

# 분류 알고리즘과 NCA를 활용한 기계학습 기반 구조건전성 모니터링 시스템

## Machine Learning Based Structural Health Monitoring System using Classification and NCA

신창교·권현석·박유림·김천곤\*  
한국과학기술원 항공우주공학과

Changkyo Shin · Hyunseok Kwon · Yurim Park · Chun-Gon Kim\*

Department of Aerospace Engineering, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, 34141, Korea

### [요 약]

본 연구는 복합재 항공기의 비행 데이터를 활용한 기계학습 기반 구조건전성 모니터링 시스템 연구의 예비 연구이다. 본 연구에서는 구조건전성 모니터링에 이용되기에 가장 적합한 기계학습 알고리즘을 선별하고, 실 기체 데이터에 대한 적용을 위해 차원 축소를 수행하였다. 이를 위해 외팔보를 통해 모사된 항공기 날개 구조와 부가 질량을 통해 손상 모사 실험을 진행하고, 분류 알고리즘을 통해 데이터를 손상의 위치와 정도에 따라 구분하였다. 이를 위해 FBG (fiber bragg grating) 센서를 부착한 외팔보의 진동 실험을 통해 정상상태와 12개의 손상상태에 대한 데이터를 취득하고, MATLAB 환경에서 tree, discriminant, SVM (support vector machine), kNN, ensemble 알고리즘의 비교와 파라미터 튜닝을 통해 가장 적합한 알고리즘을 도출하였다. 또한 NCA (neighborhood component analysis)를 이용한 특징 선택을 통해, 실 기체에서 나올 수 있는 고차원 데이터의 관리를 위해 필요한 차원 축소를 수행하였다. 그 결과, quadratic SVM이 NCA를 적용하지 않은 모델에서 98.7%, NCA를 적용한 모델에서 95.9%로 가장 높은 정답률을 보였다. 또한 NCA 적용 후 모델의 예측 속도, 학습 시간, 용량이 모두 향상되었다.

### [Abstract]

This is a pilot study of machine learning based structural health monitoring system using flight data of composite aircraft. In this study, the most suitable machine learning algorithm for structural health monitoring was selected and dimensionality reduction method for application on the actual flight data was conducted. For these tasks, impact test on the cantilever beam with added mass, which is the simulation of damage in the aircraft wing structure was conducted and classification model for damage states (damage location and level) was trained. Through vibration test of cantilever beam with fiber bragg grating (FBG) sensor, data of normal and 12 damaged states were acquired, and the most suitable algorithm was selected through comparison between algorithms like tree, discriminant, support vector machine (SVM), kNN, ensemble. Besides, through neighborhood component analysis (NCA) feature selection, dimensionality reduction which is necessary to deal with high dimensional flight data was conducted. As a result, quadratic SVMs performed best with 98.7% for without NCA and 95.9% for with NCA. It is also shown that the application of NCA improved prediction speed, training time, and model memory.

**Key word** : Classification, Dimensionality reduction, FBG sensor, Machine learning, Structural health monitoring.

<https://doi.org/10.12673/jant.2019.23.1.84>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 31 January 2019; Revised 7 February 2019  
Accepted (Publication) 27 February 2019 (28 February 2019)

\*Corresponding Author; Chun-gon Kim

Tel: +82-42-350-3719  
E-mail: cgkim@kaist.edu

## 1. 서론

복합재료는 일반적인 금속 재료에 비해 가벼운 무게와 높은 기계적 물성을 가진다. 때문에 기체의 무게가 그 성능과 직결되는 항공우주 분야에서 기체의 성능을 높이기 위해 널리 사용된다. 실제로 Boeing B787 Dreamliner의 경우, 기체 대부분에 탄소 복합재가 사용되었으며, 무게 기준으로 볼 때 50%에 달하는 사용률을 보인다. 복합재는 기존의 재료들과 다른 성질을 가지기 때문에, 항공우주 구조물의 복합재 사용이 커질수록 복합재 항공우주 구조물에 대한 구조건전성 모니터링 기술의 필요성도 커지고 있다. 한편 기계학습은 센서에서 나오는 방대한 데이터를 관리하고 활용하는데 이점을 가진다. 이러한 이점으로 인해, 이전에도 다양한 센서들과 그 데이터들을 활용한 기계학습 기반 구조건전성 모니터링 시스템에 대한 연구가 진행되어 왔다.

