.; et al. Fully-sialylated alpha-chain of complement 4-binding protein: Diagnostic utility for ovarian clear cell carcinoma. *Gynecol. Oncol.* **2015**, *139*, 520–528. [[CrossRef](#)]
16. Matsuo, K.; Tanabe, K.; Ikeda, M.; Shibata, T.; Kajiwarra, H.; Miyazawa, M.; Miyazawa, M.; Hayashi, M.; Shida, M.; Hirasawa, T.; et al. Fully sialylated alpha-chain of complement 4-binding protein (A2160): A novel prognostic marker for epithelial ovarian carcinoma. *Arch. Gynecol. Obs.* **2018**, *297*, 749–756. [[CrossRef](#)]
17. Tanabe, K.; Matsuo, K.; Miyazawa, M.; Hayashi, M.; Ikeda, M.; Shida, M.; Hirasawa, T.; Sho, R.; Mikami, M. UPLC-MS/MS based diagnostics for epithelial ovarian cancer using fully sialylated C4-binding protein. *Biomed. Chromatogr. BMC* **2018**, *32*, e4180. [[CrossRef](#)]
18. Hayashi, M.; Matsuo, K.; Tanabe, K.; Ikeda, M.; Miyazawa, M.; Yasaka, M.; Machida, H.; Shida, M.; Imanishi, T.; Grubbs, B.H.; et al. Comprehensive Serum Glycopeptide Spectra Analysis (CSGSA): A Potential New Tool for Early Detection of Ovarian Cancer. *Cancers* **2019**, *11*, 591. [[CrossRef](#)]

19. Tang, Z.; Chuang, K.V.; DeCarli, C.; Jin, L.W.; Beckett, L.; Keiser, M.J.; Dugger, B.N. Interpretable classification of Alzheimer's disease pathologies with a convolutional neural network pipeline. *Nat. Commun.* **2019**, *10*, 2173. [CrossRef]
20. Liu, W.; Liao, S.; Hu, W. Efficient Single-Stage Pedestrian Detector by Asymptotic Localization Fitting and Multi-Scale Context Encoding. *IEEE Trans. Image Process. Publ. IEEE Signal Process. Soc.* **2019**, *29*, 1413–1425. [CrossRef]
21. Ding, C.; Tao, D. Trunk-Branch Ensemble Convolutional Neural Networks for Video-Based Face Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 1002–1014. [CrossRef]
22. Kim, T.; Kim, H.Y. Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. *PLoS ONE* **2019**, *14*, e0212320. [CrossRef]
23. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
24. Chang, K.; Bai, H.X.; Zhou, H.; Su, C.; Bi, W.L.; Agbodza, E.; Kavouridis, V.K.; Senders, J.T.; Boaro, A.; Beers, A.; et al. Residual Convolutional Neural Network for the Determination of IDH Status in Low- and High-Grade Gliomas from MR Imaging. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2018**, *24*, 1073–1081. [CrossRef]
25. Oh, K.; Chung, Y.C.; Kim, K.W.; Kim, W.S.; Oh, I.S. Classification and Visualization of Alzheimer's Disease using Volumetric Convolutional Neural Network and Transfer Learning. *Sci. Rep.* **2019**, *9*, 18150. [CrossRef]
26. Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [CrossRef]
27. Raghu, S.; Sriraam, N.; Temel, Y.; Rao, S.V.; Kubben, P.L. EEG based multi-class seizure type classification using convolutional neural network and transfer learning. *Neural Netw. Off. J. Int. Neural Netw. Soc.* **2020**, *124*, 202–212. [CrossRef]
28. Lee, K.S.; Jung, S.K.; Ryu, J.J.; Shin, S.W.; Choi, J. Evaluation of Transfer Learning with Deep Convolutional Neural Networks for Screening Osteoporosis in Dental Panoramic Radiographs. *J. Clin. Med.* **2020**, *9*, 392. [CrossRef]
29. Sharma, A.; Vans, E.; Shigemizu, D.; Boroevich, K.A.; Tsunoda, T. DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architecture. *Sci. Rep.* **2019**, *9*, 11399. [CrossRef]
30. MathWorks. Alexnet. Available online: <https://jp.mathworks.com/help/deeplearning/ref/alexnet.html?lang=en> (accessed on 5 August 2020).
31. Li, F.-F.; Deng, J.; Russakovsky, O.; Berg, A.; Li, K. ImageNet. Available online: <http://www.image-net.org/> (accessed on 19 May 2020).
32. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the Neural Information Processing Systems Conference, Montreal, QC, Canada, 8–13 December 2014; pp. 2672–2680.
33. Marouf, M.; Machart, P.; Bansal, V.; Kilian, C.; Magruder, D.S.; Krebs, C.F.; Bonn, S. Realistic in silico generation and augmentation of single-cell RNA-seq data using generative adversarial networks. *Nat. Commun.* **2020**, *11*, 166. [CrossRef]
34. Prat, J. Staging classification for cancer of the ovary, fallopian tube, and peritoneum. *Int. J. Gynaecol. Obstet. Off. Organ Int. Fed. Gynaecol. Obstet.* **2014**, *124*, 1–5. [CrossRef]
35. Shinozaki, E.; Tanabe, K.; Akiyoshi, T.; Tsuchida, T.; Miyazaki, Y.; Kojima, N.; Igarashi, M.; Ueno, M.; Suenaga, M.; Mizunuma, N.; et al. Serum leucine-rich alpha-2-glycoprotein-1 with fucosylated triantennary N-glycan: A novel colorectal cancer marker. *BMC Cancer* **2018**, *18*, 406. [CrossRef] [PubMed]

