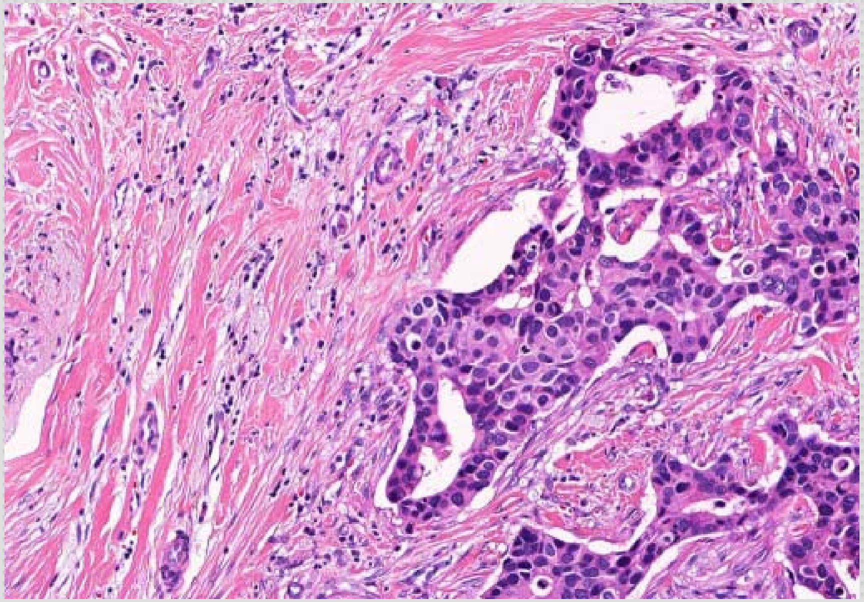
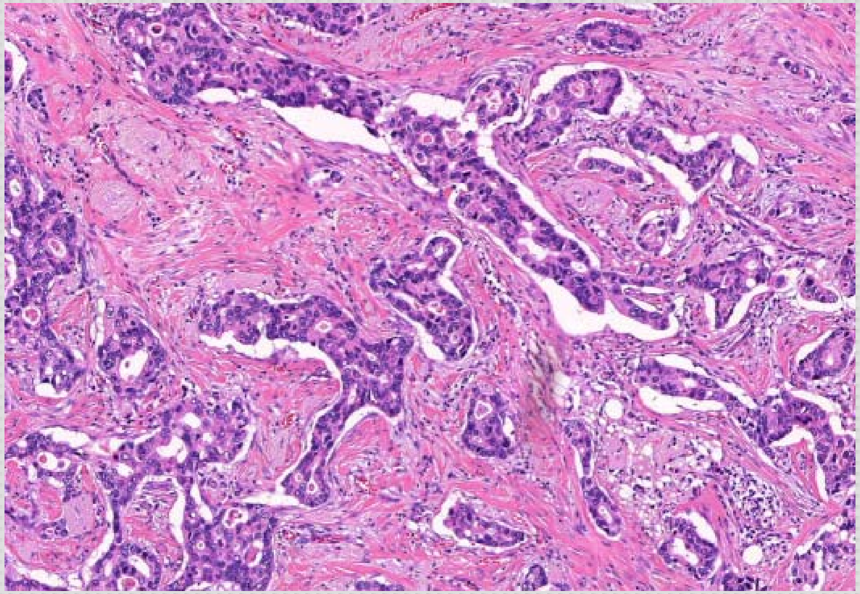
Neural network approach for histopathological diagnosis of breast diseases with images

Yuichi Ishibashi (Okayama University)
Atsuko Hara (Kitasato University)
Isao Okayasu (Kitasato University)
Koji Kurihara (Okayama University)

Abstract

- Diagnosis of breast diseases relies on recognizing diseased tissue in histopathological images. The tissues studied will contain both diseased and normal areas.





Examples of breast cancer: Invasive ductal carcinoma (scirrhous type)

Abstract

The method to insure a correct diagnosis

1. to subdivide the histopathological image into sections.
2. These subdivisions will then all be digitized by Wavelet transformation.
3. To evaluate by neural network analysis.