Panopoulou et al. [1]은 FBG 센서를 이용하여 dynamic strain을 측정하고, 인공신경망(ANN; artificial neural network)을 이용하여 손상의 위치와 정도에 대한 분류를 수행하였다. Worden et al. [2]은 항공기 날개 패널에 Lead zirconate titanate (PZT) 가속도계 센서를 통해 excitation transmissibility를 측정하고, SVM을 통해 손상 정도와 위치에 대한 분류를 수행하였다. Roy et al. [3]은 PZT transducer를 이용한 ultrasonic guided wave 측정과 ANN를 통한 손상 감지와 위치 파악을 수행하였다. Azzam et al. [4]은 ANN을 이용하여 비행 파라미터를 통한 간접적인 구조의 피로 예측을 수행하였다. Mujica et al. [5]은 자기조직화 지도(SOM; self-organizing map)를 활용하여 CBR(case based reasoning) 시스템을 구축하고, 이를 통한 손상의 위치 파악을 수행하였다. Holmes et al.은 GP(Gaussian process) regression과 비행 파라미터를 활용하여 착륙 과정의 하중을 간접적으로 추정하였고, 이후 후속 연구에서 주성분 분석(PCA; principal component analysis)과 neural network autoencoder를 활용한 dimensionality reduction과 SVM을 이용한 착륙의 특징에 따른 분류를 수행하였다 [6, 7].

이처럼 센서에서 얻은 데이터로 단순히 하중, 변형률 등만 파악하는 것이 아니라, SOM, GP, ANN과 같은 과정을 거쳐 기체의 손상 여부, 손상의 정도, 위치 등 더 의미 있는 정보를 파악하는 연구들이 진행되었다. 하지만 기존의 연구들은 랩 스케일의 시뮬레이션에 치중되어 있어 실제 기체에서 나오는 데이터에는 적용되기 힘든 한계가 있다. 따라서 본 연구는 실 기체에 적용되는 광섬유 센서와 기계학습 기반의 구조건전성 모니터링 시스템 개발 연구의 일환으로, 기계학습을 통한 구조건전성 모니터링 데이터의 활용 방안에 대하여 연구한다. 이를 위하여 항공기의 날개 구조에 발생한 다양한 손상을 외팔보를 통해 간단하게 모사하고, 진동 실험 데이터와 MATLAB 기반의 분류 모델을 이용하여 손상의 유무, 위치, 정도에 따른 분류를 수행하였다. 또한 NCA를 통한 특징 선택으로 차원 축소를 수행하여, 실제 비행체에 적용되었을 때 발생할 수 있는 고차원 데이터에 대한 효율적 관리 시스템을 구축하였다. 궁극적으로는 시

뮬레이션 데이터에 그치지 않고, 비행 중 발생하는 실 기체의 비행 데이터를 이용한 구조 건전성 모니터링에도 사용할 수 있는 통합적 관리 시스템의 구축을 목표로 한다.

## II. 방법론

### 2-1 실험 구성

실험 시편으로는 알루미늄 보의 한쪽을 고정하여 외팔보로 사용하였다. 이는 날개의 하중을 spar가 대부분 지지하고 [8], 이를 간단하게 모사할 수 있는 구조가 외팔보이기 때문이다. 실험 환경은 그림 1과 같이 구성되었으며, 실험 시편의 자세한 규격은 표 1과 같다.

손상의 모사를 위해서 주로 사용되는 방법으로는 saw cut과 부가 질량이 있다. 구조에 손상이 생기면 stiffness 값이 감소하여 고유진동수에 변화가 생기는데, 부가 질량을 이용한 방법은 이를 모사하기 위해 구조물의 stiffness의 변화 대신 질량을 늘려 고유진동수를 변화시킨다. 본 연구에서도 이러한 방법을 사용하여 손상을 모사하였으며, 부가 질량의 위치와 하중을 준 위치는 그림 2와 같다.

