



AMERICAN UNIVERSITY OF BEIRUT

CROWD ATTRIBUTES ESTIMATION USING SUPPORT  
VECTOR MACHINE AND DEEP LEARNING WITH MULTI-  
SOURCE SENSOR FUSION

by

KARIM MOHAMAD ALI HASSOUN

A thesis defense  
submitted in partial fulfillment of the requirements  
for the degree of Master of Engineering  
to the Department of Mechanical Engineering  
of the Maroun Semaan Faculty of Engineering and Architecture  
at the American University of Beirut

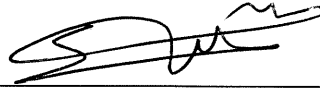
Beirut, Lebanon  
April 2019

AMERICAN UNIVERSITY OF BEIRUT

CROWD ATTRIBUTES ESTIMATION USING SUPPORT VECTOR MACHINE AND DEEP LEARNING WITH MULTI-SOURCE SENSOR FUSION

by  
KARIM MOHAMAD ALI HASSOUN

Approved by:



---

Dr. Samir Mustapha, Assistant Professor  
Department of Mechanical Engineering  
Maroun Semaan Faculty of Engineering and Architecture  
American University of Beirut

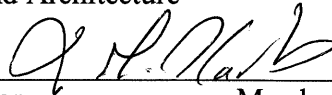
Advisor



---

Dr. Zaher Dawy, Professor  
Department of Electrical and Computer Engineering  
Maroun Semaan Faculty of Engineering and Architecture  
American University of Beirut

Member of Committee



---

Dr. Mohammad S. Harb, Assistant Professor  
Department of Mechanical Engineering  
Maroun Semaan Faculty of Engineering and Architecture  
American University of Beirut

Member of Committee

Date of thesis defense: 30<sup>th</sup> of April, 2019

# AMERICAN UNIVERSITY OF BEIRUT

## THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Hassan Karim Mohamad Ali  
Last First Middle

Master's Thesis       Master's Project       Doctoral Dissertation

I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes after:

~~One~~ --- year from the date of submission of my thesis, dissertation, or project.

Two --- years from the date of submission of my thesis, dissertation, or project.

Three --- years from the date of submission of my thesis, dissertation, or project.

  
Signature

10/05/2019

Date

## ACKNOWLEDGMENT

I would like to express my deep gratitude and appreciation towards my thesis advisor, Dr. Samir Mustapha, for his guidance and belief in me all the way leading to obtaining my degree. His insight, knowledge, encouragement and patience are just a few of the great attributes of him being a wonderful research mentor. Along with the academic lessons, research skills learnt through him will serve as guides for my future path. Simply would like to say, it was an honor being your student.

I wish to acknowledge the valuable assistance and advice given by Dr. Zaher Dawy and Dr. Abdallah Kassir during my research work. It was of great knowledge and pleasure working with them.

To Mohamad Ali Fakhri, much gratefulness in being there throughout my graduate study journey for I've always considered you as my older brother.

I would like to further extend my many thanks and appreciation to all friends who supported and assisted me throughout the research.

We would like to acknowledge the financial support received from General Electric (GE), award # 103311.

To my loving family, no amount of words can describe the unconditional support you have given me. And to again I say, no amount of words can describe how thankful I am.

# AN ABSTRACT OF THE THESIS OF

Karim Mohamad Ali Hassoun for

Master of Engineering

Major: Mechanical Engineering

Title: Crowd Attributes Estimation using Support Vector Machine and Deep Learning with Multi-Source Sensor Fusion

Unfortunate tragedies have previously been the result of high-density human crowds or pedestrian flow. In addition to, crowd behavior as a reaction to an incident aggravates the complexity and disruption of human flow, resulting in possible trampling and crushing situations. Therefore, it is important to monitor such crowd motion for danger warning and prevention. In this study, a frame work was established to provide continuous monitoring and estimation of crowd flow and load on pedestrian bridges, with particular focus on high crowd density enhancing operation safety. A main innovation under sensing instrumentation is the employment of structurally mounted Fiber Bragg Gratings (FBG) Fiber Optic Sensors (FOS), in conjunction with individually held wearable sensing devices incorporating Inertial Measurement Unit (IMU). Furthermore, the approach added innovation under machine learning employment, primarily Convolutional Neural Networks (CNN) along with conventional Support Vector Machine (SVM) algorithms thus generating crowd estimation models from gathered sensors' data. The concept was validated using experimental measurements on two phases based on crowd replication scenarios on a scaled test bridge. Generated machine learning models demonstrated effectiveness in crowd attribute classification for flow activity and load characterization, along with regression model for load estimation. Multi-modal sensor fusion at the input and feature level was further applied on strain and acceleration data collected enriching the machine learning models, thus enhancing system efficiency and robustness against noisy and time shifted input data. The results showed that the monitoring solution to be highly effective with peak testing accuracy for single class flow activity classification at 98%, multi-class flow and load characterization classification at 91%, and percentage error for load estimation regression reaching a minimum of 9%.

**Keywords:** crowd, flow, activity, density, load, weight, management, monitoring, structural health, SHM, sensor, fiber optic, fiber bragg gratings, FBG, strain, smartphones, mobile, accelerometer, fusion, pattern recognition, machine learning, classification, regression, characterization, estimation, multi-class, support vector, SVM, gaussian process, convolution neural network, CNN.

# CONTENTS

ACKNOWLEDGMENT .....	V
ABSTRACT.....	VI
CONTENTS.....	VII
LIST OF ILLUSTRATIONS.....	IX
LIST OF TABLES.....	XII
CHAPTER	
I. INTRODUCTION.....	1
A. Problem Definition.....	1
B. Current Solutions Review .....	4
C. Aim and Objectives.....	11
II. PROPOSED APPROACH AND METHODOLOGY.....	13
A. Pattern Recognition and Machine Learning.....	14
i. Support Vector Machine (SVM) .....	15
ii. Gaussian Process (GP).....	21
iii. Convolutional Neural Network (CNN) .....	22
B. Data Fusion .....	24
C. Fiber Optic Sensors (FOS).....	25
i. Fiber Bragg Grating.....	27
III. EXPERIMENTAL VALIDATION – PHASE I .....	30
A. Test Bridge.....	30
B. Instrumentation .....	32
C. Crowd Replication .....	32
IV. RESULTS AND DISCUSSIONS – PHASE I.....	35
A. Raw Signal .....	35
B. Features .....	37

C. Classification and Regression Models .....	38
<b>V. EXPERIMENTAL VALIDATION – PHASE II .....</b>	<b>41</b>
A. Test Bridge .....	41
B. Instrumentation .....	41
C. Crowd Replication .....	42
D. Raw Signal .....	43
<b>VI. RESULTS AND DISCUSSIONS – PHASE II.....</b>	<b>46</b>
A. Classification and Regression Models .....	46
B. Noise and Time Shift Invariance .....	50
C. Fusion.....	52
i. Sensor number fusion .....	52
ii. Sensor data fusion.....	57
D. SVM Results .....	59
E. CNN Results .....	66
F. Discussion .....	72
i. Single Class – Flow Activity Classification .....	72
ii. Multi Class – Flow Activity and Load Characterization Classification .....	74
iii. Load Estimation Regression .....	75
iv. Discrepancy analysis .....	76
v. Optimal Results .....	77
<b>VII. CONCLUSION AND FUTURE WORKS .....</b>	<b>83</b>
<b>REFERENCES .....</b>	<b>87</b>



# ILLUSTRATIONS

Figure	Page
Figure 1. A simple linear support vector machine – hyperplane separation for classification .....	18
Figure 2. A simple linear support vector machine – fitting for regression.....	20
Figure 3. Data processing stages in a simple convolutional neural network.....	23
Figure 4. FBG structure and working principle in respect to shifts in reflected Bragg wavelength due to strain .....	28
Figure 5. Test-bed bridge to be used for crowd data acquisition (a) technical drawing (cm) - top view, (b) technical drawing (cm) - front view, (c) assembled.....	31
Figure 6. Experiment setup demonstrating how the interrogator is connected to the three daisy-chained FBG strain sensors mounted below the middle C-beam .....	33
Figure 7. Group of eight walking fast on the test bridge .....	34
Figure 8. One person – 74 kg, (top right) fast, (bottom right) slow .....	36
Figure 9. Ten persons – 726 kg (top) fast, (bottom) slow .....	36
Figure 10. Selected extracted features (a) Maximum Strain vs. Mass, (b) Area vs. Mass .....	38
Figure 11. Experiment setup demonstrating a crowd walking along the test bridge. FBG sensors mounted below and connected to the interrogator measure strain. Smartphones held per volunteer sends accelerometer data via WiFi to the wireless access point. Workstation records strain and acceleration data for post-processing .....	42
Figure 12. Strain and acceleration levels for a group of two under (a) ‘Fast’ and (c) ‘Slow’ flow activity. Strain and acceleration levels for a group of eight under (b) ‘Fast’ and (d) ‘Slow’ flow activity.....	45
Figure 13. CNN network architecture applied for flow type and load characterization classification and load estimate regression models .....	49
Figure 14. Strain(a) and acceleration norm (b) between unmodified, noisy, and shifted signal.....	51

Figure 15. SVM single class flow activity classification on number of FBG sensors between middle and all and their corresponding signal type.....	53
Figure 16. CNN single class flow activity classification on number of FBG sensors between middle and all and their corresponding signal type.....	54
Figure 17. SVM multi-class flow activity and load characterization classification on number of FBG sensors between middle and all and their corresponding signal type .....	55
Figure 18. CNN multi-class flow activity and load characterization classification on number of FBG sensors between middle and all and their corresponding signal type .....	55
Figure 19. SVM load estimation regression on number of FBG sensors between middle and all and their corresponding signal type for both flow activities.....	56
Figure 20. CNN load estimation regression on number of FBG sensors between middle and all and their corresponding signal type for both flow activities.....	57
Figure 21. Fusion process at the (a) input level and feature level for (b) SVM and (c) CNN.....	58
Figure 22. SVM using raw signal individual and fusion data – Flow activity classification testing accuracies for unmodified, noisy, shifted, and noisy shift data.....	60
Figure 23. SVM using FFT signal individual and fusion data – Flow activity classification testing accuracies for unmodified, noisy, shifted, and noisy shift data.....	61
Figure 24. SVM using raw signal individual and fusion data – Flow activity and load characterization classification testing accuracies for unmodified, noisy, shifted, and noisy shift data .....	62
Figure 25. SVM using FFT signal individual and fusion data – Flow activity and load characterization classification testing accuracies for unmodified, noisy, shifted, and noisy shift data .....	63
Figure 26. SVM using raw signal individual and fusion data – Load estimation regression testing percentage errors for unmodified, noisy, shifted, and noisy shift data .....	64
Figure 27. SVM using FFT signal individual and fusion data – Load estimation regression testing percentage errors for unmodified, noisy, shifted, and noisy shift data .....	65

Figure 28. CNN using raw signal individual and fusion data – Flow activity classification testing accuracies for unmodified, noisy, shifted, and noisy shift data.....	66
Figure 29. CNN using FFT signal individual and fusion data – Flow activity classification testing accuracies for unmodified, noisy, shifted, and noisy shift data.....	67
Figure 30. CNN using raw signal individual and fusion data – Flow activity and load characterization classification testing accuracies for unmodified, noisy, shifted, and noisy shift data .....	68
Figure 31. CNN using FFT signal individual and fusion data – Flow activity and load characterization classification testing accuracies for unmodified, noisy, shifted, and noisy shift data .....	69
Figure 32. CNN using raw signal individual and fusion data – Load estimation regression testing percentage errors for unmodified, noisy, shifted, and noisy shift data .....	70
Figure 33. CNN using FFT signal individual and fusion data – Load estimation regression testing percentage errors for unmodified, noisy, shifted, and noisy shift data .....	71

## TABLES

Table	Page
Table 1: Volunteer composition used in Phase 1 crowd replication; dataset indicating size of groups, total groups and average mass per group size .....	34
Table 2: Classification and regression model testing results per signal type and machine learning approach .....	39
Table 3: Volunteer composition used in crowd replication; dataset indicating size of groups, total groups and average mass per group size .....	43
Table 4: Optimal testing accuracy results of single class flow activity classification per input data and signal type .....	78
Table 5: Standard deviation of the optimal testing accuracy results of single class flow activity classification per input data and signal type .....	78
Table 6: Optimal testing accuracy results of multi class flow activity and load characterization classification per input data and signal type .....	79
Table 7: Standard deviation of the optimal testing accuracy results of multi class flow activity and load characterization classification per input data and signal type .....	79
Table 8: Optimal testing percentage error results of load estimation regression per input data and signal type for both flow activities; Fast and Slow .....	80
Table 9: Standard deviation of the optimal testing percentage error results of load estimation regression per input data and signal type for both flow activities; Fast and Slow .....	81

# CHAPTER I

## INTRODUCTION

### **A. Problem Definition**

The issue of crowd safety is a matter of growing importance. Pedestrian behavior varies under different scenarios, whether in a normal everyday situation or in events of mass gatherings. Festivals, religious observances, sporting events, concerts and political rallies can be all considered as such where crowd disasters have a large likelihood of occurring. During high density crowds, there is a potential of injury with leading possibility loss of life due to the dynamics of the crowd in said events. It has been argued that crowd behavior is unlikely to be predicted easily for the reason that people are irrational and erratic. Factors such as poor crowd control, high crowd density, and constrained access may all contribute to unfortunate pedestrian disasters.

#### Stampede and related cases

Panic stampedes are one of the most relevant and serious concern during mass events. A type of threat a pedestrian may be subjected to is trampling. In these situations, although the density of the crowd is high, movement of the pedestrians is still possible and any pedestrian that falls may find that they are unable to stand again because of jostling from the motion of other pedestrians. Consequently, fatalities may occur from the percussion by the feet of those standing who are unaware of the fallen pedestrian or by asphyxiation by others tripping and falling on top. Another type of fatality caused by crowding is where pedestrians are crushed. In situations where

pedestrians are crushed, the density of the crowd is extremely high and the physical movement of pedestrians is almost impossible. When crushing occurs, the high pressures developed within the crowd, which has the force capable of bending steel barriers or knocking down brick walls, can be unbearable to some members of the crowd, producing fatalities from asphyxiation while still standing.

An annual Islamic pilgrimage to Mecca, The Hajj, is attended by more than 2.5 million visitors every year [1]. Unfortunately, the Hajj suffers from time to time tragic incidents occurring due to high crowd density movement at the Jamarat bridge [2], resulting in some of the most fatal crowd disasters in history. The notable 2015 Hajj stampede took the lives of 2236 pilgrims as a result from overcapacity and formation of shockwaves at a previously not identified converging path leading to crowd collapse [3]. A similar incident at the pilgrimage on January 12, 2006 via video analysis found trampling of people can be triggered as different crowd dynamics during high crowd densities can cause two sudden flow transition, leading from laminar to stop-and-go flow followed by turbulent crowd motion [4].

Being one of the most populated countries in the world, India has encountered mass gathering casualty cases. On October 15, 2016 at least 24 people died and many got injured in a stampede in Varanasi, India [5]. On July 14, 2015, at least 27 pilgrims died in a stampede on the banks of a Godavari river during Maha Pushkaralu festival in Andhra Pradesh, India [6].

### Structure related crowd disaster cases

Generally speaking, pedestrians themselves walking play a vital role of a transportation system. With urban area booming and developing, pedestrian bridges construction has risen with some matching the levels of road bridges. Today, an increasing number of engineers put an effort in the design stage to accommodate the potential threat of pedestrian induced lateral vibrations [7-10]. As a result of such dynamics and large lateral vibrations, pedestrian bridges around the globe have encountered instability during event of high crowd loading and density [11-14].

One of the most noted incidents of induced pedestrian lateral vibrations occurred at the Toda Park Bridge in Toda City, Japan (T-Bridge) [15]. The large cable-stayed bridge measuring 179m in length, carried 2000 people simultaneously (crowd density of approximately 2.1 pedestrian/m<sup>2</sup>) subsequently causing lateral acceleration of the bridge girder. Though accounting for the resonance step frequency of all pedestrians and mutually independent (random) phases, this was larger than expected [16].

An infrequent yet critical crowd related disaster that may occur during mass gathering events is the collapse of pedestrian bridges due to overloading or aging. In 2010 at Phnom Penh, Cambodia, about 450 casualties resulted from a stampede due to swaying suspension footbridge panic [2]. In 2013 at Ratangarh temple, India, stampede on a bridge resulted in 89 casualties with many more are published [2]. A recent incident on March 2018 occurred in Miami, U.S.A near Florida International University where at least six people lost their lives with more injured after a 53m footbridge collapsed over an eight-lane street [17]. A pedestrian bridge linked to an island in

Prague, Czech Republic collapsed entirely, with reports indicating cause being possible steel ropes corrosion [18].

### Crowd experience influences

The chaotic behavior and panic of a crowd tend to cause more injuries or casualties than the initial cause of the disaster as it has been investigated [2, 19]. An investigation into the 2010 Love Parade crowd incident in Germany via video recordings analysis was undertaken in [20], with the interest to understand the actual cause and distinguish between mass panic and coordination breakdown. Filingeria et al. [21] drew five overarching issues that influence crowd experience from analysis of a focus group and event observation data, being: physical design of crowd space and facilities, public order, crowd movement, communication and information, comfort and welfare. Yet lives are still lost in crowd associated disasters despite number of crowd control measures.

## **B. Current Solutions Review**

Collecting, organizing and analyzing data of crowd attributes during an event execution are the essential part of any crowd management system. Such monitoring operation is vital in order to detect dangerous situations that may result in pedestrian asphyxiation by chest compression or stampeding at very high densities. Hence, undertaking timely decisions in avoiding potential crowd disaster [22].

Studying a crowd behavior can be done by synthesizing derived prediction models to be used as priori event planning in a simulation environment. One of the most popular models used is the social-force model by [23], where the crowd's motion and



behavior is described and measured as a motivation to perform particular movements. Resulting model equations, in terms of motion velocity, spatial distance, and attractive and repulsive effects, all through computer simulations are realistically capable of describing of pedestrian behavior.

#### Vision utilizing machine learning solutions

System effectiveness relies on accurate crowd related information acquisition from the event. Most widely and conventionally used crowd management systems are based on vision technologies, e.g. CCTV. Vision systems can gather crowd related information from camera image and video analysis and generate near real time crowd control measures. Various research directions have been explored in more advanced computer vision-based crowd analysis work. These methods rely on attributes that can describe a crowd's status, including crowd density estimation, face recognition, head, pedestrian and crowd tracking from crowd events [24]. People counting or density estimation can be done via pixel based analysis, texture level analysis, and object level analysis. People tracking or main crowd flow estimation can be done via pedestrian tracking, and flow estimation. Understanding people's behavior can be done via event detection and collective behavior [25].

Such based systems utilize several forms of machine learning algorithms such as Support Vector Machine and Neural Networks for crowd monitoring and estimation applications [26]. Wu et al. [27] proposed an approach to estimate the crowd density using texture analysis, followed by training a Support Vector Machine (SVM) to relate the textural features with the actual density of the scene. Zhang et al. [28], devised a multi-column based Convolutional Neural Network (CNN) that takes images as the

input and outputs a local density maps (people per square meter) via head size and count, thus obtaining a final crowd count. To cope with different head sizes on the same image, caused by perspective distortion, the authors proposed the use of filters with different sizes of local receptive field to learn the map from the raw pixels, integrating into the density maps. Instead of estimating density maps, Sindagi et al. [29] and Fu et al. [30] proposed a crowd count classification model. The proposed model classified crowd images into various density levels, which in turn can correspond to a coarse estimate count in an image. Complexity reduction with increase in performance was achieved by Shang et al. [31] following an end-to-end count estimation method using CNNs, where entire images are used as an input instead of dividing them into patches.