- The collective evaluation of subdivisions will increase the accuracy of diagnosis and help to avoid missing cancerous or inflamed tissue.

Histopathological diagnosis

Determines whether the lesion is

- tumor or
  - benign or
  - malignant
- not tumor

If it is a tumor then it is important to decide what kind of treatment would be required

Diagnosing the kind of tumor is important.

Classification of breast cancer

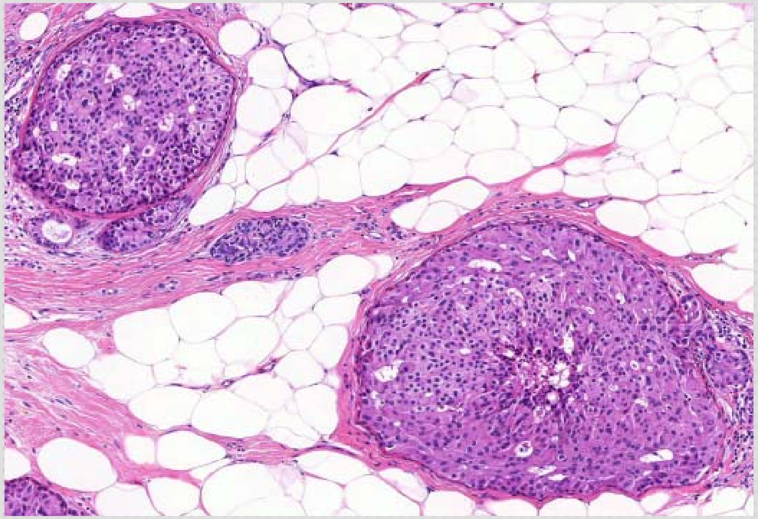
Classification					Disease
EPITHELIAL TUMORS	Benign			IA1	1.Intraductal papilloma
				IA2	2.Ductal adenoma
				IA3	3.Adenoma of the nipple
				IA4	4. Adenoma
				IA5	5.Adenomyoepithelioma
	Malignant	Noninvasive		IB1a	a.Noninvasive ductal carcinoma
				IB1b	b.Lobular carcinoma in situ
		Invasive	Invasive ductal carcinoma	IB2a1	a1.Papillotubular carcinoma
				IB2a2	a2.Solid-tubular carcinoma
				IB2a3	a3.Scirrhous carcinoma
			Special types	IB2b1	b1.Mucinous carcinoma
				IB2b2	b2.Medullary carcinoma
				IB2b3	b3.Invasive lobular carcinoma
				IB2b4	b4Adenoid cystic carcinoma
				IB2b5	b5.Squamous cell carcinoma
				IB2b6	b6.Spindle cell carcinoma
				IB2b7	b7.Apocrine carcinoma
				IB2b8	b8.Carcinoma with cartilaginous and/or osseous metaplasia
				IB2b9	b9.Tubular carcinoma
				IB2b10	b10.Secretory carcinoma（Juvenile carcinoma）
				IB2b11	b11.Invasive micropapillary carcinoma
				IB2b12	b12.Matrix-producing carcinoma
				IB2b13	b13.Others
		Paget's disease		IB3	3Paget's disease

