J. Korea Saf. Manag. Sci. Vol. 23 No. 3 September 2021
http://dx.doi.org/10.12812/ksms.2021.23.3.083
ISSN 1229-6783(Print)
ISSN 2288-1484(Online)

SVM 알고리즘을 활용한 선루프 실러도포 공정 품질검사 시스템 구축

양 희 종* · 장 길 상**
*주식회사 아이티엔제이 대표이사
**울산대학교 경영대학 경영정보학과

The Construction of Quality Inspection System for Sunroof Sealer Application Process Using SVM Algorithm

Hee-Jong Yang* · Gil-Sang Jang**
*CEO, ITNJ, Inc.
**Dept. of Management Information Systems, University of Ulsan

Abstract

Recently, due to the aging of workers and the weakening of the labor base in the automobile industry, research on quality inspection methods through ICT(Information and Communication Technology) convergence is being actively conducted. A lot of research has already been done on the development of an automated system for quality inspection in the manufacturing process using image processing. However, there is a limit to detecting defects occurring in the automotive sunroof sealer application process, which is the subject of this study, only by image processing using a general camera. To solve this problem, this paper proposes a system construction method that collects image information using a infrared thermal imaging camera for the sunroof sealer application process and detects possible product defects based on the SVM(Support Vector Machine) algorithm. The proposed system construction method was actually tested and applied to auto parts makers equipped with the sunroof sealer application process, and as a result, the superiority, reliability, and field applicability of the proposed method were proven.

Keywords : SVM, Defect detection, Sunroof, Infrared thermal imaging, Quality inspection system

1. 서 론

최근 국내 자동차 업체들은 세계 시장 점유율이 확대됨에 따라 자동차 부품 수급이 증가하며 이와 함께 품질에 대한 인식도 높아지고 있다. 제품 제조과정에 발생하는 결함을 영상처리 기반으로 검출 하는 연구가 활발히 이루어지고 있다[2]. 자동차 부품 중 선루프 제품은 제조공정 가운데 특히 실러(Sealer) 도포 공정에서 많은 결함이 발생하고 있다. 실러를 선루프 프레임에 도포하는 과정에 자동화 로봇에 의하여 균일하게 실러가 도포가 되지 않으면 프레임과 유리 압착 공정에서 유격이 발생하여 제품의 결함으로 이어진다. 실러와 선루프 프레임이 색상이 동일하여 일반 영상처리 방법으로는 제품의 결함을 판단하기엔 한계가 있다. 실러도포 공정을 분석한 결과 실러의 온도와 제품 프레임의 온도 차이가 있는 것으로 분석되었다.

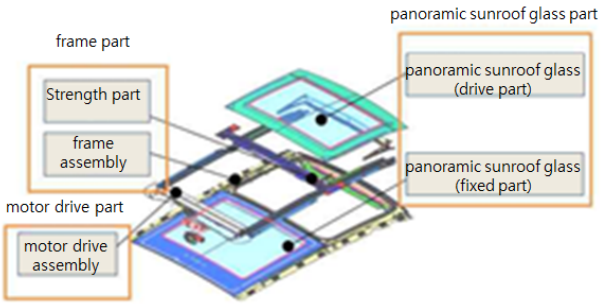
본 논문에서는 표면의 온도를 측정할 수 있는 열화상 카메라를 이용하여 실러도포 과정을 촬영하고 수집된 결과를 영상 처리를 통해 결함의 특징을 수치화한다. 수치화된 특징들을 SVM(Support Vector Machine) 기계학습 알고리즘을 활용하여 학습하여 실러 도포 결함을 검출하는 방법에 대하여 소개한다.

본 논문은 먼저, 선루프의 구조와 제조 공정 가운데 열화상 카메라를 이용하여 영상을 수집하는 방법에 대하여 설명하고, 다음으로 수집된 영상 정보에서 결함의 유형을 분석하고 찾기 위한 SVM 알고리즘을 제안하며, 끝으로 제안한 알고리즘의 평가 방법 및 결과를 제시하는 순서로 구성된다.