In the context of crowd counting, that means detecting pedestrians (whole bodies or parts, such as heads) in images or video sequences. They tend to produce more accurate results when compared to pixel-level analysis or texture-based approaches, but identifying individuals is mostly feasible in lower density crowds. In denser crowds, clutter and severe occlusions make the individual counting problem almost impossible to solve, despite the recent advances of computer vision and pattern recognition techniques [25].

Systems relying on vision technology are prone to certain limitations and concerns such as vision loss, whether due to night operation, smoke, or fog for example. Privacy, obstruction of field of view and misperception of extracted image data among other may further hinder such system capabilities. Researchers have developed systems that overcome basic vision system limitations whilst improving efficiency [32].

Passive infrared sensors [33] and thermal detector sensor array [34] with machine learning has been researched as non-image-based crowd monitoring techniques. A sensory fusion model of infrared and visual data with Artificial Neural Network (ANN) was researched with the purpose of crowd density mapping and improving crowd density estimation [35] whilst expected reliability when said limitations may occur. A real-time thermal video sequence data acquisition based crowd control and monitoring system where an alarm is triggered according to different density levels was proposed [36]. In addition, analysis of drone or satellite based high resolution aerial images of the area to be monitored have been proposed for crowd density estimation. An unmanned aerial vehicle (UAV) and unmanned ground vehicle (UGV) based joint approach was proposed for effective information gathering for surveillance and crowd control [37].

#### Human activity monitoring

Conventional approaches, as per stated literature, utilize computer vision, relying on image and video analysis with respectable results as technology and post processing advances, however yet still with a few certain disadvantages. It is of interest to employ a system that relies on a physically measurable parameter, achieving a direct and thorough monitoring. Smartphones today have become an integral carry-on item in our daily lives. With the increase in today's sensing technology, smartphones come equipped with a wide variety of sensors (such as accelerometer, gyroscope, and GPS among others) and wireless interfaces (such as Wi-Fi, 4G, and Bluetooth among others), making them interesting monitoring platforms [38]. A key attribute of a crowd to be monitored is recognition of their current activity status, known as Human Activity

recognition (HAR), as performed in [39] using a cell phone's accelerometer. Therefore, smartphones and smartwatches are able with the wide range of sensing modalities to monitor social behavior and detect crowd dynamics such as flow and bottle necks [40]. Maximizing information content and reducing errors, through the approach of multi-sensor fusion, is of interest of many researchers [41, 42]. In [43], time domain features from a smartphones' triaxial accelerometer, pressure sensor and microphone were employed to differentiate between fifteen different activities through a nonlinear discriminatory analysis (KDA) approach, together with a nonlinear support vector machine classifier (SVM) achieving a 94% in offline subject-independent test. Ronao et al. [44] proposed a multi-layer convolution neural network for human activity recognition, based on raw time-series data from smartphone sensors. The convnet utilized automatic feature extraction on the onboard smartphone's accelerometer and gyroscope to classify six different activities. A total of 10,299 data examples was in use for training and testing the convnet, with each example sensor's axis fed into a 6-channel 1D convolution layer. In addition, FFT extracted information was merged with its corresponding sensor axis at the first convolution layer. Highest accuracy of proposed model returned 95.75%, an increase in comparison to SVM with a returning accuracy of 94.61%. Jiao et al. [45] collected multi-sensor data through their smart golf club for CNN based golf swing classification. The smart golf club integrated two orthogonally affixed strain gage sensors, 3-axis accelerometer and 3-axis gyroscope, creating an intermediate 1D convolutional convolving 8-channel signals.

## Structural monitoring

Besides efforts to avoid bridge disasters by design, Structural Health Monitoring (SHM) techniques can be used to continuously monitor the structure's health. Structural health monitoring (SHM) refers to the use of instruments in a nondestructive method to measure and analyze key structural parameters under operational and environmental conditions, with the core function of warning imminent irregular and/or unfortunate incidents at early stage to avoid further damage and casualties while also providing structure maintenance information thus increasing life span. A smart structure utilizing SHM achieves this with a system consisting of sensors or an array of sensors all connected to a central brain or processing station which continuously reads measured sensor data, fed into algorithms for analysis. The system is able to comprehend the status of the structure automatically reporting to engineers or controllers a warning of an anomaly, inherent failure, or damage.

A large number of research has been focusing on the implementation of SHM systems in the aerospace, automotive and civil industry utilizing various sensor types [46, 47]. Lately employed SHM instruments namely fiber optic sensors provide accurate strain and temperature measurements which may be embedded or retrofitted onto existing structures [48]. Furthermore, applying machine learning algorithms on SHM systems implemented on bridges for damage detection has been tackled a great deal of times. In [49], a PCA baseline model was built using the signals recorded by strain FBGs sensors during experiments with an undamaged structure. During subsequent steps, experiments were performed using the structure in the different possible states (undamaged and six different damage cases). The FBGs showed a very good sensitivity

and were able to detect small strain changes in the structure under the same load conditions when small damages were induced. Sohn et al. [50] used fiber optic strain gauges with Bragg grating to measure the dynamic response of a ship. A vibration-based damage detection problem is cast in the context of statistical pattern recognition. Following an Auto-Regressive (AR) and Auto-Regressive with eXogenous inputs (ARX) prediction models and outlier analysis, the study successfully identifies features from the strain time histories that distinguish the signals recorded under the different structural conditions of the boat. In [51], an SVM-based procedure has been proposed for the detection of fatigue cracking in steel girders based on the data of the SWS. Detection of the fatigue cracking stages is a complicated task due to notable changes of the strain patterns during crack propagation. Features were inputted to SVM classifiers. It was observed that the SVM models can accurately classify most of the damage stages, specifically for cracks larger than 10 mm. In addition, tracking the performance of the SVM models gives an insight into the damage location. Tang et al. [52] applied the field of damage identification for the whole or the main parts of a bridge. The thesis jointly adopts the phase space reconstruction technique and the support vector machine to predict the monitoring data of strain and tilt angle of the Pan Yan-zi Bridge. The work is of great value for the SVM-based studies of the online security alarming technique of the bridge structural monitoring data. In [53, 54], the authors used FBG fiber optic sensors to perform dynamic strain measurements, feeding them into a neural network in order to identify damages in a composite specimen which represented a typical aeronautical construction consisting of skin, frames and stringers. Abdeljaber et al. [55] utilized raw accelerometer vibration signals instrumented on a steel frame

grandstand simulator for SHM damage detection with the adaptive implementation of a 1D CNN model. Possible adaptation of SHM systems on a bridge could be used as a mean for physically measuring crowd dynamics, comprehending and monitoring specific attributes of the crowd.

In most cases, lack of effective crowd management mechanisms and poor physical infrastructure are major contributors to unfortunate pedestrian accidents that can lead to a significant number of casualties. A study conducted by [56] emphasized how the crowd managers need to be informed in advance about a spatial overview with information that may include density, flow and movement. Consequently, real-time crowd and long-term SHM is of great importance, maintaining and avoiding potential safety disasters. Concept system provides a solution to the simultaneously monitoring of both the structure and passing crowd status ensuing intelligent crowd management decisions while reducing associated dangers.

### **C. Aim and Objectives**

The aim of work shown in this thesis is to develop a crowd monitoring solution utilizing SHM based sensors as a foundation. Precisely, we present a framework that allows for its bridge status monitoring objective whilst mainly focusing on simultaneously expanding its functions to crowd monitoring. With the appropriate selection, distribution, and combination of sensors, representation of a crowd's characteristics via machine learning postprocessing is attainable resulting in a comprehensive understanding and estimation of flow and load. A main contribution in our solution instrumentation, is the choice of Fiber Bragg Gratings (FBG) Fiber Optic

Sensors (FOS) strain sensors over conventional ones due to its numerous advantages in measuring strains exhibited by the structure. In conjunction with FBG FOS, wearable devices incorporating Inertial Measurement Unit (IMU) sensors are used to record pedestrian physical activity acceleration forces. Strain data will perceive information of the crowd as a whole, while acceleration forces data will perceive information at the individual level. Data collected are innovatively processed through machine learning methods to generate models, namely Convolutional Neural Network (CNN) and the accustomed Support Vector Machine (SVM). Acquired data from the different sensing modalities are further fused for robust and efficient models' generation. Final models are developed returning valuable crowd attribute information for future decision making, avoiding crowd safety risk.

Following sections of the thesis are organized as follows: Chapter 2 describes the concept solution approach in from a broader perspective with emphasis on machine learning, data fusion, and FOS. Chapter 3 details initial phase of experimental setup, in regards to test bed, instrumentation, and crowd replication. Chapter 4 follows with analysis of signal data obtained including results from generated machine learning models. Chapter 5 subsequently builds upon initial experiment setup primarily in terms of instrumentation. Chapter 6, discusses results generated from SVM and CNN classification and regression models along with fusion approaches and noise and time shift variances testing accordingly. Chapter 7 concludes with the thesis's closing remarks and future work.



## CHAPTER II

### PROPOSED APPROACH AND METHODOLOGY

The proposed approach and methodology present a framework for crowd monitoring and estimation of crowd attributes on data acquired from both the SHM system sensors (FBG FOS strain sensors) and wearable sensors (accelerometers). Through expanding the ability to monitor the status of a bridge structure by means of advanced instrumentation and machine learning techniques, the framework allows for possible flow activity and load characterization classifications and load estimation of overhead crowd. The thesis's main focus will be on applying machine learning on data gathered with the purpose of crowd monitoring and safety as an application.

The end target application comprises a pedestrian bridge equipped with a network of FBG FOS structurally mounted with wearable IMU sensors worn by passing crowds. As pedestrians travel across the bridge, dynamic vibration strain is translated into the structure and measured through the FBG sensors. Individuals performing a physical activity, such as walking induce body acceleration forces through which wearable accelerometers measure. Monitoring a bridge's strain level can yield information of the crowd as a whole, whereas monitoring acceleration forces can yield information at an individual level of the crowd, with the future possibility of expanding functionality to include location and biometric parameters. In our approach, data collected from different sensing modalities serve as input and training information for Support Vector Machine (SVM) and Convolutional Neural Network (CNN) machine

learning algorithm models. With the intent to complement the sensing approach while advancing the capabilities of the models developed by improving reliability and robustness of the system during decision making, sensors' data are further fused and integrated. Resulting model capabilities consist of classification models for crowd flow activity and load characterization and regression model for total crowd load estimation.

### **A. Pattern Recognition and Machine Learning**

A pattern could be said as the opposite of chaos. Pattern Recognition is an important element for problem solving in all scientific fields. Pattern Recognition focuses on the learning and recognition of patterns and regularities within a data, i.e. small intra class variation and large interclass variations. Supervised learning is when pattern recognition systems train on cases with labeled training data. While in Unsupervised learning is when algorithms are used to learn previously unknown patterns due to no labeled trained data.

Pattern Recognition systems can be used to for classification, which attempts to assign each input value to a given set of classes or classify data into one of the defined groups based on key features. Pattern Recognition systems can also be used for regression, which assigns a real valued output to each input while estimating a relationship between among the variables, between output variables, and one or more input variables.

Via statistical approach, each pattern is represented in terms of  $x$  measurements or features and is projected into a point  $x$ -dimensional space. The aim is to choose those features that allow pattern vectors belonging to different classes to occupy compact and

disjoint regions in an  $x$ -dimensional feature space. The effectiveness of the representation space (feature set) is determined by how well patterns from different classes can be separated. Given a set of training data patterns from each class, the objective is to establish decision boundaries in the feature space which separate patterns belonging to different classes [57]. In the statistical decision theoretic approach, the decision boundaries are determined by the probability distributions of the patterns belonging to each class, which must either be specified or learned [58, 59].

### *i. Support Vector Machine (SVM)*

Developed by Vapnik [60] and Cortes [61], Support Vector Machine (SVM) is a strong and robust algorithm in machine learning. SVM has become a well-known approach exploited with positive results in many pattern classification and regression applications [62, 63], such as face/object detection and recognition, handwritten digital/character recognition, speech/speaker verification and recognition, information/image retrieval, gender classification, prediction, and further more. It's regarded as the state-of-the-art tool for resolving linear and non-linear classification problems, thanks to its parsimony, flexibility, prediction capacity and the global optimum character. The basis of their formulation is the structural risk minimization, rather than the empirical risk minimization which is traditionally used in Artificial Neural Networks [60].

SVM can be basically used to determine an optimal separating hyper-plane or decision surface by embracing a novel technique based on mapping the sample points into a high-dimensional feature space and it is categorized using a nonlinear

transformation  $\Phi$ , even when the data are linearly inseparable. In other terms, SVM with different kernel functions can transform a nonlinear separable problem into a linear separable problem by projecting data into the feature space and then finding the optimal separate hyper-plane [60]. This transformation is carried out by kernel functions like linear, radial basis/gaussian function, sigmoid and polynomial. The optimal hyper-plane is gained by solving a quadratic programming problem which is reliant on regularization parameters.

In other words, the basic idea of the SVM is to construct a hyper-plane as the decision plane, which separates the positive and negative classes with the largest margin, which is related to minimizing the VC dimension of SVM. When the two classes are linearly separable, we wish to find a separating hyper-plane which gives the smallest generalization error among the infinite number of possible hyper-planes. Such an optimal hyper-plane is the one with the maximum margin of separation between the two classes, where the margin is the sum of the distances from the hyper-plane to the closest data points of each of the two classes. These closest data points to the hyper-planes are called Support Vectors (SVs), shown in Figure 1 as data points on the dashed line [63]. An extension to nonlinear decision surfaces is necessary since real-life classification problems are hard to be solved by a linear classifier. When the decision function is not a linear function of the data, the data will be mapped from the input space into a high dimensional feature space by a nonlinear transformation. In this high dimensional featured space, the generalized optimal separating hyper-plane is constructed. Cover's theorem states that if the transformation is nonlinear and the dimensionality of the feature space is high enough, then input space may be transformed

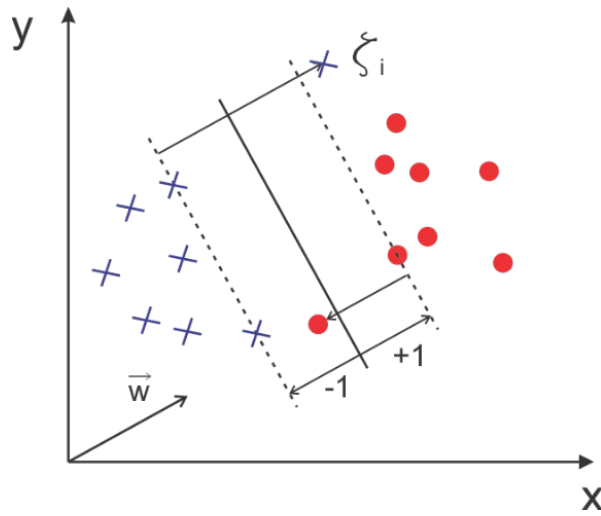
into a new feature space where the patterns are linearly separable with high probability. This nonlinear transformation is performed in implicit way through so-called kernel functions (Kernel types include linear, radial basis function, sigmoid and polynomial). In order to accomplish nonlinear decision function, an initial mapping of the data into an (usually significantly higher dimensional) Euclidean space is performed, and the linear classification problem is formulated in the new space with dimension  $d$ . The training algorithm then only depends on the data through dot product. The data can become linearly separable in feature space although original input is not linearly separable in the input space. Hence kernel substitution provides a route for obtaining nonlinear algorithms from algorithms previously restricted to handling linear separable datasets. The use of implicit kernels allows reducing the dimension of the problem and overcoming the so-called dimension curse. Variant learning machines are constructed according to the different kernel function and thus construct different hyper-planes in feature space [63]. In practical applications for real-life data, the two classes are not completely separable, but a hyper-plane that maximizes the margin while minimizing a quantity proportional to the misclassification errors can still be determined. This can be done by introducing positive slack variables. If an error occurs, a parameter  $C$  is chosen by the user that controls the tradeoff between the margin and the misclassification errors. A larger  $C$  means that a higher penalty to misclassification errors is assigned [63].

$$\frac{1}{2}w^T w + C \sum_{i=1}^N \xi_i \quad (1)$$

with constraints being:

$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i \quad (2)$$

$$\xi_i \geq 0, i = 1, \dots, N \quad (3)$$



**Figure 1. A simple linear support vector machine – hyperplane separation for classification**

The single-class classification can be extended adopting multi-class classification through either constructing and combining (known as one-against-one) or by considering all the data at once (known as one-against-all).

Traditional/statistical regression procedures are often stated as the processes deriving a function that has the least deviation between predicted and experimentally observed responses for all training examples. One of the main characteristics of Support Vector Regression (SVR) is that instead of minimizing the observed training error, SVR attempts to minimize the generalized error bound so as to achieve generalized

performance. This generalization error bound is the combination of the training error and a regularization term that controls the complexity of the hypothesis space [64]. In  $\varepsilon$ -SV regression [60], our goal is to find a function that has at most  $\varepsilon$  deviation from the actually obtained targets for all the training data, and at the same time is as flat as possible. In other words, we do not care about errors as long as they are less than  $\varepsilon$ , but will not accept any deviation larger than this. The next step is to make the SV algorithm nonlinear. This, for instance, could be achieved by simply preprocessing the training patterns by a map into some feature space, and then applying the standard SV regression algorithm. The input pattern (for which a prediction is to be made) is mapped into feature space by a map. Then dot products are computed with the images of the training patterns under the map. This corresponds to evaluating kernel functions. Finally, the dot products are added up using the weights. This, plus a constant term yields the final prediction output. The process described here is very similar to regression in a neural network, with the difference, that in the SV case the weights in the input layer are a subset of the training patterns [65].

$$\frac{1}{2}w^T w + C \sum_{i=1}^N \xi_i + \sum_{i=1}^N \xi^*_i \quad (4)$$

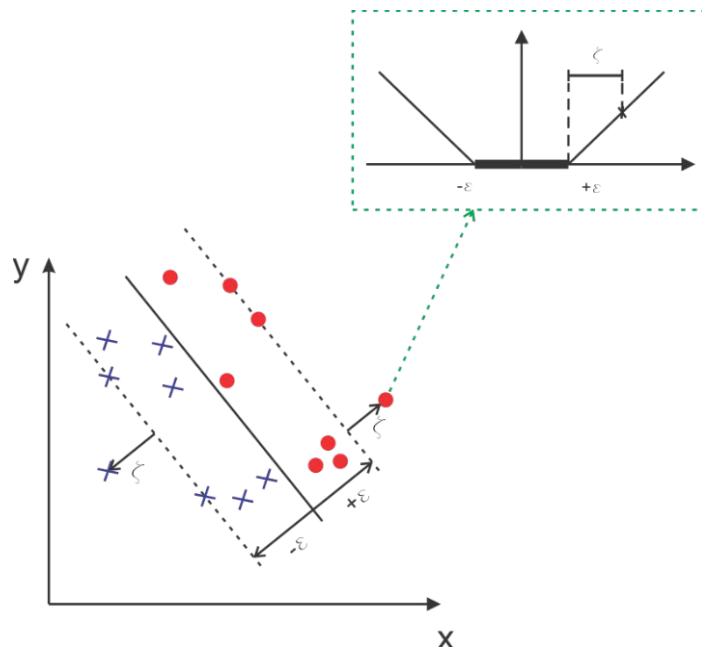
with constraints being:

$$w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi^*_i \quad (5)$$

$$y_i - w^T \phi(x_i) - b_i \leq \varepsilon + \xi^*_i \quad (6)$$

$$\xi_i, \xi^*_i \geq 0, i = 1, \dots, N \quad (7)$$

The model produced by support vector classification only depends on a subset of the training data, because the cost function for building the model does not care about training points that lie beyond the margin. Analogously, the model produced by SVR only depends on a subset of the training data, because the cost function for building the model ignores any training data that is close (within a threshold  $\epsilon$ ) to the model prediction [64], while defining the loss function that ignores errors, which are situated within the certain distance of the true value as shown in Figure 2. This type of function is often called – epsilon intensive – loss function.