Classification of breast cancer

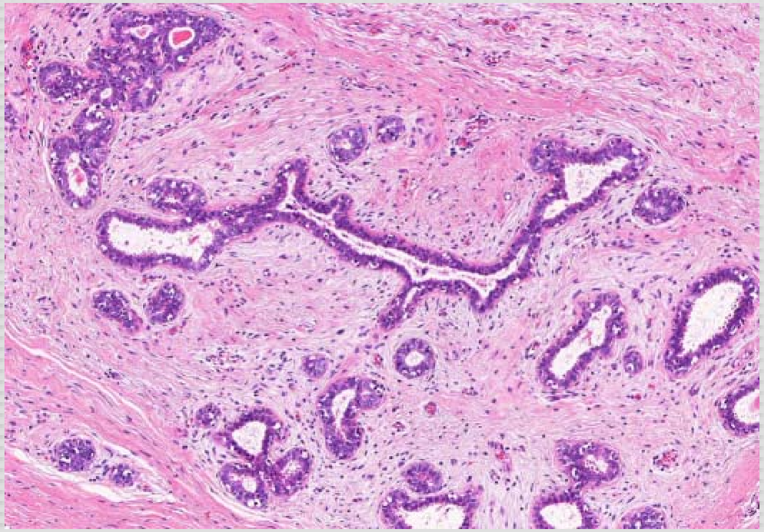
Classification		Disease
MIXED CONNECTIVE TISSUE AND EPITHELIAL TUMORS	IIA	A.Fibroadenoma
	IIB	B.Phyllodes tumor
	IIC	C.Carcinosarcoma
NONEPITHEILIAL TUMORS	IIIA	A.Stromal sarcoma
	IIIB	B.Soft tissue tumors
	IIIC	C.Lymphomas and hematopoietic tumors
	IIID	D.Others
UNCLASSIFIED TUMORS	IV	IV.UNCLASSIFIED TUMORS
MASTOPATHY	V	V.MASTOPATHY (FIBROCYTSTIC DISEASE, MAMMRY DIYPLASIA)
TUMOR-LIKE LESIONS	VIA	A.Duct ectasia
	VIB	B.Inflammatory pseudotumor
	VIC	C.Hamartoma
	VID	D.Gynecomastia
	VIE	EAccessory mammary gland
	VIF	F.Others
BORDERLINE LESION	VIIA	A.Atypical ductal hyperplasia
	VIIB	B.Atypical lobular hyperplasia
	VIIC	C.others

Diagnosis by images

- This study attempts to differentiate not only tumors but also inflammations and borderline lesions.



DCIS(cribriform-type)
Non invasive ductal carcinoma



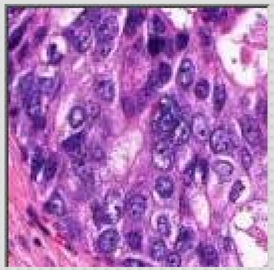
Fibrocystic disease(fibroadenomatosis)

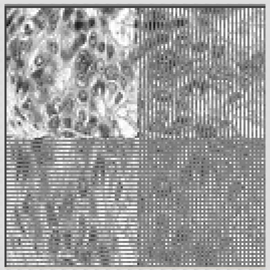
Texture analysis

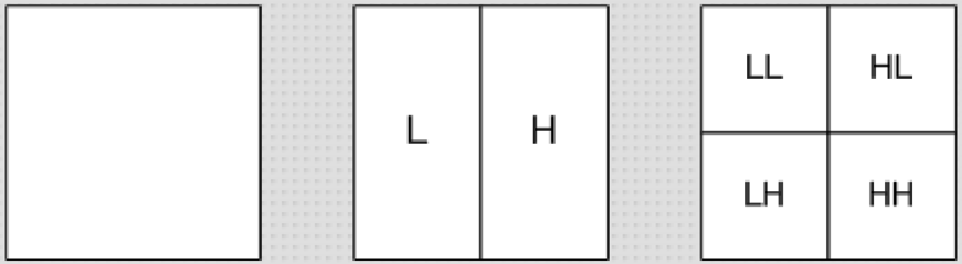
- To numerically characterize the specific variation pattern of image element values in the picture image region
- We digitized the texture information of histopathological images in order to examine the structural patterns of specimens.
- Wavelet transformation was applied

Wavelet transformation

1. In the horizontal direction one-dimensional Wavelet transform for each row divides the image into high and low frequency components.
2. Then, for each column this converted signal is performed by one-dimensional transformation in the vertical direction. One two-dimensional wavelet transform in horizontal and vertical directions divides the original signal into four components, such as LL, LH, HL and HH sub-bands.
3. Two-dimensional Wavelet transformation is adapted to LL component recursively.