부가 질량의 경우 보의 질량의 0.86%에 해당하는 1g을 구분 단위로 하여 1g, 2g, 3g의 질량이 추가된 경우에 대하여 실험하였다. 이전의 부가 질량을 사용한 연구인 Panopoulou et al.에서 1.6kg의 시편에 9g 동전을 기준 질량으로 사용하였는데(시편 질량의 0.56%), 이를 참고로 부가 질량의 최소 단위를 설정하였다 [1]. 부가 질량으로는 분동을 각 위치에 고정하여 사용하였다. 부가 질량의 위치는 그림 2와 같이 네 지점으로 하였는데, 1번과 4번은 wing root, tip에 각각 손상이 생긴 것을 모사하기 위함이며, 좀 더 자세한 위치 파악을 위해 추가로 1번, 4번 위치에서 보의 중심방향으로 50mm만큼 떨어진 지점을 부가 질량 위치로 설정하였다(위치 2번, 3번). 이러한 위치에서 세 가지 부가 질량으로 모사된 손상상태와, 부가 질량이 없는 정상상태, 총 13개의 경우에 대하여 실험과 레이블링을 진행하였다. 각 경우에 부여된 데이터 레이블은 표 2와 같다.

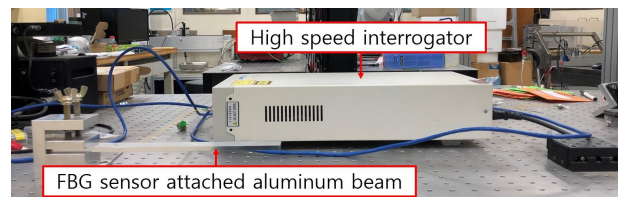


그림 1. 실험 구성  
Fig. 1. Experimental setup.

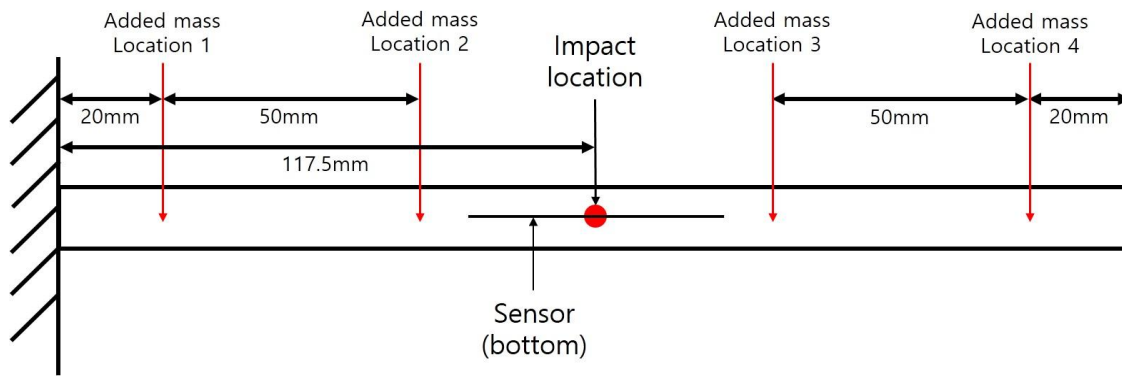


그림 2. 부가 질량 및 하중 위치 (위에서 본 모습)  
 Fig. 2. Added mass and impact location (top view).

표 1. 시편 정보  
 Table 1. Specimen size.

length (mm)	width (mm)	thickness (mm)	mass (g)
282	25	6	116.01

데이터 취득을 위한 센서로는 FBG KOREA Inc. 에서 제작된 FBG 센서를 사용하였으며, 인터로게이터로는 ‘SFI-710M Multi-Channel High Speed FBG Sensing Interrogator’를 사용하였다. 하중 부여 과정은 충격 망치를 이용하여 그림 2에서 표시된 root에서 117.5 mm 떨어진 보의 중심 지점에 일정한 하중을 가하는 것으로 수행되었다. 샘플링 주파수는 100 kHz로 설정되었으며, triggered mode의 하중 감지 경계 값을  $100\mu\epsilon$  으로 설정하여 하중 이전의 500개, 이후의 10000개의 변형률 데이터를 시간에 따라 수집하였다. 손상의 정도와 위치에 따라 총 13개의 경우로 분류하여 각 경우마다 100개의 신호 데이터를 수집하였으며, 총 1300개의 데이터를 통해 기계학습을 수행하였다.

표 2. 데이터 레이블  
 Table 2. Data label.