**Figure 2. A simple linear support vector machine – fitting for regression**

Detailed explanation and mathematical formulation of Support Vector Machine classification and regression has been thoroughly published [63-68].



## *ii. Gaussian Process (GP)*

Gaussian Process (GP) is a collection of random variables, any finite number of which has (consistent) joint Gaussian distributions. A Gaussian process is fully specified by its mean function and covariance function. The mean is a function (which is often the zero function), and the covariance is a function which expresses the expected covariance between the value of the function at the points. The actual function in any data modeling problem is assumed to be a single sample from this Gaussian distribution. The equivalent kernel is a way of understanding how Gaussian process regression works for large sample sizes based on a continuum limit. In Gaussian process regression, the covariance between the outputs at input locations is usually assumed to depend on the distance [64]. This is a natural generalization of the Gaussian distribution whose mean and covariance is a vector and matrix, respectively [69]. Meaning: “the function is distributed as a GP with mean function and covariance function” [69].

GPR has several advantages, including adaptation to nonlinearities with kernel functions, robust selection of kernel hyper-parameters via maximization of marginal likelihoods (namely type-II maximum likelihood), and a Bayesian formalism for inference that enables better generalization from small training sets [70].

Detailed explanation and mathematical formulation of GP can be found at [70, 71].

### *iii. Convolutional Neural Network (CNN)*

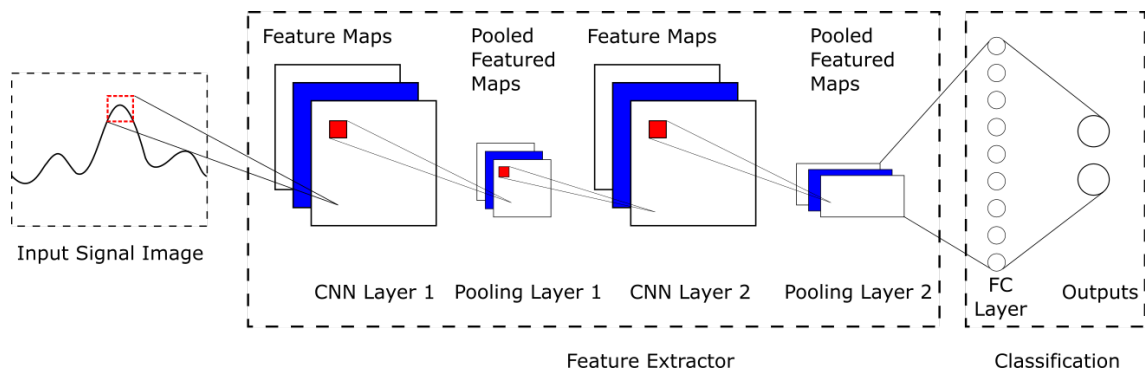
A Convolution Neural Network (CNN) is a deep artificial neural network which possesses huge representational capacity and that automatically learns the useful good features at every layer of the network hierarchy. The network consists of multiple layers each fulfilling a specific role with learnable weights and biases. It has been effectively applied to many vision-based classification problems such as visual object recognition and handwriting digit/character recognition [41-44].

Being hierarchical, multi-layer neural networks with a deep supervised learning architecture and trained with the back-propagation algorithm [43], CNNs are composed of an automatic feature extractor and a trainable classifier. CNN have the capability of extracting features information from its data without the need of manual hand-crafted feature-based selection. Advantage wise, it removes the element of the designer and relies on the algorithm to detect and extract globally relevant discriminative features in the data. The automatically extracted features from the input image have the benefit of being invariant to the shift and shape distortions such as in a case of input textual images [45]. In addition, CNN mixes three main hierarchical aspects such as local receptive fields, weight sharing and spatial sub-sampling/pooling [43].

CNNs are exploited to learn complex, high dimensional data, and differ in how convolutional and sub-sampling layers are queried into, i.e. the difference lies within their architecture with each designed for a specific form of application.

The net shown in Figure 3 represents an example of a CNN architecture with a pair of two layers. Initially, the input is convolved with a set of filters in order to obtain values of the feature map. Next, in order to reduce the dimensionality of the spatial resolution

of the feature map, each convolution layer is pursued by a sub-sampling/pooling layer. Convolutional layers alternate sub-sampling layers constitute the feature extractor to retrieves discriminating features from the raw images. Ultimately, these layers are followed by a Fully Connected (FC) layer and the output layer with the output of the previous layer taken by each layer as the input [46].



**Figure 3. Data processing stages in a simple convolutional neural network**

A convolution layer is composed of several convolution kernels which are used to compute different feature maps. Specifically, each neuron of a feature map is connected to a region of neighboring neurons in the previous layer. Such a neighborhood is referred to as the neuron's receptive field in the previous layer. The new feature map can be obtained by first convolving the input with a learned kernel and then applying an element-wise nonlinear activation function on the convolved results. Note that, to generate each feature map, the kernel is shared by all spatial locations of the input. The complete feature maps are obtained by using several different kernels. The activation function introduces nonlinearities to CNN, which are desirable for multi-layer networks to detect nonlinear features. Typical activation functions are sigmoid,

tanh and Rectified Linear Unit (ReLU). Feature representations' complexity level increases with the depth of the convolutions in the network. Placed between two convolutional layers, the pooling layer aims to achieve shift-invariance and efficiency by reducing the resolution of the feature maps. It is achieved through operations such as average and max pooling. Each feature map of a pooling layer is connected to its corresponding feature map of the preceding convolutional layer. By stacking several convolutional and pooling layers, we could gradually extract higher-level feature representations. After several convolutional and pooling layers, there may be one or more fully-connected layers which aim to perform high-level reasoning. They take all neurons in the previous layer and connect them to every single neuron of current layer to generate global semantic information. The output layer of a CNNs is typically a softmax operator that computes the probability distribution over the predicted classes, followed by a classification layer. Whereas for regression applications, the classification layer is replaced by a regression layer with the exclusion of the softmax operator. Detailed information on CNN architecture can be found in this paper [72].

## **B. Data Fusion**

A major aspect of the proposed system concept relies on data fusion between FBG and IMU sensors. The nature of measurement from the two sensing modalities allows for the monitoring of both the structure and individuals in the crowd. The fusion aspect provides redundancy against possible failure. Data collected from different sensors are combined and jointly processed for machine learning model generation, estimating specific crowd attributes. This allows for a richer representation of the

crowd-structure system. Added benefits include the capability of distinguishing between possible crowd overloading or structure deterioration or dangerous loading. Outputs with increased certainty provide information, triggering any preventative protocols.

Depending on application, data fusion is achieved at three different arbitrary levels; input, feature, and decision. Input-level fusion can be categorized when different sensors' data are directly combined, forming a new single data array. This may be then used as an input for machine learning algorithms. Fusion at the feature-level occurs when the machine learning algorithm fuses multiple data which has been simultaneously fed through. Decision-level fusion occurs at the machine learning algorithm outputs, where they are fuse by undertaking a decisive output based on discriminative results.

### **C. Fiber Optic Sensors (FOS)**

Electrical sensors have for decades been the standard mechanism for measuring physical and mechanical parameters. Despite their ubiquity, these sensors have inherent limitations such as transmission loss, susceptibility to electromagnetic interference (noise), sensitivity due to seasonal changes and operational conditions. With the addition of power supply requirement for every sensor, their usage thus becomes challenging or impractical in many applications. Another problematic challenge for SHM systems utilizing conventional sensors especially when dealing with large structures spanning a long distance with high number of sensing nodes, is wiring. Fiber optic sensing excels to these challenges, replacing light in standard optical fiber with electricity in copper wire.

Devices into which the measured object or input signal introduces modifications or modulations in some of the characteristics of light in an optical system can be considered Photonic/ Optical Sensors (OS). If fiber optic technology is used in any of the processes or parts, then the OS can be considered an Optical Fiber Sensor (OFS) or Fiber Optic Sensing (FOS) [73]. Fundamentally, a fiber optic sensor works by modulating one or more properties of a propagating light wave (including intensity, phase, polarization, wavelength, and frequency) in response to the environmental parameter being measured, which can range from mechanical, chemical, and biological. The usefulness of the fiber optic sensor therefore depends upon the magnitude of this change and our ability to measure and quantify the same reliably and accurately [74].

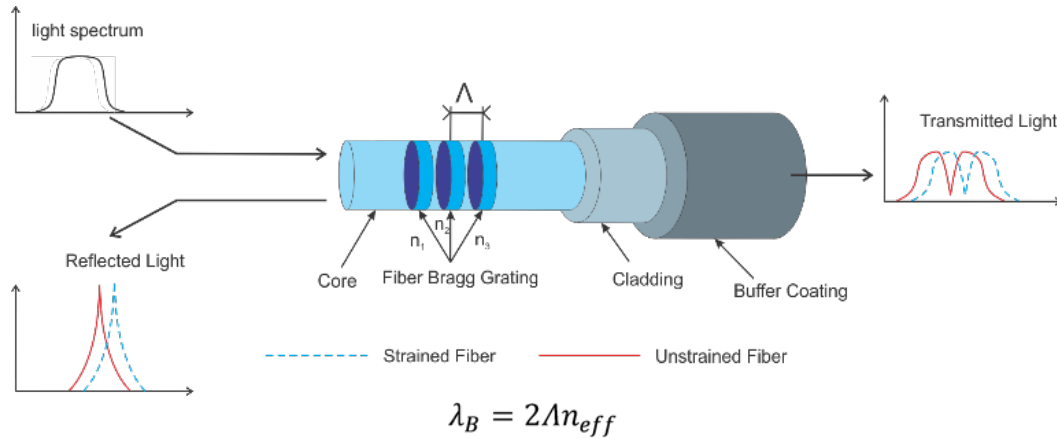
The main advantages of FOS are derived from the particular characteristics of its core material, Silica which has the following benefits: passive, dielectric, low losses at optical frequencies, electromagnetic interference immunity, chemically inert, biocompatible, ability to withstand high temperatures, multiplexing capability, quasi- and fully distributed sensing points, small and lightweight. Although expensive when compared to conventional instrumentations, their multiplexing capability and capability to run over distances as great as 20 km outweigh the cost [73]. Even though fiber optic sensors are apparently expensive for widespread use in health monitoring, they are better approaches for applications where reliability in challenging environments is essential. Price is often no longer a showstopper when the security or efficient management of very expensive systems, such as civil engineering structures, could lead to catastrophic consequences. Fiber optic sensors can even become cost effective when involving a significant number of sensors, such as in civil engineering applications. In

some extreme applications, such as in the oil and gas industry, fiber-optic sensors are sometimes the only available solution for reliable and long-term physical parameter monitoring [75].

*i. Fiber Bragg Grating*

Intrinsic optical sensors can measure temperature, strain, pressure, and other parameters by monitoring the resulting changes in the intensity, phase, polarization, wavelength or transit time of light within the fiber. One of the most common and technologically advanced types of FOS is based on Fiber Bragg Grating (FBG). FBG sensors operate in a way such that gratings or ‘wavelength selective mirrors’ are inscribed into the fiber which reflects only a specific wavelength, referred to as the Bragg wavelength (which have very small bandwidths so-called peaks), from the input light wavelength spectrum with the rest transmitting through.

Stretching a strain gage sensor causes a change in resistance of the metal grid. When an FBG sensor, hence the grating period  $\Lambda$  is stretched, compressed, or undergoes thermal expansion and contraction, the Bragg wavelength shifts accordingly as shown in Figure 4. When a broad-spectrum light beam is sent to an FBG, reflections from each segment of alternating refractive index interfere constructively only for a specific wavelength of light, called the Bragg wavelength. This effectively causes the FBG to reflect a specific frequency of light while transmitting all others.



**Figure 4. FBG structure and working principle in respect to shifts in reflected Bragg wavelength due to strain**

Strain and temperature changes are identified by a data acquisition hardware know as an interrogator. The interrogation units are responsible for reading the Bragg wavelength shift of the FBGs induced by various the physical parameters [76]. The interrogator then uses a wavelength demodulation technique to observe the change in wavelength and interpret them into meaningfully correlated measurements. Because the Bragg wavelength is a function of the spacing between the gratings, FBGs can be manufactured with various Bragg wavelengths, which enable different FBGs to reflect unique wavelengths of light. FBG techniques can either have a handful of sensing points or be fully distributed as well. By manufacturing a fiber with continuously inscribed FBGs, engineers can analyze the changes in the way the light reflects and interpret this information to provide accurate measurements. This enables interrogators that use continuously written gratings to obtain precise measurements down to millimeters of spatial resolution even in harsh and dynamic environments.

Since both strain and temperature induce reflected Bragg wavelength shift; therefore, their contributions have to be separated for accurate measurements of each



variable. The simplest one is called reference fiber method, which uses a dummy reference Bragg grating subjected to the same thermal environment but free from mechanical load.

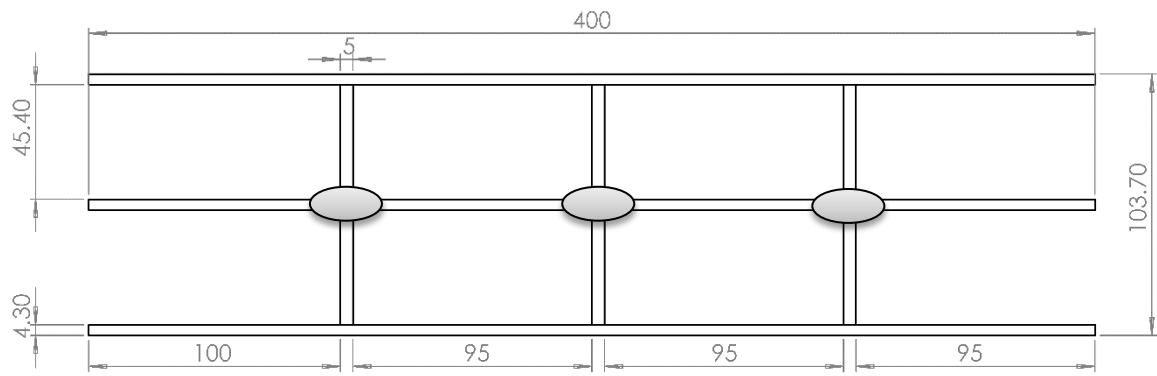
## CHAPTER III

### EXPERIMENTAL VALIDATION – PHASE I

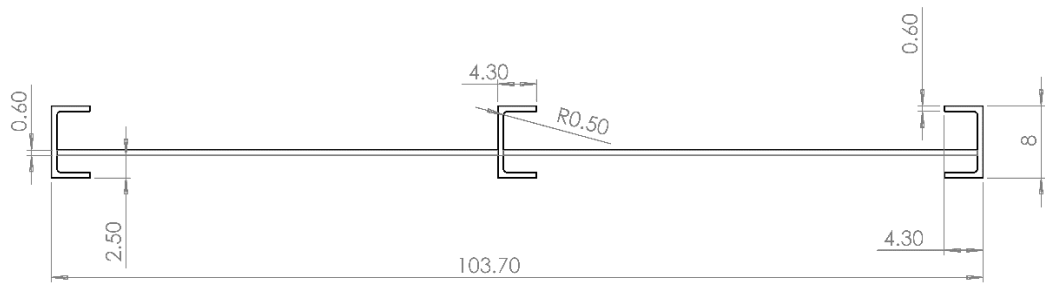
In order to proof validity of monitoring concept, a preliminary experimental setup is worked on to put the proposed approach and methodology to test, were a subset of aims and objectives are to be implemented. in Experimental Validation – Phase I will focus on a defined section regarding instrumentation and machine learning approach. FBG based FOS are used for instrumentation on a testing platform to replicate crowd flow on a structure. Data collected are used to classify crowd flow activity and load estimation using SVM and GP machine learning approach; SVM for classification, and SVR along with GP for regression accordingly.

#### **A. Test Bridge**

A scaled test bridge model, shown in Figure 5, measuring four meters in length and one meter in width and consisting of three steel C-beams is constructed. Three shoulder steel plates connect the three C-beam together laterally at approximately every 0.95 meters interval along the length. The bridge is modeled as a simply supported beam with a pin and roller. Wood panels are placed on the top surface of the bridge allowing full load distribution as well as ease for a pedestrian volunteer to walk. Subsequent instrumentation and crowd simulation for data gathering is to be done on the designed test bridge.



(a)



(b)



(c)

**Figure 5. Test-bed bridge to be used for crowd data acquisition (a) technical drawing (cm) - top view, (b) technical drawing (cm) - front view, (c) assembled**

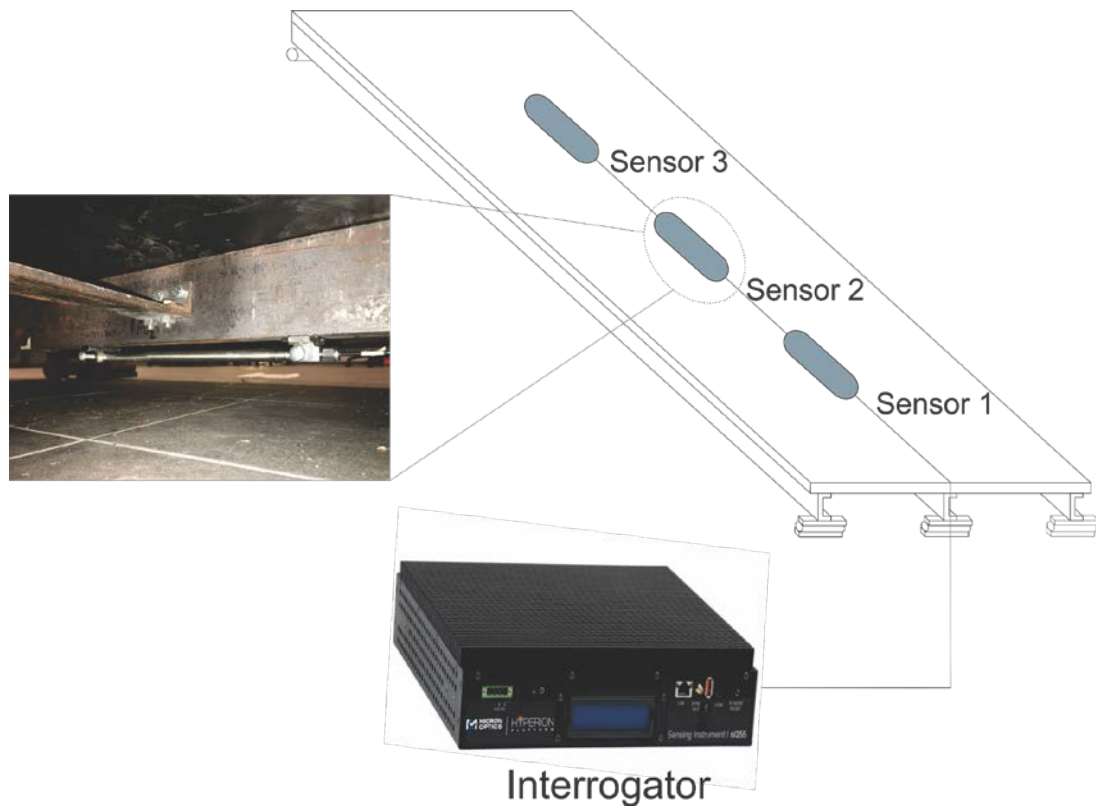
## **B. Instrumentation**

Experiment instruments were sourced from Micron Optics. Sensing consisted of three FBG os3610 (25cm gage) surface mounted strain sensors to measure structural strain levels. The sensors were daisy-chained along and mounted underneath the central C-beam (at marked location shown in Figure 5(a)) of the bridge via epoxy bonded brackets, as shown in

Figure 11 installed at the marked locations in Figure 6. Each sensor contains two FBGs, where one measures strain while the other measures temperature for compensation purposes in real-time. A Micron Optics wavelength division multiplexer interrogator, Hyperion si255, was employed to send, receive, and interpret light signals at a sampling rate of 100 Hz through the FBG sensors. A laptop with ENLIGHT software (Micron Optics propriety interrogator software) was used for recording and saving the strain data from the interrogator locally on the machine for further processing. The recording process is initialized by a trigger action where one of the end sensors read above  $5 \mu\epsilon$ , with a pre-trigger recording of 100 milliseconds and total recording length of 10 seconds.