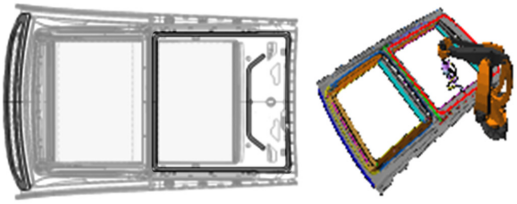
†본 연구는 울산대학교의 2016년도 연구비 지원 사업에 의하여 연구되었음.
†Corresponding Author : Gil-Sang Jang, Dept. of MIS, University of Ulsan, 93, Daehak-ro, Nam-gu, Ulsan, E-mail: gsjang@ulsan.ac.kr
Received September 13, 2021; Revision September 27, 2021; Accepted September 28, 2021

2. 선루프 구조와 영상데이터 수집 방법

선루프 모듈의 구조는 [Figure 1]과 같이 선루프 유리, 프레임, 모터 구동 장치로 크게 3가지로 이루어져 있다. 선루프 모듈 조립 과정은 [Figure 2]의 프레임 부에 실러(Sealer)를 진한 굵은 실선(bold line) 영역에 도포하고 유리를 접합하는 공정을 가진다. 이때 실러가 프레임에 균일하게 도포 되는 결과에 따라 제품의 결함이 결정된다.

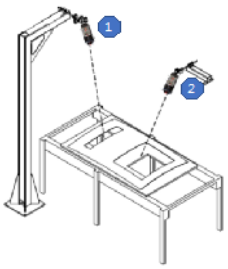


[Figure 1] Sunroof structure

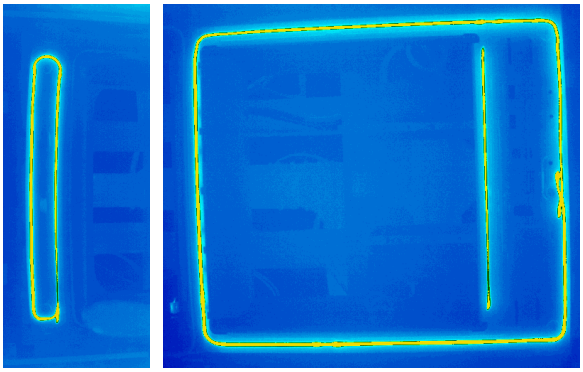


[Figure 2] The area of the sealer applied to the frame (marked with bold lines) (left), and the automated robot applying the sealer to the frame (right)

프레임 부에 실러를 도포한 결과를 수집하기 위하여 [Figure 3]과 같이 열화상 카메라 2대를 설치하였다. 1번 카메라는 front 영역을 촬영하고, 2번 카메라는 rear 영역을 촬영하여 표면의 열분포 영상을 얻는다. 공정에서 사용된 열화상 카메라 모델은 flir A65이며 640 * 512 픽셀의 해상도 가지고 있다. 한 공정에 2개의 640 * 512 픽셀의 표면 온도 정보를 데이터베이스 저장하고 온도값을 RGB 색상으로 변환하여 [Figure 4]와 같이 이미지로도 저장한다.



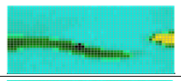
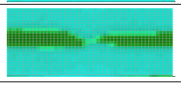
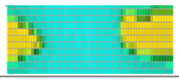
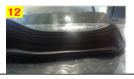
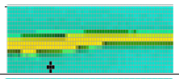
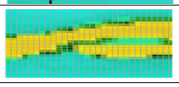
[Figure 3] Image data collection method with sealer - applied thermal imaging camera



[Figure 4] Thermal image information of sealer application

열화상 카메라로부터 얻은 실러도포 표면 영상은 327,680 픽셀로 이루어져 있으며 배경을 제외한 실제 실러도포 표면은 약 23,000픽셀의 너비를 가지고 있다. [Figure 4]에서 보는 것과 같이 실러는 프레임과 온도차이가 20 ° C 이상 차이를 나타내며, 도포 경계를 식별할 수 있다. 균일한 실러도포 판단을 위해 온도 정보는 매우 중요한 요인이다.