Data label	mass (g)	location
Normal	-	-
Damage1-1	1	1
Damage1-2	2	1
Damage1-3	3	1
Damage2-1	1	2
Damage2-2	2	2
Damage2-3	3	2
Damage3-1	1	3
Damage3-2	2	3
Damage3-3	3	3
Damage4-1	1	4
Damage4-2	2	4
Damage4-3	3	4

### 2-2 기계학습 알고리즘

MATLAB 작업환경을 기반으로 취득된 데이터의 분석을 수행하였으며, 분류학습기를 이용하여 기계학습과 알고리즘간 비교를 수행하였다. 이 과정에서 데이터의 분석 및 특징 추출을 위해 ‘Signal processing toolbox’와 ‘Wavelet toolbox’를 이용하였다. 기계학습 과정에서는 우선 인터로게이터를 통해 수집된 진동 신호 데이터에 레이블을 부여하여 reference table을 작성하였다. 이후 특징 추출 단계에서 통계적 특징, 고속 푸리에 변환(FFT; fast Fourier transform), 전력 스펙트럼 밀도(PSD; power spectrum density), 이산 웨이블릿 변환(DWT; discrete wavelet transform)에서 얻을 수 있는 특징들을 포함하여 총 20개의 특징을 사용하였다. 추출된 특징들과 사전에 작성된 reference table에서 불러온 class label로 feature table을 작성하고 이를 기반으로 분류 모델 학습을 수행하였다. 데이터의 양이 많지 않기 때문에, 성능 검증 과정에는 5-folds cross validation이 이용되었다. Tree, discriminant, SVM, kNN, ensemble 등의 총 22개 알고리즘의 학습과 성능 비교가 수행되었으며, kunnel function, k-value, multi-class method 등의 파라미터 튜닝을 통해 최적 알고리즘을 도출하였다. 이후 차원 축소를 위해 NCA를 이용한 특징 선택을 수행하였다. NCA 과정에서는 분류의 정확도를 최댓값으로 만드는 특징 가중치 행렬을 계산하고, 가중치가 0.01을 넘는 특징들을 선택하였다. 기계학습의 대략적인 알고리즘은 그림 3과 같으며, NCA를 통한 특징 선택 결과는 표 3과 같다.

### III. 결과 및 고찰

앞서 언급하였듯, 1300개의 데이터를 바탕으로 22개 알고리즘으로 모델을 학습시키고, 각 모델에 NCA를 적용시켰을 때 전체 모델의 정확도는 표 4와 같다.

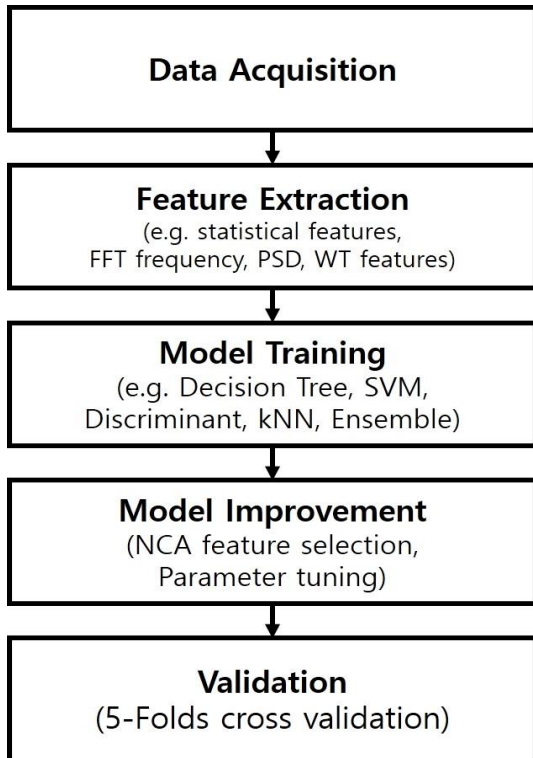


그림 3. 기계학습 알고리즘  
Fig. 3. Machine learning algorithm.

표 3. NCA 특징 선택 결과  
Table 3. NCA feature selection result.