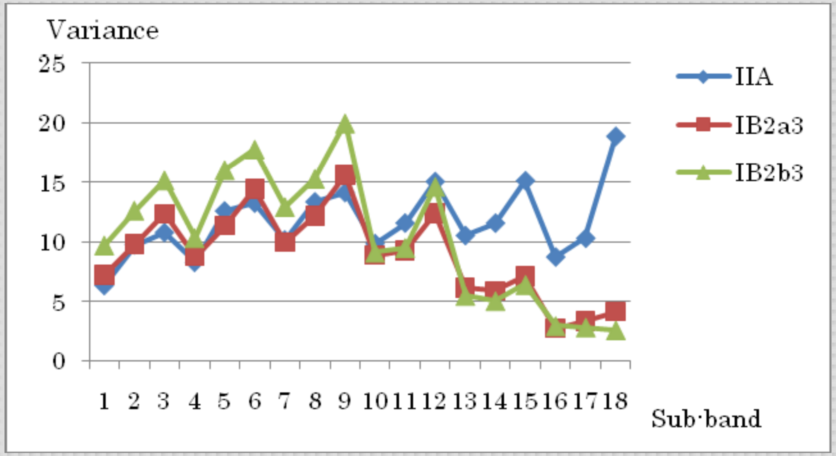


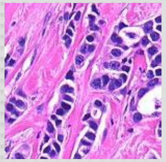




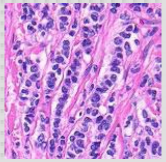
Original image | Dual-partitioning for each row | Dual-partitioning for each column

The variances in each sub-band

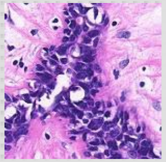




IB2a3



IB2b3

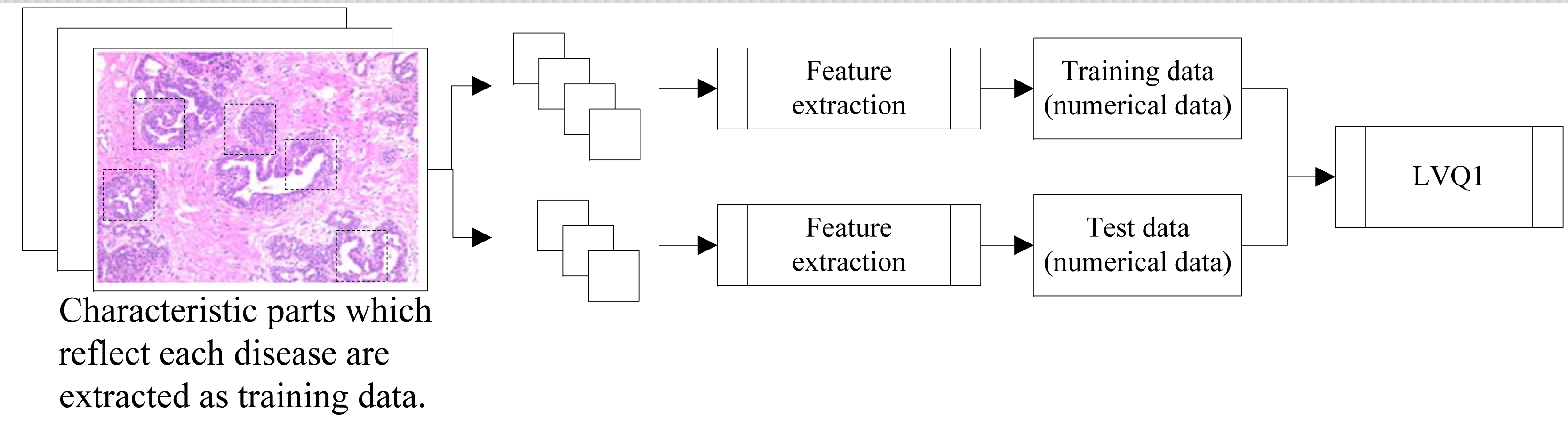


IIA

128X128 pixel images on the right side are extracted as characteristic parts.