본 논문에서 탐지하고자 하는 결함은 실러의 폭이 6mm이하 또는 12 mm이상의 결함과 2mm 이상의 실러의 끊김을 발견하는 것이다. [Figure 5]에서 다양한 결함의 유형을 보여주는 실러 실사 이미지와 열화상 프레임 영상이다.

Real picture	thermal imaging	Error type
		Broken
		Fine broken
		Broken
		Width error
		Fine broken

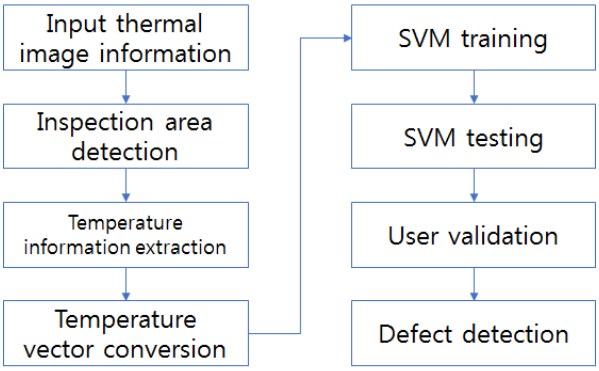
[Figure 5] Examples of defect types

3. 제안하는 SVM 알고리즘

앞서 설명한 결함 유형을 영상에서 검출하기 위해 [Figure 6]의 프로세스로 진행한다. 먼저 열화상 카메라로부터 표면의 온도분포 정보를 수집한다. 수집된 온도정보로부터 검사영역(실러 도포 영역)만 추출하기 위해 분

ISSN 1229 - 6783(Print)
ISSN 2288 - 1484(Online)

할(segmentation) 연산을 수행한 뒤 탐상하고자 하는 영역의 온도 값들을 추출한다. 추출된 온도 값들은 분류기에서 학습 할 수 있는 형태의 백터로 수치화 하고 수치화된 백터 정보를 학습 및 판정 하는 일련의 과정을 통해서 실러도포의 결함을 검출한다.



[Figure 6] Defect detection process

[Figure 6]의 결함검출 프로세스에 대한 주요 내용을 살펴보면 다음과 같다.

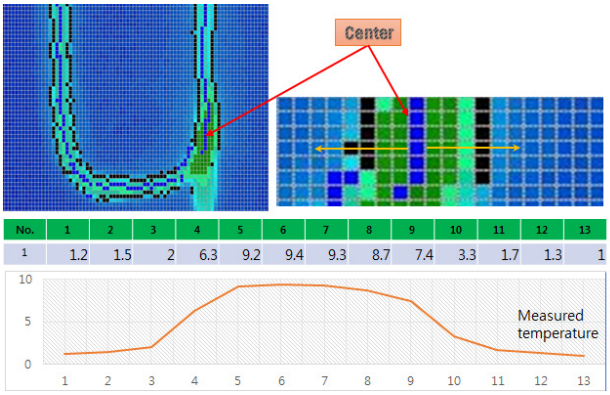
3.1 검사영역 검출

수집된 열분포 정보에서 탐상하고자 하는 실러 도포 영역을 검출하기 위해 실러도포 전에 열분포 정보 A에서 실러 도포 후 열분포 정보 A'차를 구하여 실러도포 된 영역 S을 도출시킨다.