	Before NCA	After NCA
Basic statistical features	Mean, median, standard deviation, mean absolute deviation, quartile25, quartile75, Interquartile Range(IQR), skewness, kurtosis, signal entropy	Mean, median, quartile25, quartile75, IQR
FFT features	Dominant frequency value, magnitude, power ratio	-
PSD features	PSD standard deviation, skewness, kurtosis	PSD standard deviation, kurtosis
DWT features	Energy of A1, D1, D2, D3	Energy of D1, D2, D3

표 4에서 NCA를 적용하면서 subspace kNN을 제외하면 모두 정확도가 떨어지는 모습을 확인할 수 있다. Tree 모델의 경우, 사전 실험에서 class가 2개였을 때(정상, 비정상) 손상의 발생 여부 판단)나 class의 숫자가 적을 때는 높은 정답률을 보여주었으나, 실제 실험에 들어와 class의 수가 늘어나면서 급격하게 성능이 저하되는 모습을 보여주었다. 표 4에서 simple tree에서 complex tree로 갈수록 정답률이 상승하는데, 이는 simple tree로는 13개 class에 대한 분류를 충분히 수행할 수 없고 complex tree 수준의 자세한 분류 기준이 필요하며, complex

표 4. 전체 알고리즘 정확도 표  
Table 4. Classification accuracy.

		Without NCA (%)	With NCA (%)
Tree	Complex	83.2	75.9
	Medium	70.3	64.5
	Simple	38.0	30.0
Discriminant	Linear	98.5	91.7
	Quadratic	Failed	Failed
SVM	Linear	98.4	93.9
	Quadratic	<b>98.7</b>	<b>95.9</b>
	Cubic	98.3	94.7
	Fine	82.1	77.0
	Medium	94.9	91.5
	Coarse	88.5	80.8
kNN	Fine	89.6	77.8
	Medium	84.4	74.8
	Coarse	73.5	58.7
	Cosine	85.3	72.5
	Cubic	85.5	74.6
	Weighted	88.6	78.6
Ensemble	Boosted Trees	80.5	74.1
	Bagged Trees	92.8	86.8
	Subspace Discriminant	96.5	87.8
	Subspace kNN	14.6	62.5
	RUSBoosted Trees	68.2	63.3

tree의 문제점인 overfitting은 주요하게 작용하지 않음을 의미한다. 단, tree 모델 중 가장 높은 정답률을 보여준 complex tree도 충분한 성능을 보여주지는 못하였다. Discriminant 모델은 다중 클래스 분류에서 분류에 실패하는 모습을 보여주었는데, 이는 discriminant 알고리즘이 다중 클래스 분류에 적합하지 않기 때문으로 보인다. SVM은 전반적으로 가장 우수한 성능을 보여주었으며, kernel function으로 quadratic function을 사용하고 kernel scale을 1로, multi-class method를 one-vs-one으로 설정하였을 때 가장 우수한 성능을 보여주었다. 단, SVM의 경우 linear SVM을 제외하면 결과 해석이 어렵고, 이산 분류가 아닌 다중 클래스 분류에서 메모리 사용량과 분류 속도 성능이 좋지 않으므로, 이에 대한 보완이 필요할 것으로 보인다. kNN의 경우 tree와 마찬가지로 class의 수가 늘어나면서 급격하게 성능이 저하되었으며, ensemble 모델의 경우 모두 충분하지 못한 성능을 보여주었다. 이는 kNN 알고리즘이 데이터의 차원이 커지면 정확도가 떨어지는 문제점으로 인한 것인데, 본 연구의 데이터는 20개의 특징 (NCA의 경우 10개)을 가지는 고차원 데이터이기 때문에 적합하지 않다 [9].

다음으로 NCA를 통해 모델이 얼마나 가벼워지고 빨라졌는지를 알아보기 위해서, 가장 우수한 성능을 보여준 quadratic

SVM에 대하여 NCA 전후의 비교를 수행하였다. 그 결과 표 5에서처럼 정확도는 98.7%에서 95.9%로 소폭 하락하였지만, 분류 속도, 학습 시간, 모델 용량에서 모두 유의미한 성능 향상을 보여주었다.

다음으로, 실험을 통해 학습한 분류 모델과 NCA를 적용한 quadratic SVM 모델의 각 class에 대한 분류 혼동 행렬은 그림 4, 5와 같다.

모델의 혼동 행렬에서 알 수 있듯, 같은 질량이 다른 위치에 있는 것은 잘 분류해 내지만, 다른 질량이 같은 위치에 있는 경우에 혼동이 발생하였다. 이는 구조에 발생한 손상의 위치는 잘 판단할 수 있지만, 손상의 정도를 판단하는데 있어 어려움이 있음을 의미한다. 때문에 손상의 정도 판단 알고리즘의 성능 향상이 필요하며, 이를 위해 분류 대신 회귀 알고리즘 또는 ANN 등의 기법을 사용하는 것이 가능할 것으로 보인다.