## **C. Crowd Replication**

In order to generate a dataset for the machine learning models, emulating crowd flow on a pedestrian bridge is essential. Groups of volunteers were gathered to walk on the test bridge under various sizes (as shown in Table 1) and two flow activity types: walking leisurely (slow) or walking hurriedly (fast). As a group walked across the bridge, strain levels from the structure's three FBG sensors were recorded for the collective crowd. Each recorded observation was labeled according to the total mass of

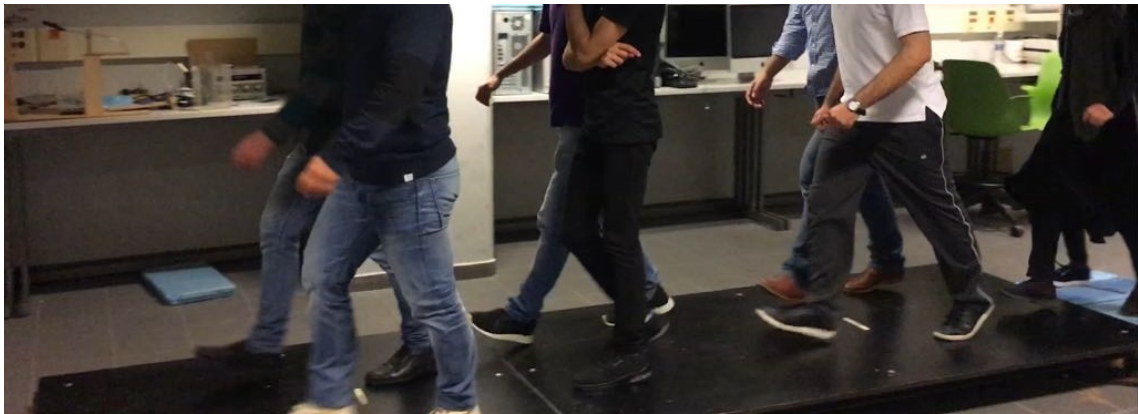


**Figure 6. Experiment setup demonstrating how the interrogator is connected to the three daisy-chained FBG strain sensors mounted below the middle C-beam**

pedestrian(s) walking, group size, and flow activity type; slow or fast. Each control group walked across the bridge forwards and back five times for each slow and fast respectively, resulting in a dataset of 1400 observations.

**Table 1: Volunteer composition used in Phase 1 crowd replication; dataset indicating size of groups, total groups and average mass per group size**

Group Size	Total Groups	Average Mass (Kg)
1	18	81
2	8	158.7
3	8	221.8
4	8	289
5	8	366.7
6	8	449.6
8	8	580.8
10	4	707.4



**Figure 7. Group of eight walking fast on the test bridge**

## CHAPTER IV

### RESULTS AND DISCUSSIONS – PHASE I

Recorded strain data was further processed further using MATLAB 2017b for generating machine learning models. Models were generated for three different inputs types as training data. For future application in mind, Raw strain data was used as a training data. In addition, Fast Fourier Transfer (FFT) function applied on the Raw signal was used. Furthermore, several manually extracted features from the Raw signal were used as training data, of which some included: Maximum Strain, Average Maximum Strain, Average Peaks\*, Maximum Peaks\*, Number of Peaks\*, Area under Strain Curve, Maximum FFT value and location between 1.5Hz and 2.5Hz, and Mean Frequency.

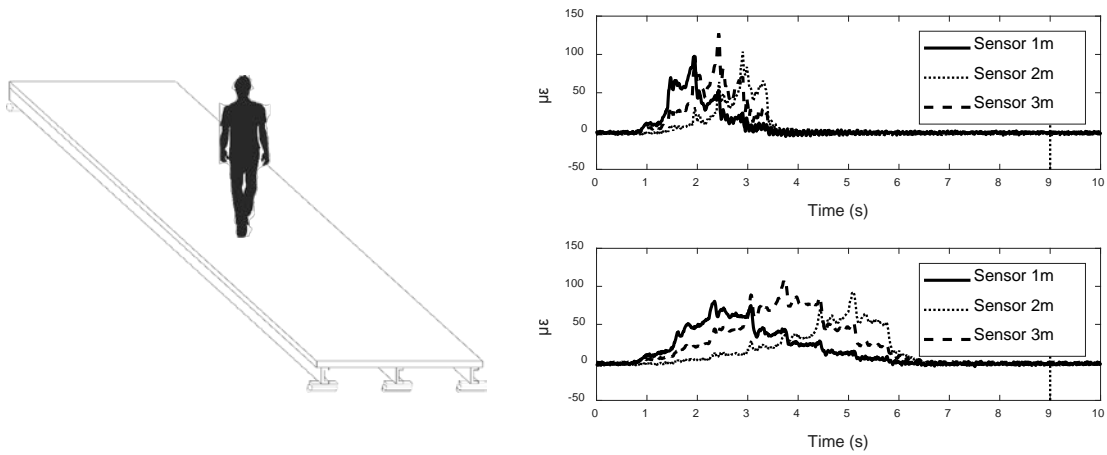
\*(Low pass Filter was applied to the Raw signal for features related to peaks)

#### **A. Raw Signal**

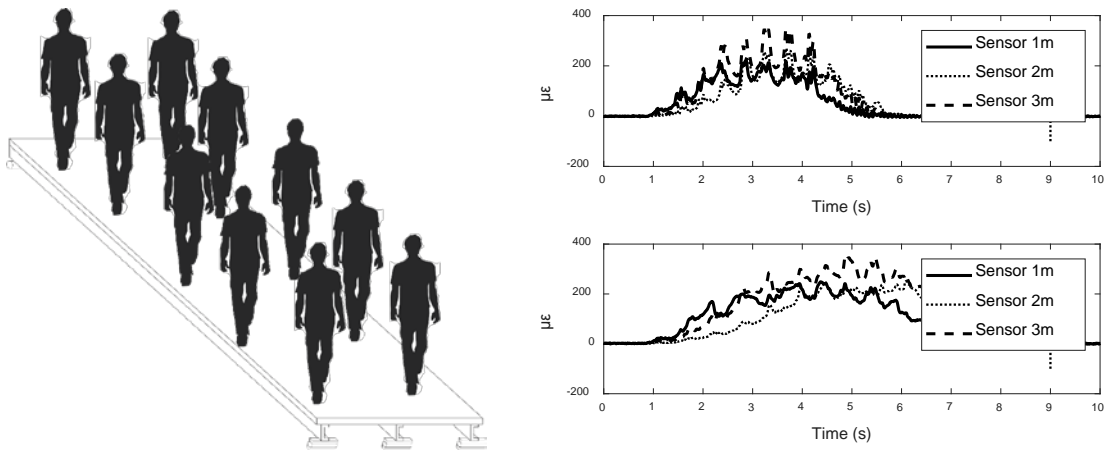
Raw data signal of a group of one and a group of ten walking under both flow activity of Slow and Fast is shown is shown in Figure 8 and Figure 9. Visible distinction in terms of peaks, length, and area under signal curve can be seen between same groups of either Slow or Fast, in addition to between groups of different sizes and total mass.

It was evident from the strain signal that a person's/collective crowd steps can be visually differentiated and identified as the person/crowd walks pass the sensors.

According to Figure 8, maximum read strain value per sensor occurs when



**Figure 8. One person – 74 kg, (top right) fast, (bottom right) slow**



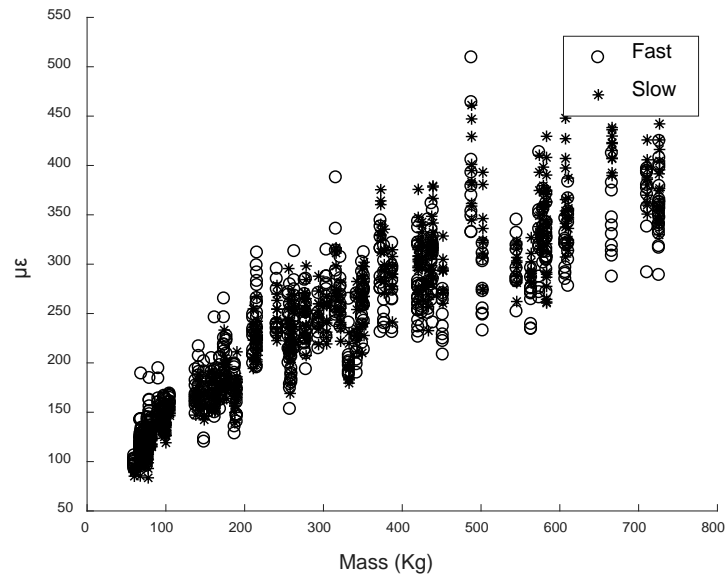
**Figure 9. Ten persons – 726 kg (top) fast, (bottom) slow**

the volunteered single person walking is directly above it, with maximum global strain occurring when the walking passed the middle of the bridge and above the middle sensor. A similar trend pertaining to a group of ten can also be seen in Figure 9, along with an observation of the strain signal returning similar to that of a static load during the continuous flow of crowd.

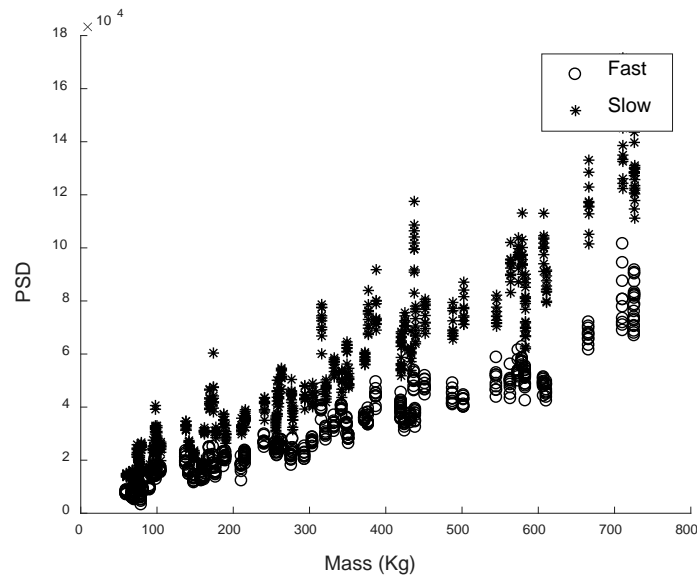


## B. Features

Features extracted demonstrate a visible pattern whether in terms of a linear relationship to weight or distinction between slow and fast. A couple of manually extracted features from the middle sensor are presented in Figure 10. Features related to FFT, Mean Frequency, and Area under Strain Curve are used in generating classification models, with Maximum Peak, Maximum Strain and Area under Strain Curve as well can in regression model generation.



(a)



(b)

Figure 10. Selected extracted features (a) Maximum Strain vs. Mass, (b) Area vs. Mass

### C. Classification and Regression Models

Raw signal (consisting of 3000 data points), FFT signal (consisting of 1500 data points) and extracted features (consisting of 25 data points) are each served as types of training data for the machine learning models. SVM/R is used for both single class classification of crowd flow activity; Slow or Fast, and load estimation prediction of crowd load mass accordingly, with both utilizing three different kernels; Quadratic, Cubic. GP will be additionally used for regression utilizing an Exponential kernel. Data was split with 20% not used in training assigned to Testing. Classification test results are that of true positive and true negative sensitivity from a confusion matrix and their corresponding accuracy percentage, with a higher number indicating a better performance. Regression test results are based on the percentage difference between

estimated and actual mass followed by the mean for all the groups in both flow activities, with lower numbers indicating a better performance.

**Table 2: Classification and regression model testing results per signal type and machine learning approach**

Classification – crowd flow activity; Slow or Fast		
	Testing Accuracy (%)	Sensitivity (Slow   Fast)
Training Data: Raw Signal		
SVM - Quadratic	90.59	0.91   0.91
SVM - Cubic	88.53	0.91   0.86
Training Data: FFT Signal		
SVM - Quadratic	93.82	0.96   0.92
SVM - Cubic	93.53	0.95   0.92
Training Data: Extracted Features		
SVM - Quadratic	92.06	0.99   0.85
SVM - Cubic	90.88	0.95   0.87
Regression – crowd load mass estimation		
	Testing Mean Error (%)	
Training Data: Raw Signal		
SVM - Quadratic	11.749	
SVM - Cubic	14.875	
Gaussian - Exponential	8.1117	
Training Data: FFT Signal		
SVM - Quadratic	17.954	
SVM - Cubic	19.084	
Gaussian - Exponential	16.889	
Training Data: Extracted Features		
SVM - Quadratic	13.408	
SVM - Cubic	13.181	
Gaussian - Exponential	9.6033	

It is evident according to the classification section of Table 2, that model based on FFT signal as input returned the highest accuracy on our test data set. However Raw signal input and extracted features still perform with relatively with high accuracies

similar to that of FFT signal. In terms of kernels, Quadratic always seemed to return a better flow activity classification with their sensitivity being higher than that of Cubic.

It is evident according to the regression section of Table 2, that regression based on Raw signal as input returns the least error, followed by Extracted Features and FFT signal. In addition, SVM - Quadratic kernel presented a lower mean error when compared to SVM - Cubic. However, Gaussian - Exponential generally returned the lowest error out of all models, reaching as low as about 10%.

# CHAPTER V

## EXPERIMENTAL VALIDATION – PHASE II

Succeeding the proof of concept regarding the proposed approach and methodology in Experimental Validation – Phase I, complete aim and objective application scope could now be tackled within Experimental Validation – Phase II.

The working experimental setup from Experimental Validation – Phase I is supplemented in the files of instrumentation, fusion, machine learning approach with testing. FBG based FOS alongside wearable accelerometers are instrumented on the test bridge. Crowd replications are proceeded with their data serving as SVM and CNN machine learning model inputs for flow activity and load characterization classification and load estimation.

### **A. Test Bridge**

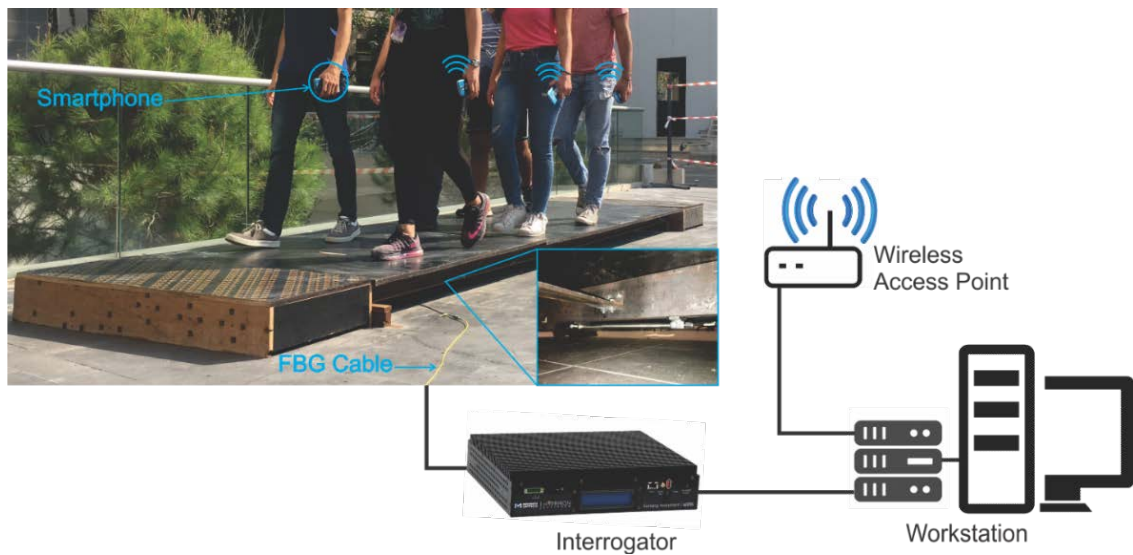
The test bridge utilized in Experimental Validation – Phase I was re-allocated from an indoor temperature-controlled laboratory to an outdoor location, where it was exposed to representative field parameters in terms of temperature cycles among others.

### **B. Instrumentation**

In addition to identical instrumentations utilized in Experimental Validation – Phase I, sensing instruments of a different modality was added to obtain a better understanding of a crowd in motion along the test bridge. The pedestrian's physical

activity status was measured through wearable IMU sensors. Measurement data of body acceleration forces were acquired by a hand-held smartphone with inbuilt accelerometers along the x-, y-, and z-axis, as shown in

Figure 11. The smartphone ran a custom application that records all three axis acceleration forces and transmits them with a unique identification number via Wi-Fi over to the workstation's connected wireless access point at a rate of 20Hz where it is saved locally to a database for further processing.



**Figure 11. Experiment setup demonstrating a crowd walking along the test bridge. FBG sensors mounted below and connected to the interrogator measure strain. Smartphones held per volunteer sends accelerometer data via WiFi to the wireless access point. Workstation records strain and acceleration data for post-processing**

### C. Crowd Replication

In order to generate a new dataset for the machine learning models similarly as one before, emulating crowd flow on a pedestrian bridge is essential. Groups of volunteers were gathered to walk on the bridge under various sizes (as shown in Table 3) and two flow activity types: walking leisurely (slow) or walking hurriedly (fast). As a

group with each of its members holding a smartphone in their hand as he or she walked across the bridge, strain levels from the structure’s three FBG sensors were recorded for the collective crowd along with individual group members’ flow-activity induced acceleration forces from smartphones. Each recorded observation was labeled according to the total mass of pedestrian(s) walking, group size, and flow activity type; slow or fast. Each control group walked across the bridge forwards and back twice for both slow and fast respectively, resulting in a dataset of 488 observations.