$$abs(A) - abs(A') = S \quad (1)$$

3.2 온도정보 추출

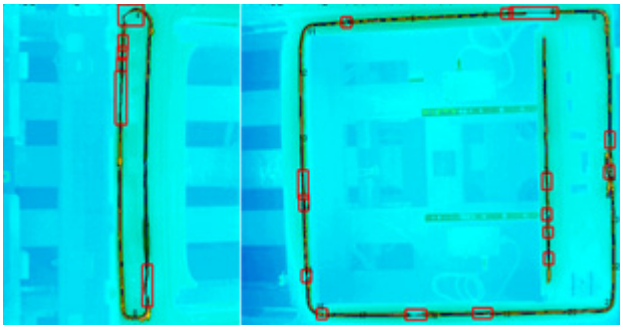
도출된 도포 영역의 시작점 t(x,y)로부터 좌6 픽셀과 우6 픽셀의 온도 값을 추출한다. 이때 실제 온도 값과 시작점으로 부터 2블럭의 온도 값을 추출한다. [Figure 7]에서 도포된 영역에서 온도 값을 추출한 예시를 보여주고 있다. 한 제품에 검사해야 할 검사 포인트는 약 3,300곳이며, <Table 1>과 같은 방법으로 각 포인트 마다 온도 백터를 만든다.



[Figure 7] Example of temperature extraction from the sealed area

3.3 결함특성수치 샘플링

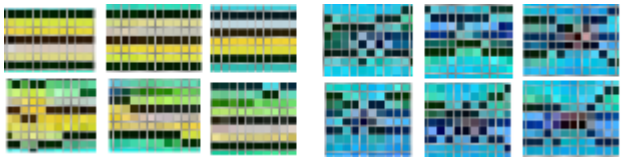
제조 공정상 비결함 온도 정보 샘플은 쉽게 구할 수 있지만 결함 온도 정보의 경우 비결함 샘플보다 상대적으로 적어서 분류기로 학습하는데 있어 어려움이 있다. 본 연구에서는 결함 온도 정보를 샘플링하기 위하여 임의적으로 자동화 로봇을 티칭하여 [Figure 8]과 같이 발생시켰다. 끊김, 미세끊김, 폭기준 미달, 폭기준 초과 샘플을 한 공정에 여러 곳에 발생시켰다. 다수 샘플을 만들어 검사영역 기준으로 비결함 샘플 500, 결함 샘플 500 곳을 [Figure 9]의 예시와 같이 추출하였다. 적은 샘플수의 환경에서 상대적으로 좋은 성능을 보이는 SVM 분류기를 사용하고자 한다.



[Figure 8] Defect sample (the area marked in red is where the defect occurred)

<Table 1> Temperature vector table by inspection point

t(x - 6, y+2)	t(x - 5, y+2)	t(x - 4, y+2)	t(x - 3, y+2)	t(x - 2, y+2)	t(x - 1, y+2)	t(x,y+ 2+2)	t(x+1, y+2)	t(x+2, y+2)	t(x+3, y+2)	t(x+4, y+2)	t(x+5, y+2)	t(x+6, y+2)
t(x - 6, y+1)	t(x - 5, y+1)	t(x - 4, y+1)	t(x - 3, y+1)	t(x - 2, y+1)	t(x - 1, y+1)	t(x,y+ 1+1)	t(x+1, y+1)	t(x+2, y+1)	t(x+3, y+1)	t(x+4, y+1)	t(x+5, y+1)	t(x+6, y+1)
t(x - 6,y)	t(x - 5,y)	t(x - 4,y)	t(x - 3,y)	t(x - 2,y)	t(x - 1,y)	t(x,y)	t(x+1,y)	t(x+2,y)	t(x+3,y)	t(x+4,y)	t(x+5,y)	t(x+6,y)



[Figure 9] Non-defective sample(right) and defective sample(left)

3.4 SVM 학습

두 클래스를 선형 분리할 수 있는 결정평면에 가장 가까이 존재하는 각 클래스의 샘플을 SV(Support Vector)라고 하며, SVM은 결정평면과 SV사이의 최대가 되는 마진을 결정하는 이진 분류기이다. 결정평면은 $wx+b=0$으로 표시된다. W는 가중치 백터를 의미하고 b는 바이어스를 뜻한다. 두 클래스의 SV사이의 분리 마진은 $2/\|w\|$가 된다. 이를 최대화 하기 위해서는 $\|w\|$를 최소화 하는 결정평면은 아래 식과 같다[5][7].