Restibrachium is found in IB2a3(Scirrhous carcinoma)and IB2b3(Invasive lobular carcinoma) and the forms of changes in the graph are similar. But IIA(Fibroadenoma) is different from the others in the graph and image. As described above Wavelet feature reflects texture information, therefore classification and recognition using Wavelet feature is appropriate.

Feature extraction and recognition by Neural Network (LVQ1)



Pattern recognition using neural network

The algorithm of LVQ1

Input data: $\mathbf{x} \in R^p$ Label: $y \in \{1,2,..,G\}$

Training data: $\{(\mathbf{x}_1, y_1),...,(\mathbf{x}_n, y_n)\}$

Assuming that k sets of codebook vector and label: $\{(\mathbf{m}_i, l_i), i = 1,..,k\}$

LVQ divides an input space using a finite number of labeled codebook vectors and differentiates. In sequential type one data is selected at time t and the codebook vector is updated. In LVQ1 the codebook vector and the label are updated by the following expression.

$$\mathbf{m}_c(t+1) = \begin{cases} \mathbf{m}_c(t) + \alpha(t)(\mathbf{x}(t) - \mathbf{m}_c(t)), y(t) = l_c(t) \\ \mathbf{m}_c(t) - \alpha(t)(\mathbf{x}(t) - \mathbf{m}_c(t)), y(t) \neq l_c(t) \end{cases}$$

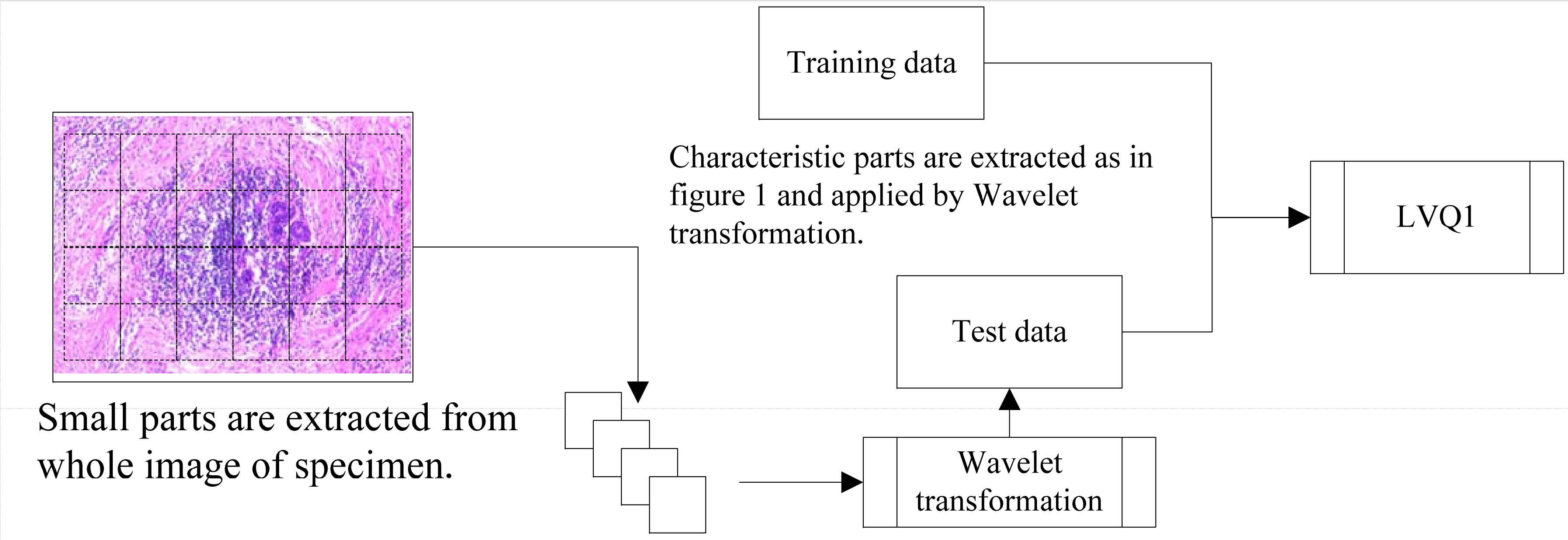
Recognition results by LVQ1

There were 211 small images extracted from 9 kinds of diseases. Each disease contains 3 to 5 different cases. 211 images were divided into 141 training data and 70 test data.