**Table 3: Volunteer composition used in crowd replication; dataset indicating size of groups, total groups and average mass per group size**

Group Size	Total Groups	Average Mass (kg)
1	10	78.3
2	10	154.3
3	10	225.5
4	10	291.8
5	10	392.4
6	10	444.3
8	1	605.9

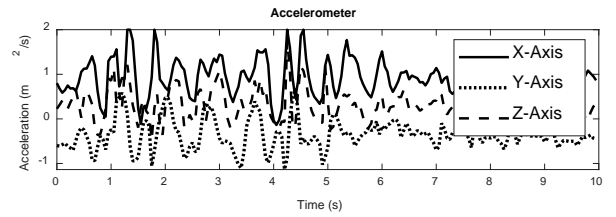
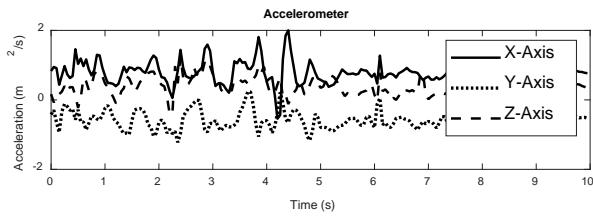
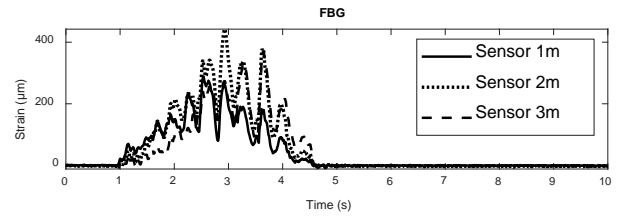
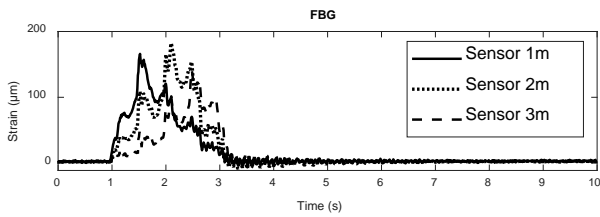
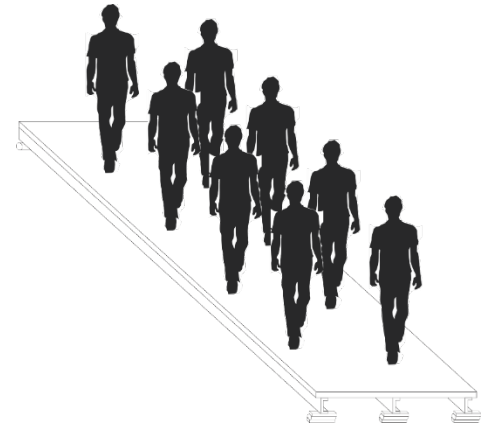
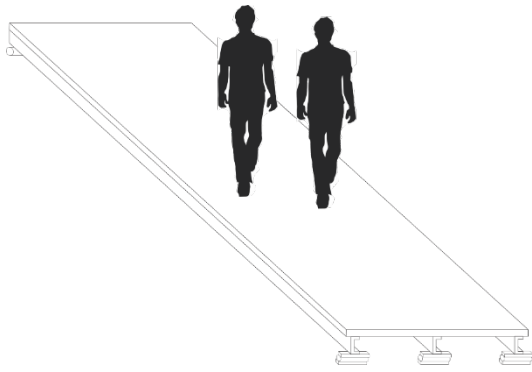
#### **D. Raw Signal**

Figure 12 illustrates the time synced strain and acceleration levels exhibited by a group of 2 individuals weighing 143.9Kg while travelling on the bridge under both flow activities; ‘Fast’ Figure 12(a) and ‘Slow’ Figure 12(c). In addition to, a group of 8 individuals with a total mass of 605.9Kg is shown under both flow activities as well, Figure 12(b) and Figure 12(d).

It is evident from the FBG strain signal and IMU acceleration signal that crowd dynamics are translated with distinguishable patterns. The crowd’s steps can be visually

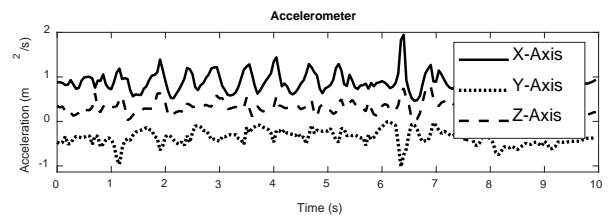
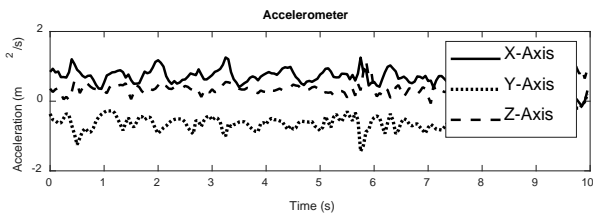
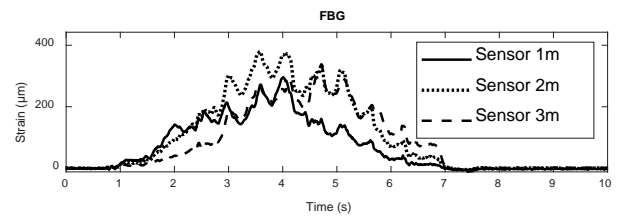
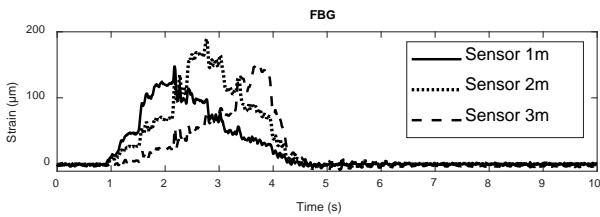
differentiated and identified as peaks while passing through the FBG sensors, with global peak strain level reading occurring over the centrally mounted bridge sensor 'Sensor 2m'. Noticeable strain difference between a crowd of two and eight can be observed, regardless of their flow activity (Figure 12(a) and Figure 12(b) or Figure 12(c) and Figure 12(d)), in term of the maximum amplitude measured by the FBG sensors, with higher values indicating total mass and generally larger group sizes. Generally, a 'Fast' flow activity when compared to 'Slow' flow activity can be seen with a higher amplitude and shorter activity time span for FBG strain readings regardless of group size. Under the same time span for a 'Flow' flow activity, accelerometer readings for x- and y- axis show sharper rate of change and gradient when compared to 'Slow' flow activity. Both sensors demonstrate potential utilization of advanced machine learning algorithms to accurately classify flow and load characterization in addition to load estimation.





(a) Group 02 – Fast

(b) Group 08 – Fast



(c) Group 02 – Slow

(d) Group 08 – Slow

**Figure 12. Strain and acceleration levels for a group of two under (a) ‘Fast’ and (c) ‘Slow’ flow activity. Strain and acceleration levels for a group of eight under (b) ‘Fast’ and (d) ‘Slow’ flow activity**

## CHAPTER VI

### RESULTS AND DISCUSSIONS – PHASE II

#### A. Classification and Regression Models

Machine learning models require input data for training purposes. Recorded data from crowd replication on the test bridge is to be used for the purpose of machine learning model's generation. Collected data which included both strain and acceleration data with their corresponding labels, is used for both classification and regression models. In regards to outputs, a two-class classification model yields the capability of categorizing a crowd's flow activity type, i.e. whether the crowd is moving under 'Fast' or 'Slow' flow activity. Data are consequently split equally under the two different labels. In order to return output with more detailed characterization of the crowd, a multi-class classification model is trained with the added labelling to include a load characterization class, i.e. 'Light' or 'Heavy' based on the total weight of the crowd. Multi-class output labels include 'Slow-Light', 'Slow-Heavy', 'Fast-Light', and 'Fast-Heavy'. Load characterization labels are based on total crowd mass of a 250 Kg threshold, further split obtaining roughly equal label distribution. Regression models are used to estimate the total crowd mass on bridge. Labels used during training relied only on total mass recorded during data acquisition. Due to the varying nature of human dynamics and behavior, two regression models are to be generated depending on the flow activity type; a model for each of the 'Slow' and 'Fast' activity types. After classification, the total crowd mass is estimated using the appropriate regression model according to its flow activity

All processing was performed on a workstation with the following specifications: Intel® Xeon® W-2123 CPU @ 3.60Ghz, 64.0 GB RAM, NVIDIA Quadro P600. Classification models were generated on the workstation utilizing Matlab 2018b Statistics and Machine Learning Toolbox for SVM and Deep Learning Toolbox for CNN.

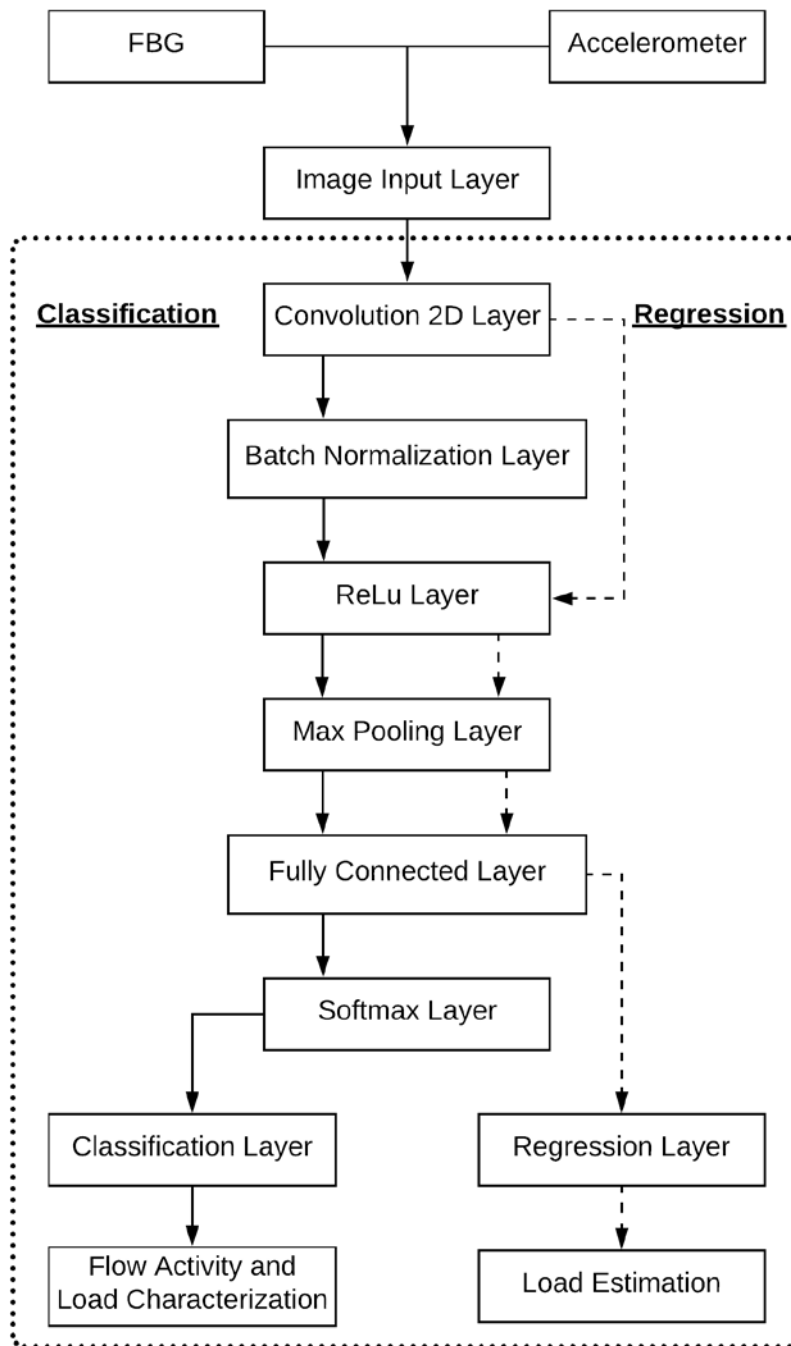
Strain data per observation was trimmed among the three FBG sensors ensuring equal sets. Utilizing FBG strain data's timestamps as reference, accelerometer data were time synchronized accordingly. Measurements from the three FBG sensors were concatenated into a single vector achieving a sequential strain curve according to crowd flow. Due to the various natural ways a volunteer might hold the smartphone in their hands, it is important not to discriminate on the orientation of the activated axis. This was guaranteed by summing, per group size, all the norm of the tri-axis smartphone's accelerometer. The dataset consisted of observations of FBG and accelerometer data with their respective labels.

As mentioned previously, the main advantage and motive for employing CNN is due to its automatic feature extraction. Therefore, raw signal data for FBG and accelerometers are used as 1D image inputs to generate the CNN estimation models. In addition, Fast Fourier Transform (FFT) function is applied on the sensors' raw signal, to be used as input for CNN model generation. To maintain equal assessment, no manually extracted features will be used for SVM model generation with equivalent inputs as that of CNN to be used accordingly.

The CNN optimization solvers used for obtaining optimum classification and regression results were Stochastic Gradient Descent with Momentum (SGDM), Root

Mean Square Propagation (RMSPROP) and ADAM. In the case of SVM, classification and regression kernel functions relied on Quadratic and Cubic Polynomial and Gaussian for highest results, with multi-class SVM employing one-vs-one approach on same kernel function.

The CNN architecture employed for crowd classification and regression is shown in Figure 13. Considered a shallow network, it consists of an image input layer followed by a single convolution layer constituting of 20 filters. When performing a classification, flow path of architecture continues with a batch normalization layer, a ReLu layer, and a max-pooling layer. Final weights are connected at the fully connected layer with a softmax operator, superseded by a classification layer. The network architecture used when performing a regression is similar to that of classification while omitting the batch normalization layer and softmax operator, and substituting the final layer for a regression layer instead. Though CNN has been mainly applied in vision applications with images used as inputs, each observation of our FBG and accelerometer data can be considered as a 1D signal image. Each image's width comprises the width of the signal data vector with a height and channel depth of one.



**Figure 13. CNN network architecture applied for flow type and load characterization classification and load estimate regression models**

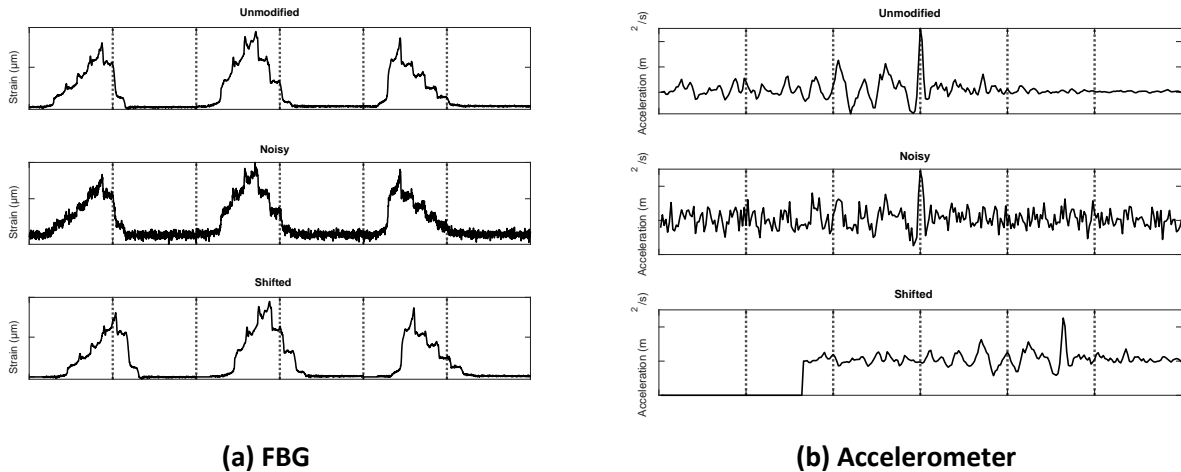
Unbiased evaluation of selected classification and regression models require the random splitting of the dataset into training, validation, and testing sets of percentages (80:10:10) respectively. Training data is used for model training while validation data is used for optimization and tuning of model's parameters. Testing data is treated as unseen data which is then used to evaluate the model's performance and accuracy. Evaluation of classification models is reported through the sum of true positive and true negative over total number of testing samples as a percentage accuracy with higher values indicating a better performance. Evaluation of regression models is reported as the percentage error, which is the difference between estimated and actual output followed by a mean for all observations, were lower values indicate a better performance. In order to verify generalization, the dataset was randomly split and re-trained and tested 25 times. Final assessment of performance was based on the mean and standard deviation of all generated model's resulting accuracies. In order to obtain objectives, each of the sensing instrument's dataset are initially used independently for model generation. FBG and accelerometer models are evaluated to perceive how these sensors may perform individually under crowd attributes classification and regression; flow activity class, load characterization class, and load estimation.

## **B. Noise and Time Shift Invariance**

Testing and evaluating a machine learning model's robustness is vital. The aim of robustness testing is to demonstrate how such models might perform under real world scenarios and unexpected complications. For robustness testing, two main error-inducing factors were tested on the models trained using the original unmodified

training data. As a sensor may suddenly become *noisy* or its data may experience *time shift*, the models were tested on modified data simulating such circumstances.

Simulating noise is implemented via the addition of a white gaussian noise to the sensors' data. Time shift is accomplished at every observation for both sensors through the random shifting along the time axis (maximum of 1 second) and padding with zeros. These factors are applied to both the individual (sensor's raw signal data and re-applying FFT) and fusion testing portion of the dataset as well.



**Figure 14. Strain(a) and acceleration norm (b) between unmodified, noisy, and shifted signal**

An observation sample from the dataset with simulated noise and time shift for both strain and acceleration norm signal is shown in Figure 14. Adding noise to sensors data points, clearly alters some important features within the signal while time-shifting shift these important features along the time axis.

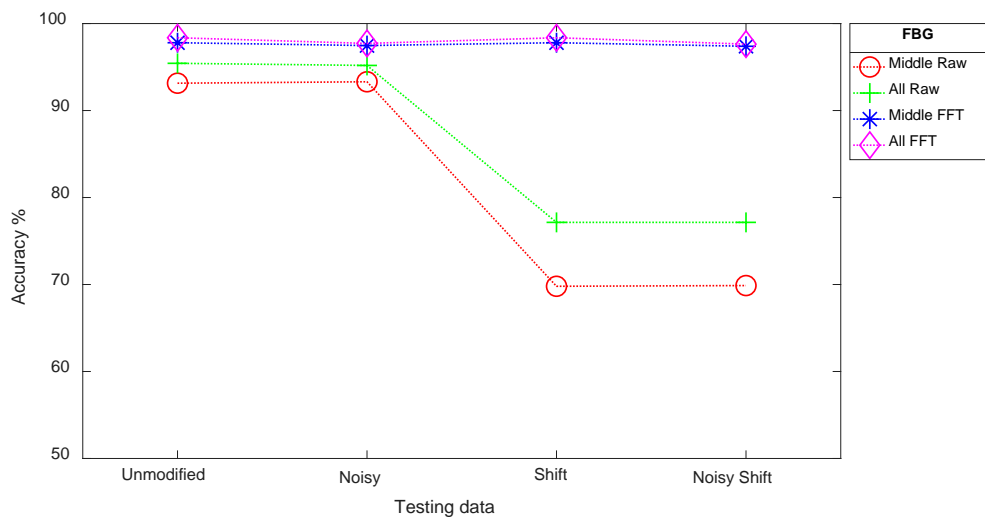
## C. Fusion

### *i. Sensor number fusion*

Three multiplexed FBG sensors in series mounted below the central C-beam of the test bridge measure strain at the three different locations as discussed previously and shown in Figure 5(a). It is important to understand if and how the fusion of number of sensors and distribution affect results in terms of crowd attributes estimation. In this case, focus will be, considering total length and size of test bridge, on whether a single FBG strain sensor is sufficient or on a multitude of matching sensors will return increased effectiveness. The middle sensor installed at the central of the test bridge ought to read the greatest strain exhibited by the structure as the crowd walks on. Along with the middle sensor to perceive how it may fair, all three sensors will be trained and used to classify flow and load characterization and estimate load accordingly, while testing robustness against noise and time shift under both signal types of Raw and FFT and SVM and CNN machine learning approach.