$$\Phi(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 \tag{2}$$

두 클래스의 샘플이 선형적으로 완전 분리가 어려운 경우에는 식 (3)과 같이 오류에 대한 허용 변수를 적용한 비용함수를 통해 최적화가 가능하다[4][6].

$$\Phi(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i \tag{3}$$

분류하고자 하는 두 클래스가 비선형 분리면을 가지고 있게 때문에 커널 함수 K(xi,x)를 적용하여 더 높은 차원의 특징공간에서 hyper plane을 찾음으로써 선형적으로 분리 할 수 있게 된다. 커널함수를 이용할 경우 최종적으로 얻어 지는 SVM 분류기는 식 (4)와 같이 표현되며, 이를 이용 하여 두 개의 비선형분리면을 갖는 클래스를 분류할 수 있다[1][4].

$$f(\mathbf{x}) = sign\left(\sum_{i=1}^{N}\alpha_{o,i}d_iK(\mathbf{x}_i,\mathbf{x}) + b_o\right) \tag{4}$$

비선형 특성을 갖는 분류 식별에 있어서 RBF 커널함수의 식별 성능이 polynomial 커널 함수의 식별 성능 보다 우수한 기존 연구결과를 바탕으로 본 연구에서도 입력 벡터를 고차원의 특징 벡터로 전환하여 내적을 구하기 위해 RBF 커널을 사용하였으며, 그 식은 아래와 같다[3][8].

$$\exp\left(-\frac{1}{2\sigma^2}\|\mathbf{x}-\mathbf{x}_i\|^2\right) \tag{5}$$

결함이 있는 영역 500개와 비결함이 있는 영역 500개의 데이터를 샘플링 하여 두 집합의 비율을 맞추고 SVM 학습을 실시하였고 분류 한 결과를 4장에서 제시한다.

4. SVM기반 품질검사 평가 방법 및 결과

4.1 알고리즘 평가 방법 및 과정

알고리즘 정확도를 평가하기 위하여 민감도(sensitivity)와 특이도(specificity)를 이용하여 평가하였으며, 에러 타입II의 비율도 중요하게 판단하여 중대결함비율(critical error rate)을 아래의 수식으로 정의하였다[9].

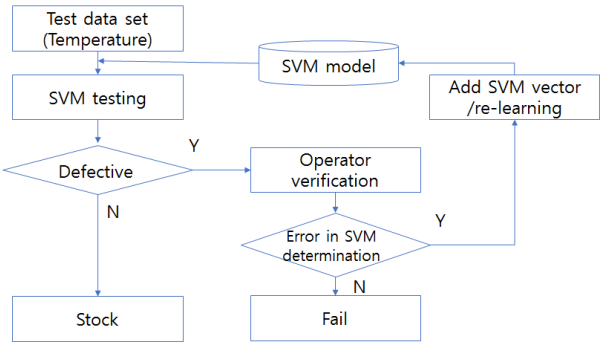
$$sensitivity = \frac{TP}{TP+FN} \tag{6}$$

$$specificity = \frac{TN}{TN+FP} \tag{7}$$

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{8}$$

$$critical\ error\ rate = \frac{FN}{TP+FP+FN+TN} \tag{9}$$

위 식에서 TP(True Positive)는 실제 결함을 결함이라고 판정한 개수를 나타내며, FN(False Negative)은 에러 타입 II 로 실제 결함을 비결함으로 판단한 분류 개수를 의미한다. TN(True Negative)는 실제 비결함을 비결함으로 판정한 개수를 말하고, FP(False Positive)는 에러 타입 I로 실제 비결함을 결함으로 분류한 개수를 말한다. 본 연구에서는 에러 타입 II 가 발생하였을 때 가장 큰 결함으로 평가되기에 중대결함비율(critical error rate)를 추가 판단하는 기준으로 삼았다. 초기 학습 모델을 가지고 현장에서 품질 검사를 하여 FN 또는 FP가 나올 때 해당 벡터를 추가하여 재학습 하는 단계를 [Figure 10]과 같이 진행하였다.