Classification		IB1a	IB1b	IB2a1	IB2a2	IB2a3	IB2b3	IIA	IX	VIIA	Error Rates
IB1a	Noninvasive ductal carcinoma	8	0	1	0	0	0	0	0	1	0.200
IB1b	Lobular carcinoma in situ	0	10	0	0	0	0	0	0	0	0.000
IB2a1	Papillotubular carcinoma	0	1	3	0	1	0	0	1	1	0.571
IB2a2	Solid-tubular carcinoma	0	0	0	5	0	0	0	0	0	0.000
IB2a3	Scirrhous carcinoma	0	1	0	1	9	0	0	0	0	0.182
IB2b3	Invasive lobular carcinoma	0	0	0	0	0	5	0	0	0	0.000
IIA	Fibroadenoma	0	0	0	0	0	0	3	1	0	0.250
IX	Normal	0	1	0	0	0	0	1	5	0	0.286
VIIA	Atypical ductal hyperplasia	0	0	0	0	0	0	0	2	9	0.182
Total											0.186

Wavelet transformation for a whole case image and the method of recognition by LVQ1

Test data are transformed values by Wavelet transformation from the 128X128 pixel areas which are all over the image of a new case.



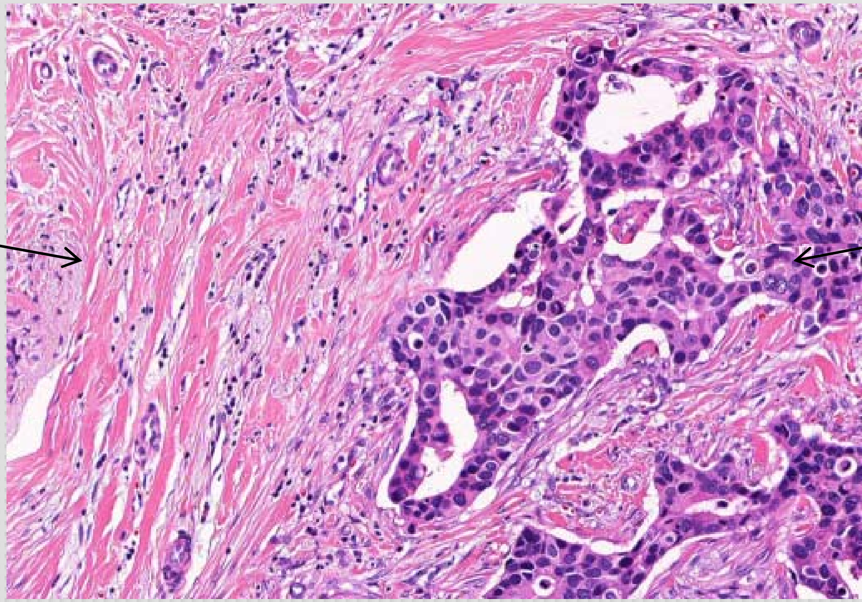
Recognition results by LVQ1 for a whole case image