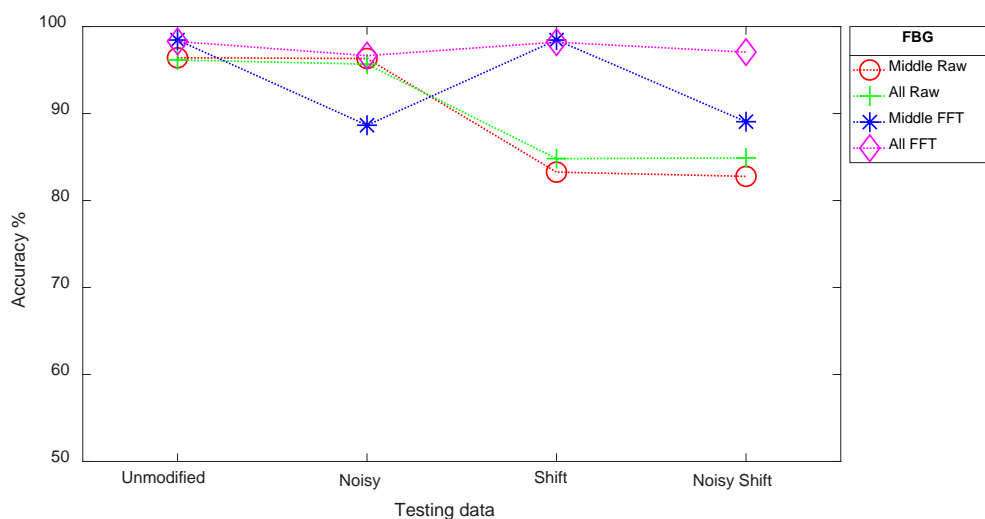
The following figures indicate single and multi-class classification and regression results presenting the trend of either middle or all three FBG sensors when tested.





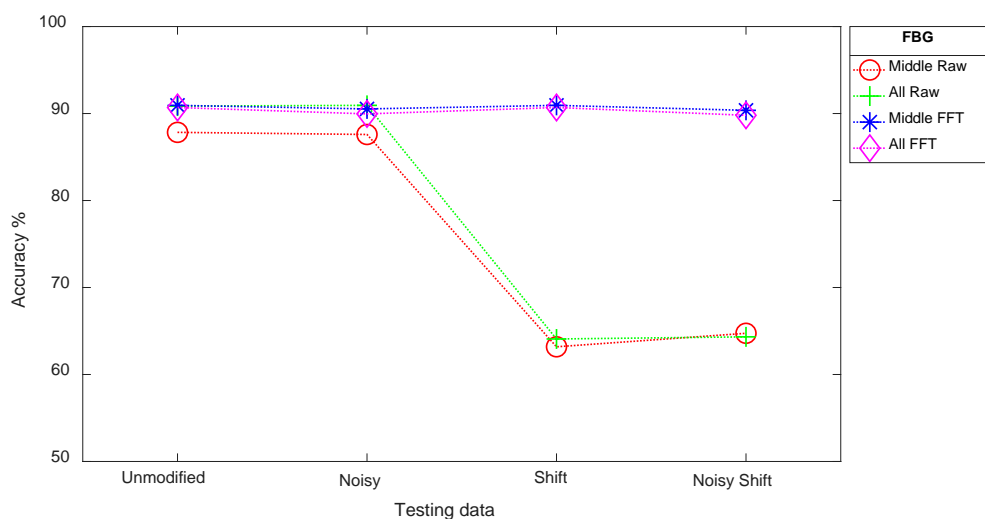
**Figure 15. SVM single class flow activity classification on number of FBG sensors between middle and all and their corresponding signal type**

It can be readily be observed from Figure 15 that all three sensors return higher testing accuracies through SVM than simply using a single middle sensor. Utilizing raw signal type data from showed little to non-existent margin difference between all and middle sensor. However under the case of FFT signal type, a larger difference can be observed, especially when experiencing shift.



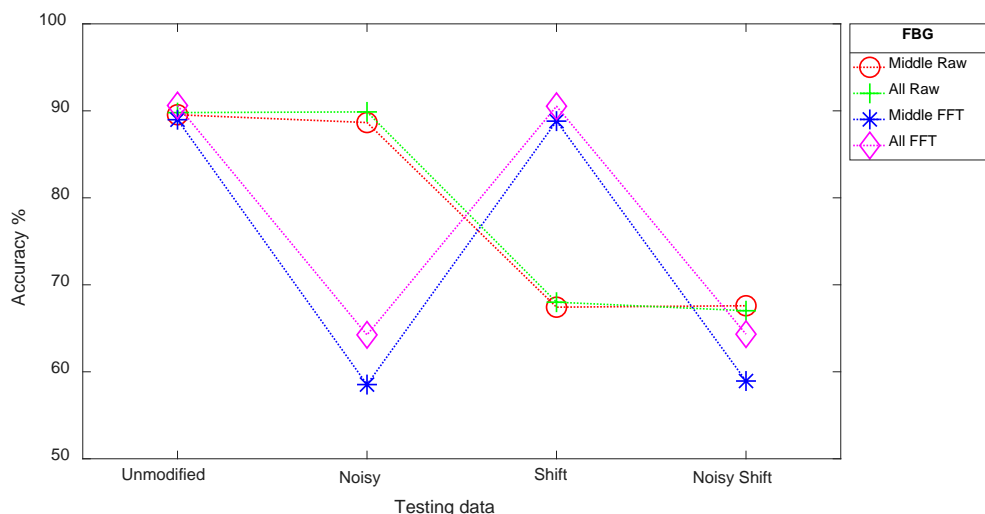
**Figure 16. CNN single class flow activity classification on number of FBG sensors between middle and all and their corresponding signal type**

Figure 16 presents almost no difference between middle and all sensors for unmodified testing accuracies under both signal types. A similar trend can be found for noisy testing, with the exception of middle sensor using FFT signal type returning a lower accuracy. The same observation trend as unmodified testing can be seen for shift testing, with all three sensors outperforming their middle counterpart under raw signal type.



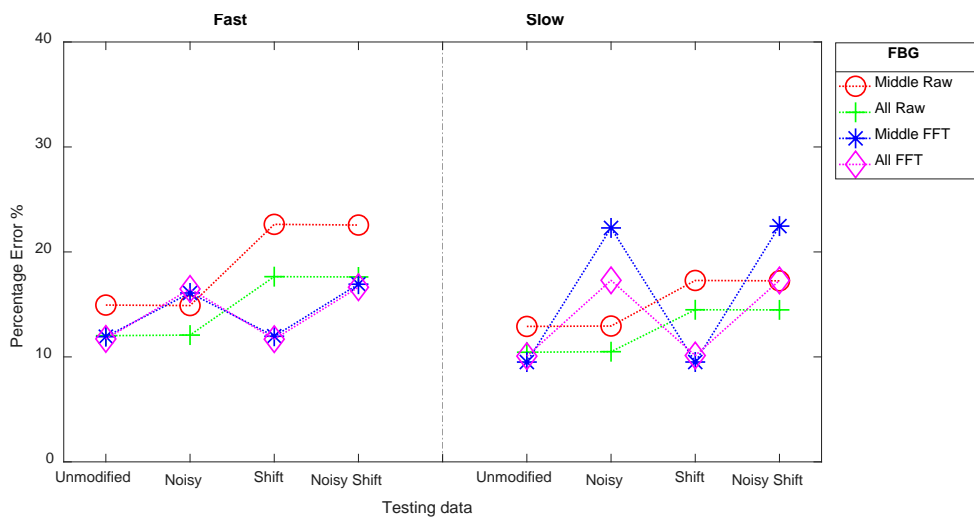
**Figure 17. SVM multi-class flow activity and load characterization classification on number of FBG sensors between middle and all and their corresponding signal type**

It can be observed in Figure 17 both middle and all sensors using FFT signal type return virtually same testing results. For unmodified, noisy, and shift testing scenarios under raw signal type, the three FBG sensors together outperform the stand-alone middle sensor, while returning similar results for noisy shift testing scenario.



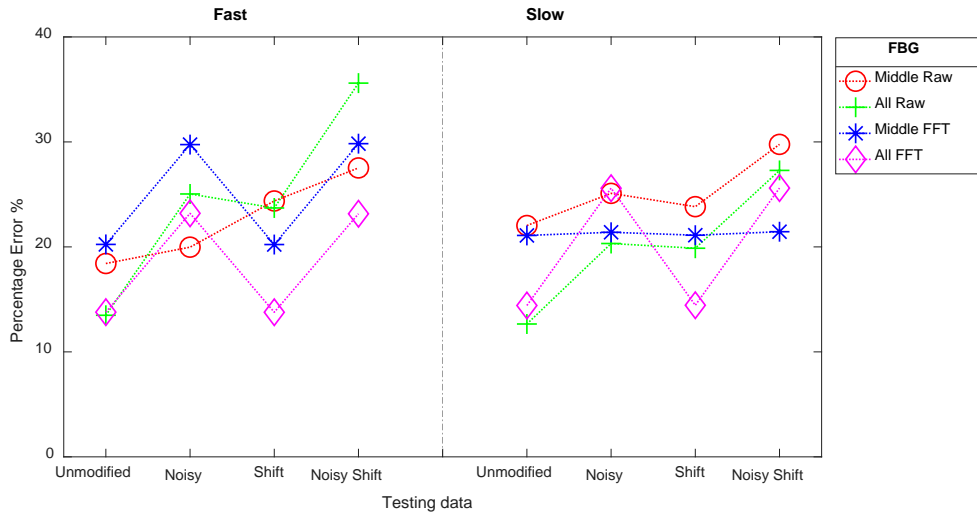
**Figure 18. CNN multi-class flow activity and load characterization classification on number of FBG sensors between middle and all and their corresponding signal type**

Difference between middle and all FBG sensors when using raw signal types, little to no difference between their returned testing accuracies can be observed in Figure 18, with all slightly higher under noisy testing. Ignoring the trend exhibited using FFT signal type for now, the middle sensor returned lower accuracies over all testing scenarios when compared to all three at once.



**Figure 19. SVM load estimation regression on number of FBG sensors between middle and all and their corresponding signal type for both flow activities**

For both flow activities, a similar trend for both middle and all sensors within the same signal type in Figure 19. Under raw signal type, percentage errors from all three sensors are lower than that of just the middle sensor. However under FFT signal type, both sensors number resulted in similar percentage errors for Fast activity, while middle sensor returned higher percentage errors in noise related testing for Slow activity.



**Figure 20. CNN load estimation regression on number of FBG sensors between middle and all and their corresponding signal type for both flow activities**

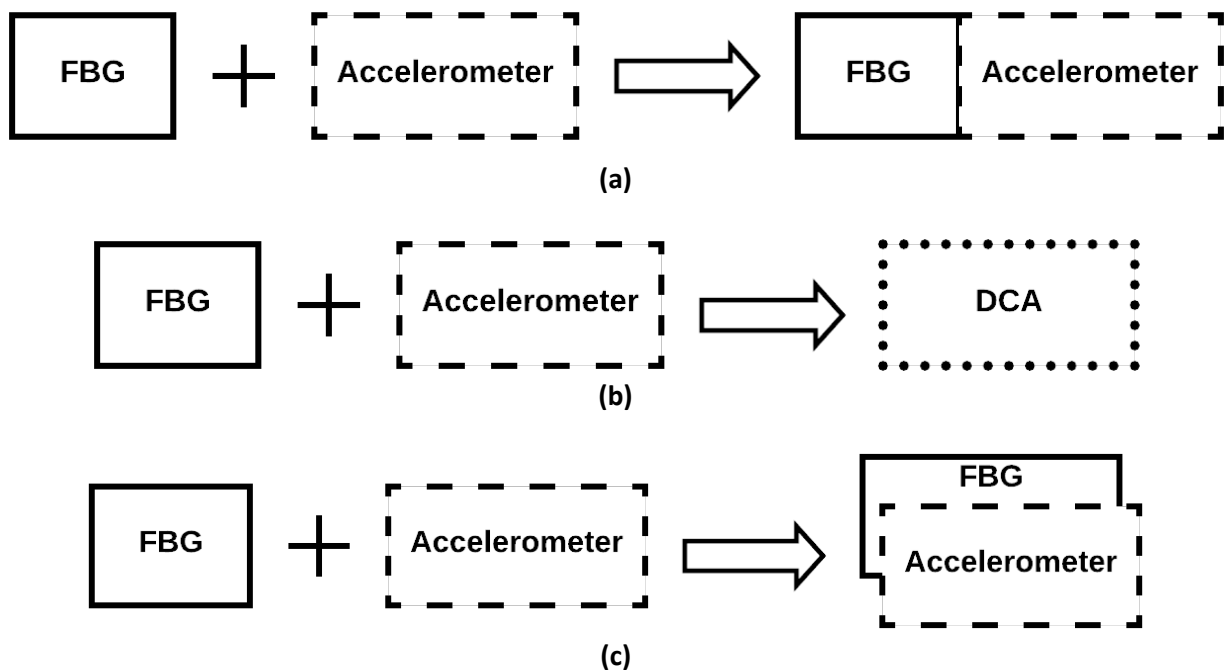
For both flow activities Figure 20, the use of all the three sensors using raw signal type returned generally lower percentage errors. The only exception where middle sensor returned lower percentage error is in the case of noisy testing under Fast activity. Under Fast activity, both middle and all sensor showed a similar trend using FFT signal type with the latter returning lower percentage errors. Yet again under Slow activity, the middle sensor returned lower percentage error for noisy testing, with all three sensors fused returning lower percentage errors for rest of testing.

It can be determined that the all three fused number of FBG strain sensors either generally returned similar or better results than the middle sensor when used alone. Inputs regarding FBG will constitute of the three fused sensors data for further crowd attribute classification and regression tasks.

### *ii. Sensor data fusion*

One of the main objectives, in this work, is to combine and jointly process the data collected from different sensing modalities using machine learning models to better

estimate crowd flow and load attributes. Rather than relying on information from an individual sensor, the data fusion from structural monitoring and wearable sensing yields a smarter representation of the unified structure-crowd system. The successful fusion of strain at the crowd level and acceleration data at the individual level is expected to produce greater robust classification and regression models. Future wise, the fusion may possess advanced capabilities such as differentiating between crowd overloading, environmental loading, and structural weakening. Outputs from fusion models could be utilized to decisively trigger preventative actions with higher and avoid possible incidents and disasters.



**Figure 21. Fusion process at the (a) input level and feature level for (b) SVM and (c) CNN**

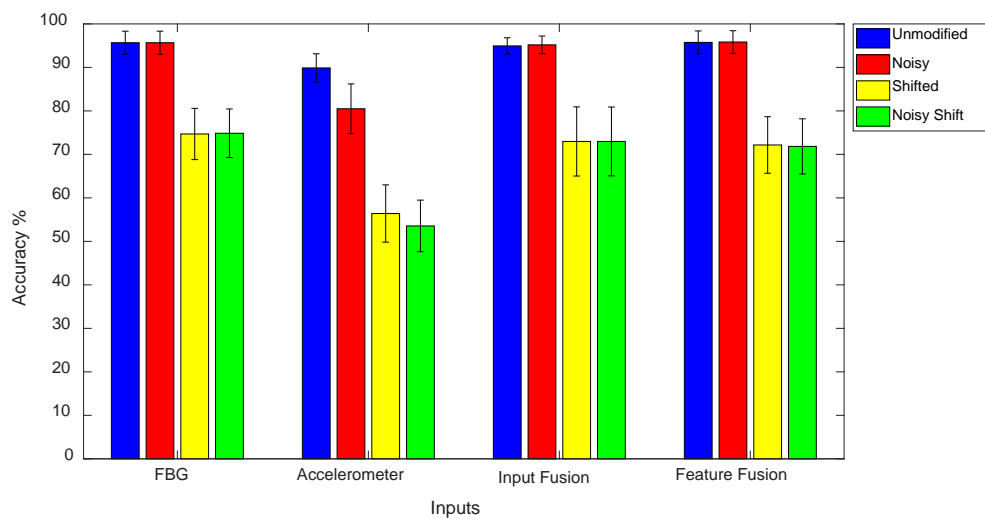
One method of sensor data fusion is achieved by fusion at the input level where each pair of observations from the FBG and accelerometer was merely concatenated

into a single signal vector is shown in Figure 21(a). The newly fused signal vector is employed in the same manner as FBG and accelerometer data was individually used to generate classification and regression models for both SVM and CNN. An alternative sensor data fusion method employed specifically for SVM is *Discriminant Correlation Analysis* (DCA) [77] occurring at the feature level. As result of DCA, a new more discriminative data vector, shown in Figure 21(b), is formed based upon FBG and accelerometer data along with their corresponding labels, which in turn is also used for generation of classification and regression models. A CNN-specific fusion method at the feature level occurs at the convolution layer. The convolution layer is configured to function on a two channels basis. Similar to how individual FBG and accelerometer signal data arrays are input into the CNN, fusion takes place through the concatenation of the individual sensors' data arrays along the 3<sup>rd</sup> dimension, as portrayed in Figure 21(c). Convolution per channel occurs based on the sensor data on each of the 3<sup>rd</sup> dimension. Extracted features per channel are fused as it is forwarded onto the next layer for further processing.

#### **D. SVM Results**

The following figures present several classification and regression testing results plot for a combination of inputs and signal types. Inputs include data for both individual (FBG and Accelerometer) and fusion (input level and DCA feature level). Two plots are presented for each of the input data, based on raw and FFT signal type used. Results are based on inputs were unmodified, noisy, shifted, and noisy shifted testing data was used. Utilized kernels used for models' training and testing include

gaussian, quadratic and cubic polynomial. However, only kernels returning the best accuracies based on unmodified testing data for every input under each of the signal data type are discussed. Due to the different dynamics under each of the two flow data type are discussed. Due to the different dynamics under each of the two flow activities, separately trained and tested regression models' results are presented. In regards to accuracy evaluation, higher scores indicate better performance for classification while lower scores indicate better performance for regression.

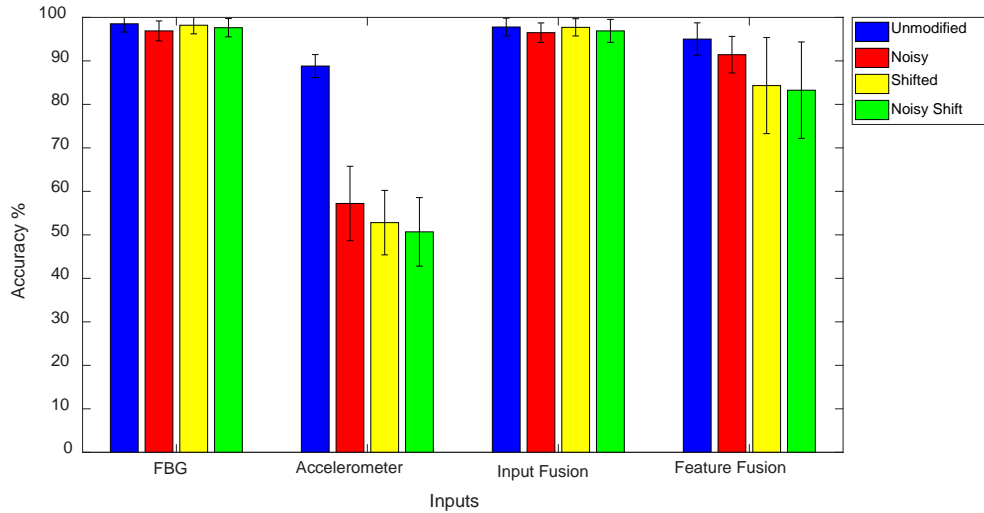


**Figure 22. SVM using raw signal individual and fusion data – Flow activity classification testing accuracies for unmodified, noisy, shifted, and noisy shift data**

Figure 22 presents the SVM testing accuracies of a single class flow activity classification using raw signal type. Plotted unmodified testing accuracy generally for all inputs are at the ballpark of 95%, with the exception of accelerometer at 89%. Both fusion approaches returned similar comparable accuracies to that of FBG individual input, with feature level fusion higher at 95.76%. Noise has little to no effect on accuracies when compared to unmodified accuracy with the exception of accelerometer,

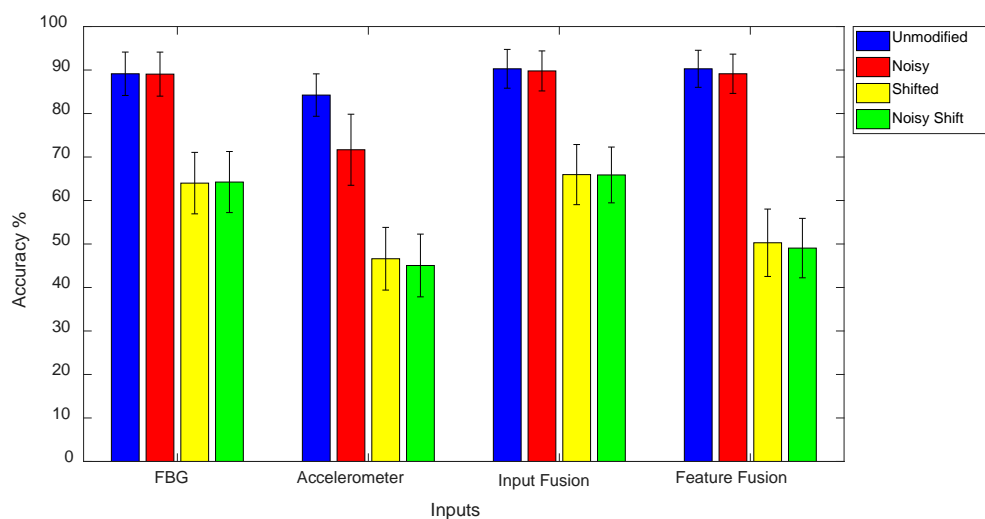


which reduced to 80%. Shift and noisy shift testing reduced accuracies of all inputs to the range of 70%, with accelerometer lower at 50% range.