[Figure 10] SVM learning and quality inspection process

ISSN 1229-6783(Print)
ISSN 2288-1484(Online)

4.2 알고리즘 평가 결과

초기 학습한 모델로 결함 검출 성능을 <Table 2>에 정리하였다. 초기 학습 모델이 인위적인 결함 정보로 학습한 결과 임에도 불구하고 <Table 3>에서 나타내는 것과 같이 정확도가 97.52%로 결함과 비결함을 판별해 낼 수 있었다. 사전 결함의 유형을 잘 파악하여 결함과 비결함의 샘플링이 잘 되었음을 알 수 있다. 그러나 문제는 에러 타입 II (FN)이 4건 정도 발생하면서 0.31%의 critical error rate가 발생하였다. 이러한 불일치 판정의 벡터를 지속적으로 모델에 자동 반영되도록 하여 실제 생산과정에서 얻은 벡터들로 학습 모델을 업그레이드 하였다.

6개월 후 학습에 사용된 벡터 수는 9,618개로 초기 1,000개보다 9배 늘어났고, 학습은 초기 모델과 동일하게 RBF 커널로 재 학습 하였다. 업그레이드 된 모델의 결함 검출 성능은 <Table 4>와 같고, 업그레이드 된 모델의 재평가한 결과는 <Table 5>에서와 같이 정확도는 초기모델보다 1.43% 감소한 것으로 나타났으나, sensitivity가 99.67% 에서 100%로 증가하였다. 품질에서 절대로 발생하지 않아야 하는 에러 타입 II(FN)가 발생하지 않았음을 확인할 수 있다.

<Table 2> Test results of the initial model

Judgment category	Count
TP	1,237
TN	29
FP(Error Type I)	25
FN(Error Type II)	4
Total	1,295

<Table 3> Algorithm evaluation results of the initial model

Sensitivity	99.67%
Specificity	53.70%
Accuracy	97.76%
Critical error rate	0.31%

<Table 4> Test results of upgraded model

Judgment category	Count
TP	1,186
TN	47
FP(Error Type I)	47
FN(Error Type II)	0
Total	1,280

<Table 5> Algorithm evaluation result of the upgraded model

Sensitivity	100.00%
Specificity	50.00%
Accuracy	96.33%
Critical error rate	0.00%

[Figure 11]에서는 실제 공정 현장에서 결함을 검출한 사례 이미지이다. TL 차종 Front 영역에서 실러 도포 시 끊김이 발생했으며, 본 논문에서 제안한 결함 검출 방법으로 검출한 첫 사례이다.

[Figure 11] Sealer application defect detected at the production site(Defect inspection(left), defect system image(right))

인위적 결함을 통해 만들어진 초기 모델에 사용된 벡터 수는 총 1,000 건에 불과 했으나, 6개월 후 생산 현장에서 수집한 결과를 모두 적용한 결과, 벡터 수는 9,618건으로 증가하였고, 초기 모델 보다 정교한 판단 능력을 보유하게 되었음을 알 수 있다. 그러나 벡터들의 수가 증가함에 따라서 SVM 학습 시간도 증가하였다. 1,000 건일 때 학습 시간은 약 2초에 불과 했으나, 9,618건 학습하는데 소요되는 학습시간은 약 6,420초로 약 1시간 47분 정도 소요되는 문제가 있었다. 그러나 이것은 데이터 학습으로 최적의 SVM 모델을 생성하는데 증가되는 시간이기 때문에, 실질적으로 생성된 모델을 적용한 시스템의 운영상에서는 문제가 발생하지 않는다. 다만 발생할 수 있는 문제점이라고 하면 모델을 생성하는 주기가 느려지기 때문에 실시간 모델 갱신은 어렵고, 스케줄러를 통해 하루 단위로 모델을 재생성하고 시스템에 반영하여야 한다는 것이다. 즉, 학습 시간이 증가한다고 해서 실제 시스템 운영상의 결함 판단 시간이 증가하는 것은 아니므로, 실시간 결함 판단에는 아무런 문제가 발생하지 않고 실시간 검사가 수행된다.