표 5. Quadratic SVM 모델의 NCA 적용 결과  
Table 5. NCA result of quadratic SVM.

	Accuracy (%)	Classification time (data/sec)	Training time (sec)	Memory usage (KB)
Without NCA	98.7	~1500	15.119	1525
With NCA	95.9	~3100	9.2376	892

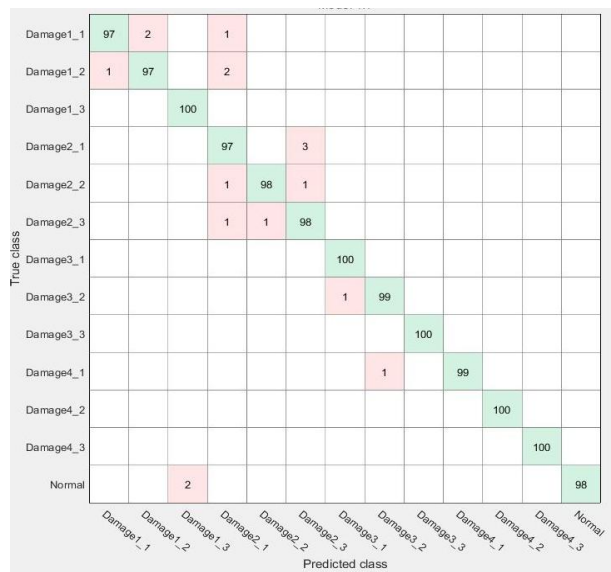


그림 4. NCA 미적용 모델 혼동 행렬  
Fig. 4. Confusion matrix of quadratic SVM (non-NCA).

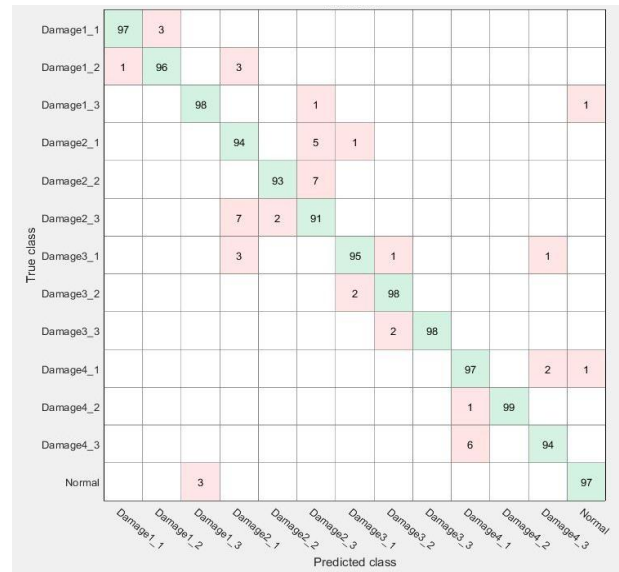


그림 5. NCA 적용 모델 혼동 행렬  
Fig. 5. Confusion matrix of quadratic SVM (NCA).

#### IV. 결론

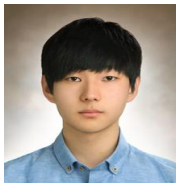
본 연구는 복합재 항공기에서 나오는 데이터의 관리 및 활용과 기계학습을 통한 구조 건전성 모니터링을 위한 연구의 일환으로, 외팔보에 모사된 손상에 대한 분류를 수행하였다. 이를 위해 손상 위치와 정도에 대한 실험을 수행하고 FBG 센서를 통해 데이터를 수집하였으며, 분류 알고리즘의 구성 및 NCA를 통한 모델 압축을 수행하였다. 그 결과 quadratic SVM이 가장 우수한 성능을 보여주었으며, NCA를 통해 특징을 선택하고, 압축된 모델을 사용하여도 95% 이상의 정답률을 얻을 수 있었다. 또한 NCA를 적용한 모델에서 예측 시간, 학습 시간, 모델의 크기 성능이 모두 향상되어, quadratic SVM이 다중 클래스 분류에서 가지는 모델의 큰 메모리 사용이나 긴 예측 시간의 문제점을 NCA를 통해 개선할 수 있을 것으로 보인다. 이는 모델이 실기체에 적용될 때, 고차원 데이터를 다루는 데 있어서 도움이 될 것이다. 또한 혼동 행렬을 분석한 결과, 손상의 위치보다 정도 파악에서 오류가 발생함을 확인하였다. 때문에 손상 정도 파악 부분에서 성능 향상이 필요하며, 이를 위해 분류 알고리즘 대신 회귀 알고리즘이나 ANN을 사용하여 정확도를 향상시키는 것이 필요할 것으로 보인다. 실험 시편 또한 실제 복합재 항공기 구조에 좀 더 가까운 복합재 외팔보, 평판으로 교체하여, 복합재에서 발생하는 층간 박리에 대한 분류 연구를 수행할 예정이다. 궁극적으로는 본 연구의 목표인 실기체 데이터를 활용한 구조건전성 모니터링 시스템 구성을 위해서 고차원의 다량의 실제 비행 데이터에 대한 검증 과정이 필요할 것이다.