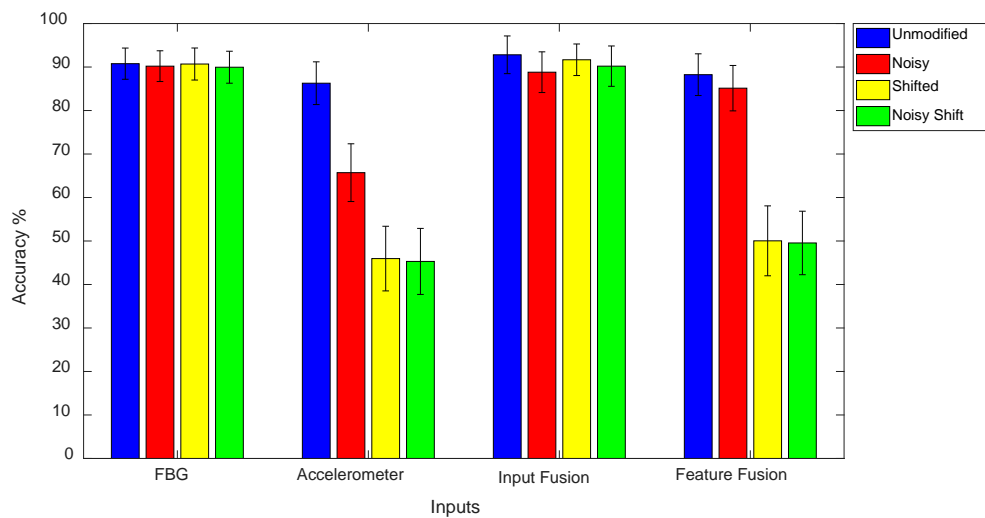
**Figure 23. SVM using FFT signal individual and fusion data – Flow activity classification testing accuracies for unmodified, noisy, shifted, and noisy shift data**

Figure 23 presents the SVM testing accuracies of a single class flow activity classification using FFT signal type. Plotted unmodified testing accuracies of all inputs in order are 98.5%, 88.8%, 97.8%, and 95% respectively. Noisy testing shows drop in accuracy with varying margins across the inputs, with accelerometer at 57% accuracy. Shifted and noisy shift testing resulted in little to no reduction in accuracies for FBG and input level fusion. However, a larger margin drop can be seen for accelerometer and feature level fusion, the former being the lowest in accuracy. Input level fusion is observed to perform better than feature level fusion.



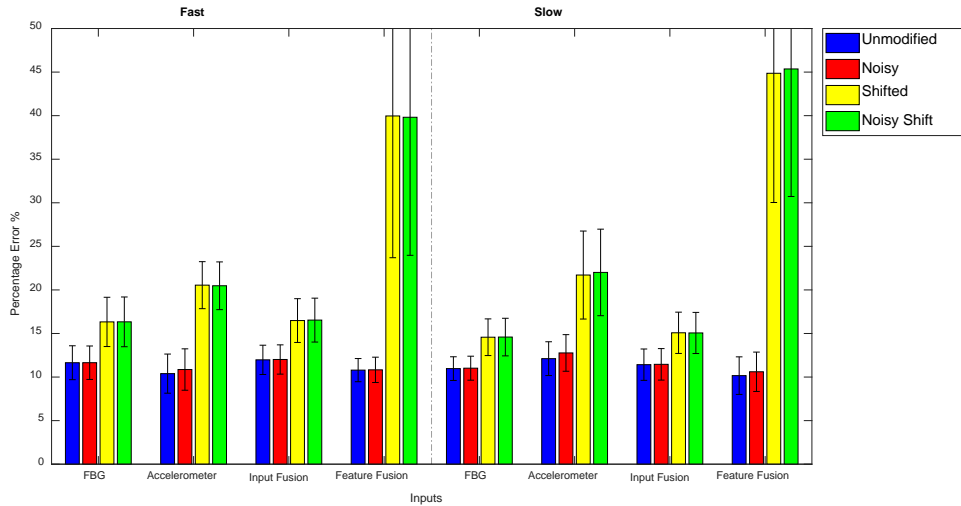
**Figure 24. SVM using raw signal individual and fusion data – Flow activity and load characterization classification testing accuracies for unmodified, noisy, shifted, and noisy shift data**

Figure 24 presents the SVM testing accuracies of a multi class flow activity and load characterization classification using raw signal type. Plotted unmodified testing accuracy for FBG and accelerometer are 89% and 84% respectively. Input and feature level fusion returned slightly higher accuracies to that of FBG at 90%. Noisy testing barely lowered accuracies across majority of inputs with accelerometer having the largest margin at 71%. In a similar case but with greater effect, accuracies of all inputs drastically suffer when testing using shifted and noisy shift with input level fusion maintaining highest accuracies around 66%.



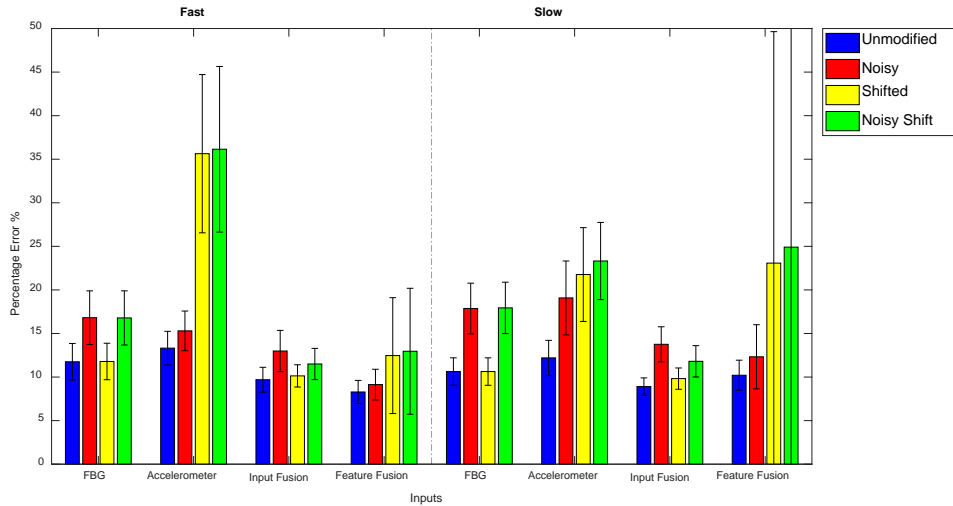
**Figure 25. SVM using FFT signal individual and fusion data – Flow activity and load characterization classification testing accuracies for unmodified, noisy, shifted, and noisy shift data**

Figure 25 presents the SVM testing accuracies of a multi class flow activity and load characterization classification using FFT signal type. Plotted unmodified testing accuracy for FBG and accelerometer are around 90% and 86% respectively. Input level fusion returned accuracies slightly higher accuracy at around 92% with DCA feature level fusion at 88%. Noisy testing had varying returned accuracies across all inputs. Interestingly FBG and input level fusion yielded accuracies for shifted and noisy shift testing similar to that of their unmodified testing, with the latter being higher. However, along with the largest standard deviation accelerometer yielded accuracies of 46%, with DCA feature level fusion at 50%.



**Figure 26. SVM using raw signal individual and fusion data – Load estimation regression testing percentage errors for unmodified, noisy, shifted, and noisy shift data**

Figure 26 presents the SVM testing percentage errors of load estimation for each of the two flow activities, Fast and Slow, using raw signal type. Unmodified percentage errors for all inputs data presented are generally below 12%, with DCA feature fusion lowest under Slow activity at around 10%. Noise testing either slightly increase or maintained percentage errors for all inputs. Shift and Noisy Shift testing had a large increase in percentage errors and standard deviation across all inputs, with DCA feature fusion returning with the highest margins followed by accelerometer input.

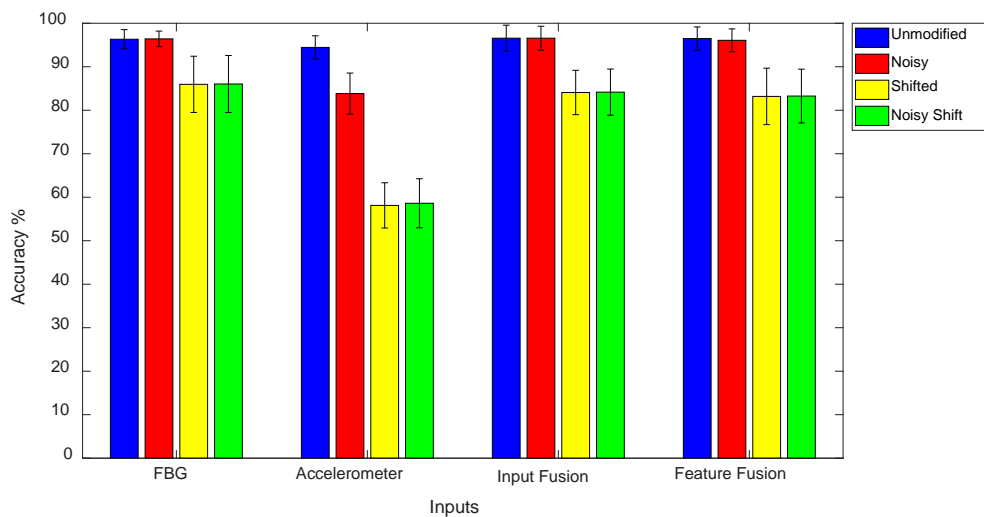


**Figure 27. SVM using FFT signal individual and fusion data – Load estimation regression testing percentage errors for unmodified, noisy, shifted, and noisy shift data**

Figure 27 presents the SVM testing percentage errors of load estimation for each of the two flow activities, Fast and Slow, using FFT signal type. For both flow activities, unmodified percentage errors for FBG and accelerometer presented are about 10-11% and 12-13% accordingly. Fusion at the input level and DCA feature level returned lower accuracies than their individual counterparts at around 8-10% respectively. Noise can be seen to increase percentage error with both fusion approaches having a lower percentage error when compared to their individual inputs. Shift had little to no change in percentage error for FBG input level feature input, while increasing for Accelerometer and feature level fusion with larger standard deviation, greater for the former. Noisy shift can be observed to increase results with largely varying margins. Both accelerometer and feature fusion have the highest percentage error along with larger standard deviation.

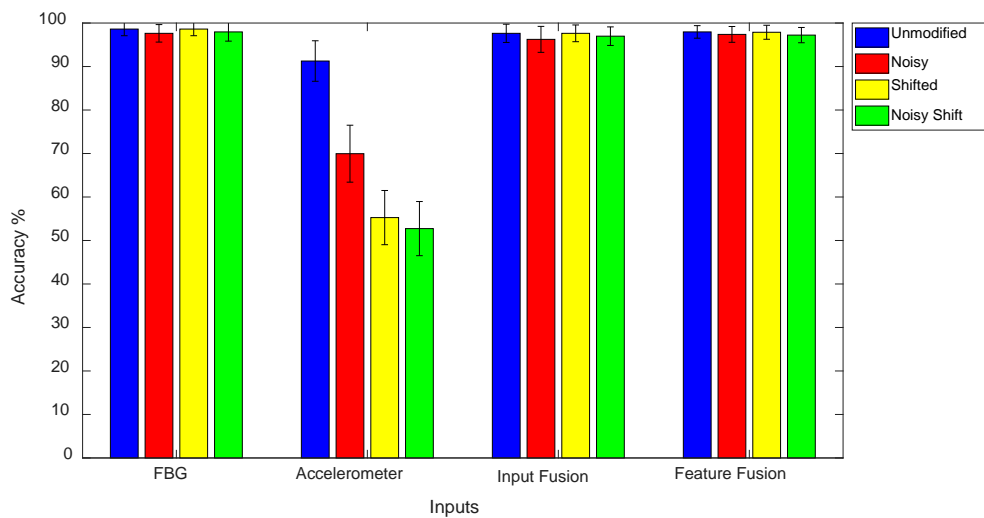
## E. CNN Results

The following figures presents several classification and regression testing results plot for a combination of inputs and signal types. Inputs include data for both individual (FBG and Accelerometer) and fusion (input level and feature level). Two plots are presented for each of the input data, raw and FFT signal type, used. Plotted results are based on inputs were unmodified, noisy, shifted, and noisy shifted testing data was used. Utilized solvers used for models' training and testing include SGDM, RMSPROP, and ADAM. However, only solvers returning best unmodified testing accuracies for every input under each of the signal data type are discussed. Due to the different dynamics under each of the two flow activities, separately trained and tested regression models' results are presented. In regards to evaluation, higher scores indicate better performance for classification while lower scores indicate better performance for regression.



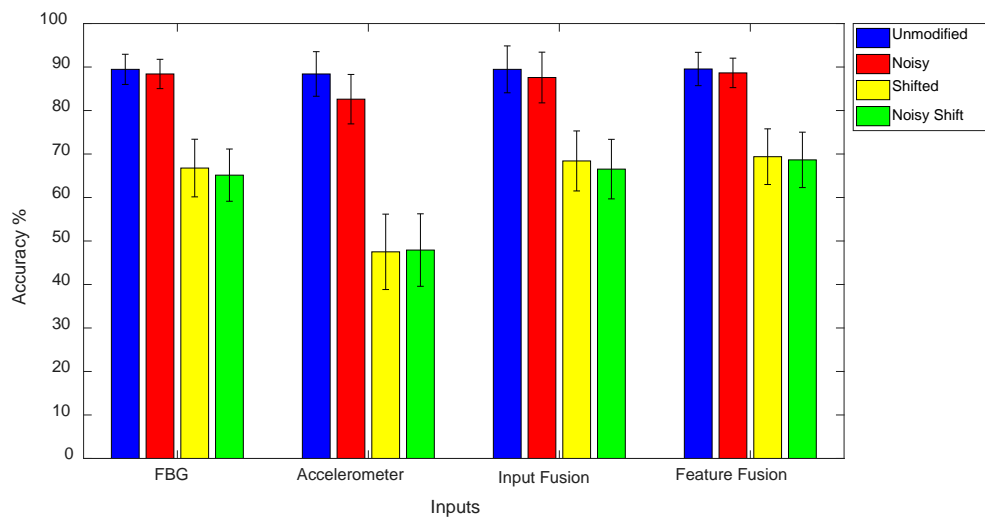
**Figure 28. CNN using raw signal individual and fusion data – Flow activity classification testing accuracies for unmodified, noisy, shifted, and noisy shift data**

Figure 28 presents the CNN testing accuracies of a single class flow activity classification using raw signal type. Plotted unmodified testing accuracy for FBG and accelerometer are at 96% and 94% respectively, with both fusions marginally higher. Noisy testing can be seen to drop accelerometers accuracy to 83%, with no effect on other inputs. Shift and Noisy shift reduce accuracy with FBG around 86% with accelerometer returning about 58%, while both fusion at around 83-84%.



**Figure 29. CNN using FFT signal individual and fusion data – Flow activity classification testing accuracies for unmodified, noisy, shifted, and noisy shift data**

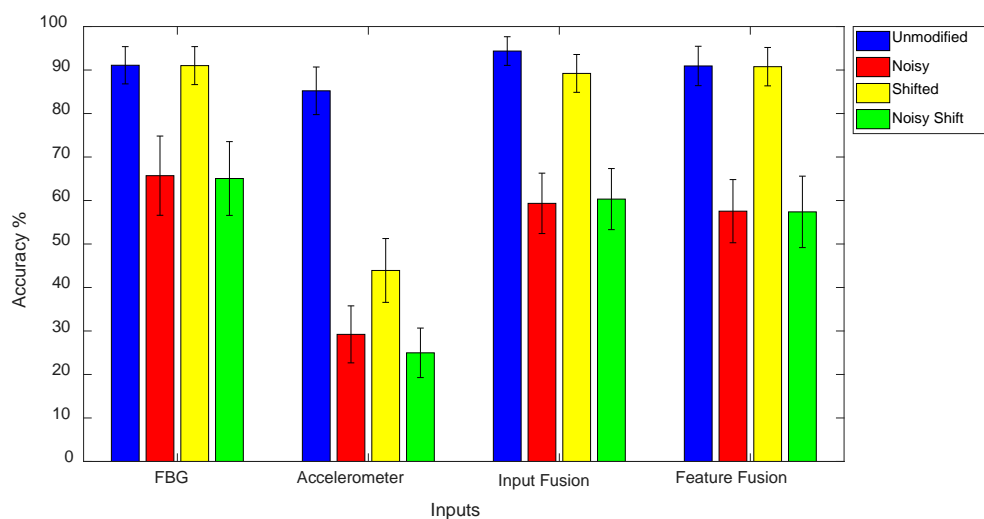
Figure 29 presents the CNN testing accuracies of a single class flow activity classification using FFT signal type. FBG and accelerometer yielded unmodified testing accuracies of 98% and 91% accordingly. Fusion at both the input and feature level returned accuracies of 97-98%. Testing using Noisy data had little to no effect on accuracy drop with feature fusion maintaining the marginally higher results. Testing on shifted data yielded accuracies of about 98% for FBG and both fusion approaches used, with accelerometer accuracy dropping to 55%. Noisy shift had little to no change on accuracies when compared to shifted testing accuracies across all inputs.



**Figure 30. CNN using raw signal individual and fusion data – Flow activity and load characterization classification testing accuracies for unmodified, noisy, shifted, and noisy shift data**

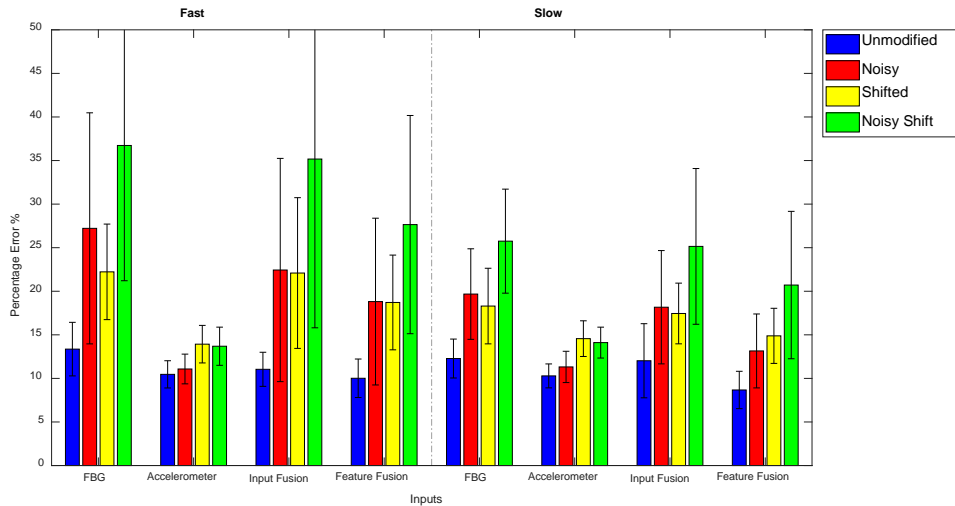
Figure 30 presents the CNN testing accuracies of a multi class flow activity and load characterization classification using raw signal type. FBG yielded an accuracy of 89% for unmodified testing, with accelerometer returning 88%. Both fusion approaches for unmodified testing resulted in accuracies around 89%. Accuracies are reduced by a slight margin during Noisy testing, with the exception of with a margin of 6% for accelerometer. Accuracies during Shift and Noisy shift is drastically lower between the range of 47% and 69% for accelerometer and feature level fusion accordingly.