	Classification	IB1a	IB1b	IB2a1	IB2a2	IB2a3	IB2b3	IIA	IX	VIIA	Error Rates
IB2a1	Papillotubular carcinoma	0	0	138	0	0	0	2	0	0	0.014
IB2a2	Solid-tubular carcinoma	5	4	4	69	39	4	0	1	0	0.452
IB2a3	Scirrhous carcinoma	0	61	0	0	61	0	0	1	2	0.512
IB1a	Noninvasive ductal carcinoma	55	0	0	0	15	0	0	28	29	0.567
IB1b	Lobular carcinoma in situ	0	122	0	0	3	0	0	0	2	0.039
IB2b3	Invasive lobular carcinoma	6	4	0	3	10	102	0	0	1	0.190
VIIA	Atypical ductal hyperplasia	0	0	0	0	0	0	0	0	127	0.000
IIA	Fibroadenoma	2	1	0	0	50	0	14	30	30	0.890
IX	Normal	1	44	0	3	5	0	0	36	37	0.714

Including non-characteristic parts for training data

Training data are extracted from characteristic parts of each disease, but a specimen contains not only characteristic parts but also non characteristic parts, such as interstitium tissue etc. Neural network tries to recognize non-characteristic parts as some sort of disease.

Interstitium tissue



Cancerous tissue

Invasive ductal carcinoma (scirrhous type)

Recognition results of improved method by LVQ1 for a whole case image

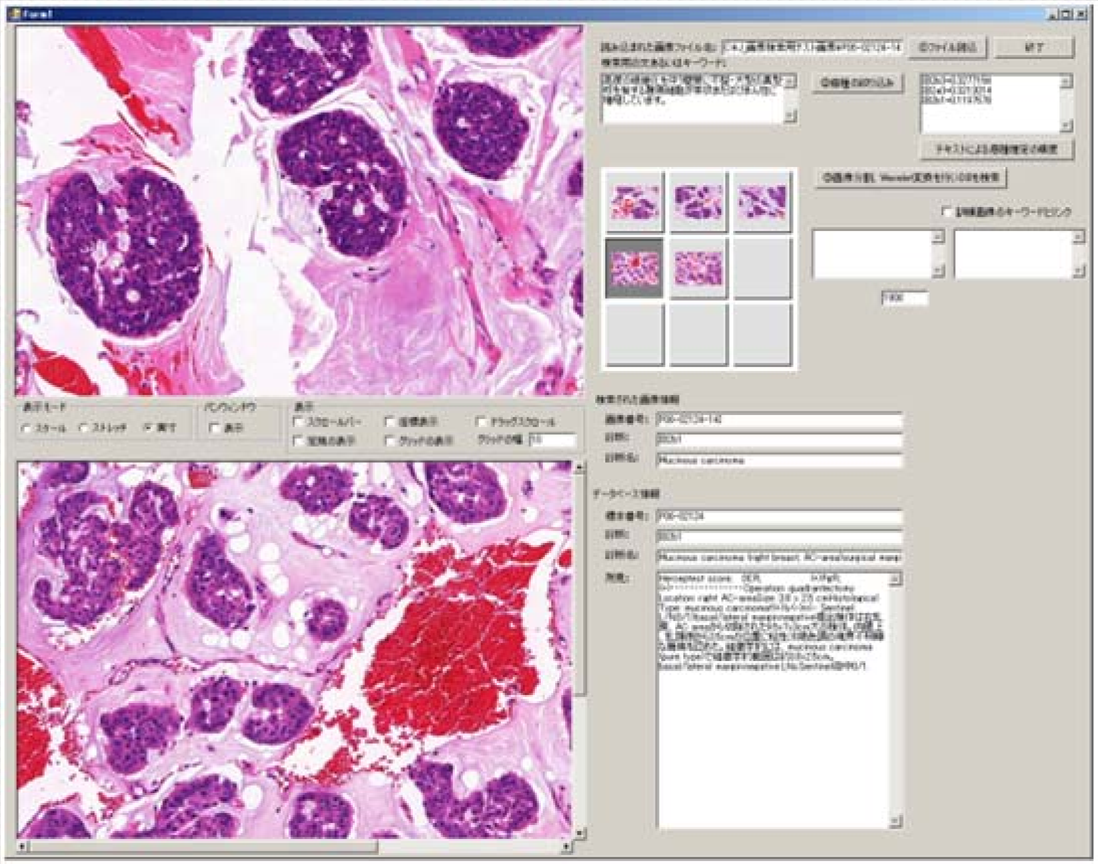
Classification		IB1a	IB1b	IB2a1	IB2a2	IB2a3	IB2b3	IIA	IIAN	IX	IXN	VIIA	Error*
IX	Normal	1	3	0	17	2	0	0	4	32	52	15	0.333
IIA	Fibroadenoma	5	0	0	1	5	0	18	51	30	0	17	0.457