## Acknowledgement

이 연구는 2018년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임(‘20002741’)

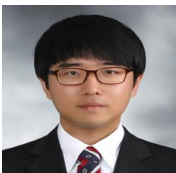
## References

- [1] A. Panopoulou, D. Roulias, T. H. Loutas and V. Kostopoulos, “Health monitoring of aerospace structures using fibre Bragg gratings combined with advanced signal processing and pattern recognition techniques,” *Strain*, Vol. 48, No 3, pp. 267-277, May. 2012.
- [2] K. Worden and G. Manson, “The application of machine learning to structural health monitoring,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 365, No. 1851, pp. 515-537, December. 2006.
- [3] S. Roy, FK Chang, S. J Lee, P. Pollock, V. Janapati, “A novel machine-learning approach for structural state identification using ultrasonic guided waves,” in *Proceedings of the 11<sup>th</sup> International Conference of Structural Safety and Reliability*, New York, NY, pp.321-328. June. 2013.
- [4] H. Azzam, “A practical approach for the indirect prediction of structural fatigue from measured flight parameters,” in *Proceedings of the Aerotech 95*, Birmingham, October. 1995.
- [5] L. E. Mujica, J. Vehi, W. Staszewski and K. Worden, “Impact damage detection in aircraft composites using knowledge-based reasoning,” *Structural Health Monitoring*, Vol. 7, No. 3, pp. 215-230, September. 2008.
- [6] G. Holmes, P. Sartor, S. Reed, P. Southern, K. Worden and E. Cross, “Prediction of landing gear loads using machine learning techniques,” *Structural Health Monitoring*, Vol. 15, No. 5, pp. 568-582, September. 2016.
- [7] G. Holmes, A. Thomas, W. Carpener, K. Worden, E. J. Cross, “Data-based classifiers for identification of aircraft landing characteristics,” in *Proceeding of the 8<sup>th</sup> European Workshop on Structural Health Monitoring*, Bilbao: Spain, July 2016.
- [8] P. Garre and G. V. Arjun, “Modeling and analysis of a ribs and spars of an airplane wing for bending and shear loads,” *International Journal For Research in Applied Science and Engineering Technology*, Vol. 5, No. 2, pp. 295-315, February. 2017.
- [9] Mathworks. Choose classifier options. [Internet] Available: <https://kr.mathworks.com/help/stats/choose-a-classifier.html>



**신 창 교 (Changkyo Shin)**

2015년 3월 ~ 현재 : 한국과학기술원 항공우주공학과 학사과정  
 ※관심분야 : 구조 건전성 모니터링, 기계학습



**권 현 석 (Hyunseok Kwon)**

2017년 2월 : 한국과학기술원 항공우주공학과 (공학석사)  
 2017년 3월 ~ 현재 : 한국과학기술원 항공우주공학과 박사과정  
 ※관심분야 : 스마트 센서를 이용한 구조 건전성 감시 시스템



**박 유 림 (Yurim Park)**

2011년 2월 : 한국과학기술원 항공우주공학과 (공학사)  
 2013년 8월 : 한국과학기술원 항공우주공학과 (공학석사)  
 2013년 9월 ~ 현재 : 한국과학기술원 항공우주공학과 박사과정  
 ※관심분야 : 광섬유 센서, 구조 건전성 감시, 초고속 충격



**김 천 곤 (Chun-Gon Kim)**

1987년 ~ 1991년 : 한국기계연구원 선임연구원  
 1991년 ~ 현재 : 한국과학기술원 항공우주공학과 교수  
 ※관심분야 : 스마트 복합재, 스텔스 구조, 광섬유 센서를 이용한 스마트 구조 설계