5. 맺음말 및 향후 연구방향

본 논문에서는 선루프 제조공정 가운데 실러도포 공정상에서 발생하는 결함을 적외선 열화상 카메라로 수집한 온도데이터로 결함을 찾는 알고리즘을 제안하였다. 본 논

문의 특징은 실러가 도포된 영역을 찾는 방법과 해당 영역에서 온도 벡터를 추출하여 분류기에 학습 적용하는 것이다. 또한 시간에 따라 자동 보정한 벡터들은 벡터의 양이 증가함에 따라 sensitivity가 증가하여 품질에서 발생하면 안 되는 에러 타입 II가 개선되는 것을 확인할 수 있었다.

향후 연구방향으로는 본 연구에서는 SVM 알고리즘만을 적용하였는데, 향후에 인공신경망(Artificial neural network), 랜덤 포리스트(Random forest), 그리고 로지스틱 회귀(Logistic regression) 모델 등 다양한 기계학습 알고리즘 등을 적용하여 비교 분석할 필요성과 다양한 제조공정의 불량검사에도 적용할 필요성이 제기된다. 또한, 시스템 운영으로 인한 학습 데이터의 증가에 따른 학습시간 증가 문제는 SVM이 최대 마진(margin) 초평면과 분류 작업이 오직 마진(margin) 위에 놓여 있는 서포터 벡터(support vector)들로부터만 영향을 받는다[1]는 이론 기반으로 볼 때 불필요한 벡터들을 최적화할 수 있는 추가 연구가 필요할 것으로 사료된다.

6. References

[1] C. W. Hsu, C. J. Lin(2002), “A comparison of methods for multiclass support vector machines.” Neural Networks, IEEE Transactions on, 13(20):415 - 425.

[2] Markets and Markets(2017), Machine vision market. R&D Special Zone Promotion Foundation. www.innopolis.or.kr

[3] E. M. Park(2013), “Classifying imbalanced data using an SVM ensemble with K - means clustering in semiconductor test process.” Master’s thesis, Sungkyunkwan University, Graduate School.

[4] J. Park, C. Hwang, K. Bae(2013), “Analysis of target classification performances of active sonar returns depending on parameter values of SVM kernel functions.” Journal of the Korea Institute of Information and Communication Engineering, 17(5): 1083 - 1088.

[5] V. Vapnik(1995a), The nature of statistical learning theory. New York: Springer - Verlag.

[6] V. Vapnik(1995b), “Support - vector networks.” Machine Learning, 20(3):273 - 279.

[7] K. Veropoulos, C. Campbell, N. Cristianini(1999), “Controlling the sensitivity of support vector machines.” Proceedings of the International Joint Conference on AI, 55 - 60.

[8] Wikipedia(2015), Support vector machine. https://ko.wikipedia.org/wiki/support_vector_machine (Accessed on 2017.10.20.).

[9] Y. Yang, I. Oh, R. Kang(2019), Data science with R. Hanbit Academy.

저자 소개

양 희 종
울산대학교 자동차선박대학원
메카트로닉스/IT 석사,
울산대학교 경영학 박사.
현재 ㈜아이티엔제이 대표이사로 재직 중
관심분야 : 머신러닝, 딥러닝, 스마트팩토리 솔루션 등
주소 : 울산광역시 동구 화잠 9길 45 102동 902호

장 길 상
울산대학교 산업공학과 학사,
KAIST 산업공학과 석사, KAIST 경영정보공학과 박사
현재 울산대학교 경영정보학과 교수로 재직 중
관심분야 : ERP/SCM/CRM, 시스템개발방법론, 빅데이터 응용, 산업안전보건 등
주소 : 울산광역시 남구 대학로 93, 울산대학교 경영정보학과