Classification		Error rate	
		Only characteristic parts	Including non-characteristic parts
IX	Normal	0.714	0.333
IIA	Fibroadenoma	0.890	0.457

Conclusion

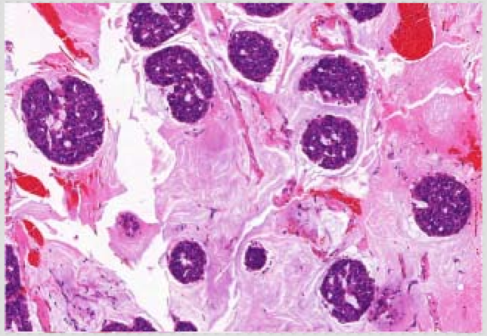
- LVQ with Wavelet transformation of different diseases as training data enables the diagnosis of breast disease.
- There are more than 50 types of breast disease and some types contain different patterns of lesion, such as atypical ductal hyperplasia.
- Many more kinds of image data should be accumulated in order to diagnose these diseases.

Histopathological information data base

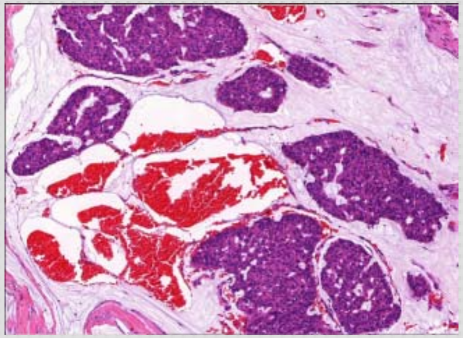


Similar image retrieval in database

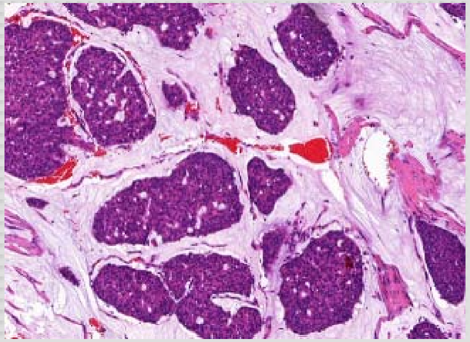
Test data



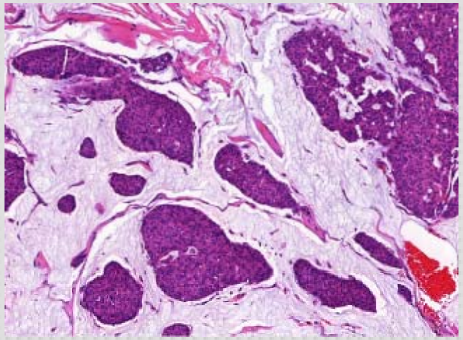
Retrieved images



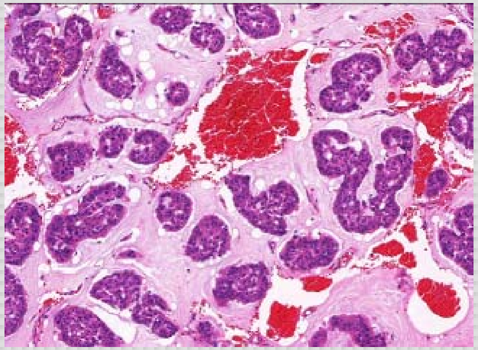
IB2b1(Mucinous carcinoma)



IB2b1(Mucinous carcinoma)



IB2b1(Mucinous carcinoma)



IB2b1(Mucinous carcinoma)