**Figure 31. CNN using FFT signal individual and fusion data – Flow activity and load characterization classification testing accuracies for unmodified, noisy, shifted, and noisy shift data**

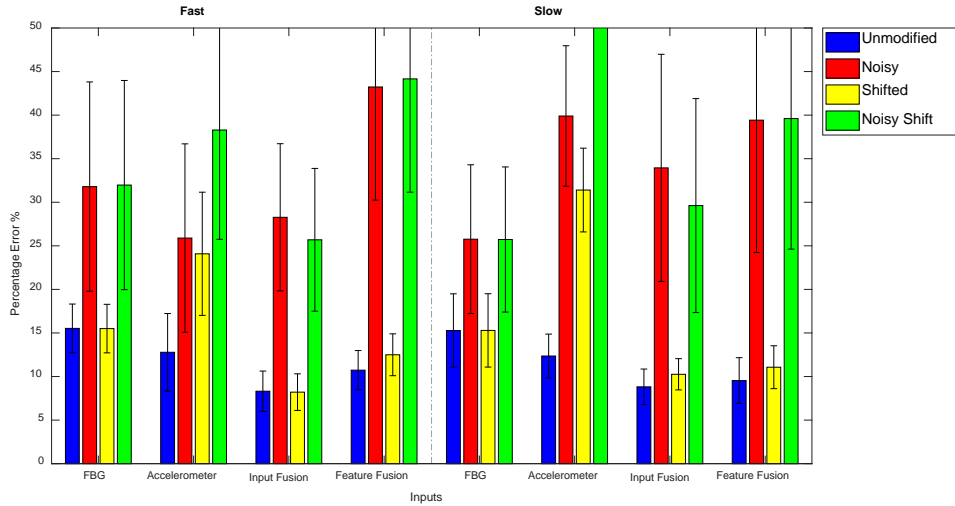
Figure 31 presents the CNN testing accuracies of a multi class flow activity and load characterization classification using FFT signal type. Returned unmodified testing accuracies for FBG and accelerometer are at about 91% and 85% respectively. Input level and feature level fusion yielded accuracies of 94% and 91% accordingly. With at exception of accelerometer at same result, Noisy testing drastically lowered accuracy to a low of 29% for accelerometer, with both fusion approaches at around 58%. Testing on Shifted data resulted in accuracies of 91% for FBG, 44% for accelerometer, and 89-90% for both fusion inputs. Testing Noisy shift further reduces accuracies the ballpark of 60-65%, with accelerometer further lower at 25%.



**Figure 32. CNN using raw signal individual and fusion data – Load estimation regression testing percentage errors for unmodified, noisy, shifted, and noisy shift data**

Figure 32 presents the CNN testing percentage errors of load estimation for each of the two flow activities, Fast and Slow, using raw signal type. Under both flow activities, unmodified percentage errors for all inputs data presented are in the range between 9-13%, with feature fusion being the lowest at both flow activities. Testing on Noise data increases FBG percentage error to relatively 27% for Fast activity and 20% for Slow activity, with accelerometer marginally returning about 1% increase. Both input and feature level fusion return lower percentage errors than FBG yielding 22% and 18% under Fast activity and 18% and 13% accordingly. Testing on Shifted data returned a substantial increase in percentage errors, with accelerometer relatively maintaining a lower result when compared to the other input data. Noisy Shift, with the exception of accelerometer also, further increases percentage errors to generally highest values returned per input. Both fusion techniques are generally equal or lower than FBG when comparing their shift and noisy shift percentage errors. With the exclusion of

accelerometer input, other inputs return accuracies on noisy and shift related testing with extreme standard deviations.



**Figure 33. CNN using FFT signal individual and fusion data – Load estimation regression testing percentage errors for unmodified, noisy, shifted, and noisy shift data**

Figure 33 presents the CNN testing percentage errors of load estimation for each of the two flow activities, Fast and Slow, using FFT signal type. Under both flow activities, unmodified percentage errors for FBG and accelerometer inputs data presented are between 12% and 15%. Input level and feature level fusion returned percentage error between 8% and 11% accordingly under both flow activities. FBG and accelerometer percentage error increases to approximately 32% and 26% with noise testing under Fast flow activity. Both fusion approaches results increase with input level fusion better performing. Under Slow flow activity, FBG percentage error is lower at 26%, while accelerometer is higher at 40% yet input level fusion being the better fusion technique as well at 34%. For Shift testing under both flow activities, FBG yielded a percentage error of 15% with accelerometer returning 24% and 31% accordingly. Both

fusions resulted in a lower percentage errors for both flow activities when compared to their individual input counterparts, with input level fusion being the lowest. Noisy shift can be seen to drastically increase percentage errors with large margin peaking at more than 50% for accelerometer under Slow flow activity. Input level fusion returns percentage error lower than both individual inputs for Fast activity. Both noise related testing returns large standard deviations to their inputs.

## **F. Discussion**

It is to be noted that all classification and regression results are based on the mean of generalized models' results repeated over the course of 25 times, each with a random split of training and testing data. The final results displayed per machine learning approach for each of the signal type used is based on the kernel/solver that returned results with overall better unmodified testing performance. It is important when fusion at either level perform or return values similar or higher than the lowest performing individual input, otherwise it is to be considered to fail its purpose.

### *i. Single Class – Flow Activity Classification*

Returning testing accuracy results of single class flow activity classification of both signal types can be seen in Figure 22 and Figure 23 for SVM, and Figure 28 and Figure 29 for CNN.

Approaching though SVM and using raw signal type allows for classification of flow activity with high accuracies. While accelerometer results are lower than those of FBG, fusion at the input level maintains similar results as that of FBG with feature level fusion achieving a slightly higher accuracy on unmodified and noisy testing data.

Applying shift related testing data drop returned accuracies for both FBG and accelerometer, with both fusions having similar results as FBG. Using FFT signal type showed an overall increase in accuracies for all inputs across the various testing data, with the exception of accelerometer which maintained a relative similar accuracy. Varying reduction in accuracies occurs when testing noise across all inputs with accelerometer having the largest margin drop. Shift related testing also fairs well across inputs with the exception on accelerometer. Input level fusion maintains being the better fusion approach.

Approaching though CNN and using raw signal type returns high accuracies as well for flow activity classification. Both FBG and accelerometer are of high accuracies, yet both fusion approaches achieved slightly higher accuracies. Noise testing had little to no effect on accuracies, with the exception of accelerometer input where accuracies reduced. Shift and noisy shift further reduce for FBG and both fusion levels to similar levels, with accelerometer having the largest reduction in accuracy. Using FFT signal type, accelerometer input yet showed high returns when tested using unmodified, but dropped when testing noise and further significantly under shift and noisy shift testing. FBG and both level fusion returned similar high results with noise related testing marginally below. However, feature level fusion returned results are as high as input level fusion with close matching accuracies through all subjected testing data.

*ii. Multi Class – Flow Activity and Load Characterization  
Classification*

Testing accuracy results of multi class flow activity and load characterization classification of both signal types is displayed in Figure 24 and Figure 25 for SVM, and in Figure 30 and Figure 31 for CNN.

Utilizing raw signal type on SVM, unmodified testing results for both fusion techniques was higher than accelerometer and slightly above FBG input. Noise testing had no effect on FBG and little reduction effect on both levels of fusion, while accuracies dropped for accelerometer. Shift and noisy shift significantly reduced accuracies with input fusion returning highest results, though slightly, among the inputs. However, feature fusion returned shift related testing results lower than the individual inputs. Through same machine learning approach but using FFT signal type, accuracies of FBG and accelerometer unmodified testing maintained high values, with input level fusion returning slightly higher accuracy values. Noise testing reduced accuracy of varying levels with accelerometer being the lowest. A similar observance under noise and noisy shift testing with input fusion being the highest. However, accelerometer and feature level fusion accuracies dropped significantly, with the latter slightly higher.

CNN approach using raw signal type returned high unmodified testing accuracies for individual inputs, with both input and feature fusion similar, with the former being the highest overall. Little reduction in accuracies with varying small margins across inputs is observed when subjected to noise testing. Shift related testing significantly reduced accuracies among the individual inputs, while both input and feature fusion returned slightly higher accuracies when compared to the highest

individual input, FBG. Approaching through the use of FFT signal type returned high unmodified and similar shift testing accuracies with the exception of accelerometer input, and input level fusion returning highest unmodified testing value. Noise related testing dropped accuracies significantly, accelerometer being the worst. With the exception of shift testing input level fusion returned the marginally higher accuracies when compared to feature level fusion.

### *iii. Load Estimation Regression*

Testing accuracies results of load estimation regression using both signal types are shown in Figure 26 and Figure 27 for SVM, and in Figure 32 and Figure 33 for CNN.

Estimating load through SVM while using raw signal type returned low percentage errors with noise having little increase. Feature level fusion employed is observed to have a better effect under both flow activity. However, shift related testing resulted in increase in percentage errors of individual inputs with input level fusion performing similar to FBG while feature level fusion failing with large standard deviations. Discrepancy is observed during shift related testing subjected on feature level fusion under both flow activity, were returned values are higher than the individual inputs of comparable testing. Using FFT signal type resulted in low percentage errors as well with both fusions returning lower unmodified testing values than their individual inputs respectively. Noise testing affected and increased individual inputs, while having a smaller increase for both fusion, feature level being the lowest. Shift related testing had a significant increase in percentage error for accelerometer

under Fast flow activity, yet input level fusion returned lower results across all inputs for both flow activities.

When estimating using CNN machine learning approach and raw signal type, individual inputs returned low percentage errors with feature level returning slightly lower values. Noise and shift had smaller effect on accelerometer data when compared to others with largely varying margins. With the exception of accelerometer, noisy shift affected greatly all inputs with the greatest standard deviations. Using FFT signal type returned unmodified testing percentage errors of both fusion lower than their comparable individual inputs on both flow activities, input level fusion being the lower. Noise and noisy shift increase drastically percentage errors of FBG and accelerometer, hence both fusion techniques suffer with large standard deviations. Yet fusion approaches are still lower than accelerometer under Slow flow activity. Shift has little to no effect on FBG input with a larger change observed for accelerometer input. Both fusion approaches return lower percentage errors than FBG. Discrepancy is observed during noise related testing subjected on feature level fusion under Fast flow activity, were returned values are higher than the individual inputs of comparable testing.

#### *iv. Discrepancy analysis*

Two regression cases of inconsistency in fusion results were observed. Using raw signal type on SVM load estimation, Figure 26, caused feature level fusion to return higher percentage errors than the individual inputs when subjected to shift and noisy shift testing under both flow activities. Another case is when using FFT signal type on CNN for load estimation, Figure 33, caused input level fusion to return higher



percentage errors than the individual inputs when subjected to noise and noisy shift testing under Fast flow activity

In order to investigate underlying cause for such failure of fusion at their respective cases, we need to understand how the trained model deals with such inputs. In regards to SVM, it is suspected due to the dimensionality reduction feature of DCA under this specific circumstance returned poor results. FBG and accelerometer input individually had 3000 and 300 features respectively. DCA works by extracting and correlating discriminant features based on their classes. The DCA process resulted in fusion input training data consisting of 6 features only, substantially reducing dimensionality hence vital information. This may lead to possible mis-estimation of crowd load.

In regards to CNN, it is suspected due to the current nature of defined CNN architecture in terms of how the convolution layer based on its parameters such as filter and stride size may not properly discriminate accelerometer features due to the dimensionality difference in terms of both axes. Information is also lost at the ends of the input image, with added loss from pooling layer.

#### ***v. Optimal Results***

Following tables present the optimal results of single and multi-class classification, and regression, along with their standard deviation. Shown results are based on SVM and CNN machine learning applied on the individual and fusion inputs on both raw and FFT signal type, indicating most effective approach in tackling said

monitoring application taking into consideration performance evaluation on the several testing cases of noise and shift.

**Table 4: Optimal testing accuracy results of single class flow activity classification per input data and signal type**

	Solver	Testing Accuracy			
		Unmodified	Noisy	Shifted	Noisy Shift
<b>Training Data - Raw Signal</b>					
<b>FBG</b>	CNN-SGDM	96.33	96.41	85.96	86.04
<b>Accelerometer</b>	CNN-ADAM	94.45	83.84	58.12	58.61
<b>Input level fusion</b>	CNN-ADAM	96.57	96.57	84.08	84.16
<b>Feature level fusion</b>	CNN-ADAM	96.49	96.08	83.18	83.27
<b>Training Data - FFT Signal</b>					
<b>FBG</b>	CNN-ADAM	98.61	97.63	98.61	97.96
<b>Accelerometer</b>	CNN-SGDM	91.27	69.96	55.27	52.73
<b>Input level fusion</b>	SVM-Quad	97.80	96.49	97.71	96.90
<b>Feature level fusion</b>	CNN-RMSPROP	97.96	97.39	97.88	97.22

**Table 5: Standard deviation of the optimal testing accuracy results of single class flow activity classification per input data and signal type**

	Standard Deviation			
	Unmodified	Noisy	Shifted	Noisy Shift
<b>Training Data - Raw Signal</b>				
<b>FBG</b>	2.20	1.79	6.46	6.55
<b>Accelerometer</b>	2.67	4.71	5.21	5.64
<b>Input level fusion</b>	2.99	2.75	5.10	5.31
<b>Feature level fusion</b>	2.67	2.63	6.49	6.18
<b>Training Data - FFT Signal</b>				
<b>FBG</b>	1.53	2.01	1.53	2.12
<b>Accelerometer</b>	4.66	6.55	6.23	6.23
<b>Input level fusion</b>	2.03	2.24	1.98	2.64
<b>Feature level fusion</b>	1.44	1.82	1.61	1.76

**Table 6: Optimal testing accuracy results of multi class flow activity and load characterization classification per input data and signal type**

	Solver	Testing Accuracy			
		Unmodified	Noisy	Shifted	Noisy Shift
<b>Training Data - Raw Signal</b>					
<b>FBG</b>	CNN-SGDM	89.47	88.41	66.78	65.14
<b>Accelerometer</b>	CNN-SGDM	88.41	82.61	47.51	47.92
<b>Input level fusion</b>	SVM-Quad	90.29	89.80	65.96	65.88
<b>Feature level fusion</b>	CNN-ADAM	89.55	88.65	69.39	68.65
<b>Training Data - FFT Signal</b>					
<b>FBG</b>	Gaussian	90.78	90.20	90.69	89.96
<b>Accelerometer</b>	Gaussian	86.29	65.71	45.96	45.31
<b>Input level fusion</b>	SVM-Quad	92.82	88.82	91.67	90.20
<b>Feature level fusion</b>	SVM-Quad	88.24	85.14	50.04	49.55

**Table 7: Standard deviation of the optimal testing accuracy results of multi class flow activity and load characterization classification per input data and signal type**

	Standard Deviation			
	Unmodified	Noisy	Shifted	Noisy Shift
<b>Training Data - Raw Signal</b>				
<b>FBG</b>	3.47	3.37	6.63	6.01
<b>Accelerometer</b>	5.13	5.69	8.67	8.33
<b>Input level fusion</b>	4.46	4.60	6.91	6.41
<b>Feature level fusion</b>	3.83	3.39	6.40	6.37
<b>Training Data - FFT Signal</b>				
<b>FBG</b>	3.59	3.53	3.68	3.68
<b>Accelerometer</b>	4.91	6.64	7.43	7.59
<b>Input level fusion</b>	4.33	4.68	3.63	4.64
<b>Feature level fusion</b>	4.80	5.22	8.04	7.30

**Table 8: Optimal testing percentage error results of load estimation regression per input data and signal type for both flow activities; Fast and Slow**

	Solver	Percentage Error							
		Fast				Slow			
		Unmodified	Noisy	Shifted	Noisy Shift	Unmodified	Noisy	Shifted	Noisy Shift
<b>Training Data - Raw Signal</b>									
<b>FBG</b>	SVM-Cubic	11.64	11.64	16.33	16.33	10.97	11.02	14.57	14.59
<b>Accelerometer</b>	CNN-ADAM	10.46	11.08	13.93	13.69	10.29	11.32	14.56	14.11
<b>Input level fusion</b>	SVM-Cubic	11.97	12.01	16.48	16.53	11.42	11.46	15.08	15.06
<b>Feature level fusion</b>	CNN-ADAM	10.02	18.82	18.71	27.65	8.67	13.15	14.88	20.71
<b>Training Data - FFT Signal</b>									
<b>FBG</b>	SVM-Quad	11.75	16.81	11.79	16.78	10.63	17.86	10.63	17.94
<b>Accelerometer</b>	SVM-Quad	13.32	15.30	35.63	36.14	12.20	19.08	21.76	23.32
<b>Input level fusion</b>	SVM-Quad	9.69	12.99	10.13	11.50	8.90	13.76	9.82	11.80
<b>Feature level fusion</b>	SVM-Quad	8.28	9.13	12.46	12.96	10.20	12.32	23.07	24.91

**Table 9: Standard deviation of the optimal testing percentage error results of load estimation regression per input data and signal type for both flow activities; Fast and Slow**

	Standard Deviation							
	Fast				Slow			
	Unmodified	Noisy	Shifted	Noisy Shift	Unmodified	Noisy	Shifted	Noisy Shift
<b>Training Data - Raw Signal</b>								
<b>FBG</b>	1.95	1.92	2.83	2.85	1.36	1.38	2.10	2.16
<b>Accelerometer</b>	1.56	1.71	2.15	2.19	1.37	1.79	2.05	1.77
<b>Input level fusion</b>	1.68	1.68	2.51	2.52	1.80	1.81	2.37	2.36
<b>Feature level fusion</b>	2.21	9.57	5.44	12.52	2.14	4.24	3.16	8.46
<b>Training Data - FFT Signal</b>								
<b>FBG</b>	2.10	3.08	2.09	3.11	1.57	2.91	1.58	2.95
<b>Accelerometer</b>	1.93	2.28	9.08	9.50	2.01	4.24	5.38	4.43
<b>Input level fusion</b>	1.43	2.37	1.28	1.79	1.00	2.01	1.22	1.80
<b>Feature level fusion</b>	1.33	1.76	6.65	7.23	1.73	3.68	26.57	28.71

Focusing on the application of single class flow activity, approaching through CNN using raw signal type returned achieved higher results. A similar case can be said when using FFT signal type, with the exception at the input level fusion where SVM returns were slightly higher. Classifying a multi-class for both flow activity and load characterization may be intricate, however high results were still generally achievable. Using raw signal type, CNN approach was used achieving the optimum results with the exception at input level fusion where SVM was used instead. Using FFT signal type, SVM approach was used as it achieved better results. With the exception of accelerometer and feature level fusion under raw signal type for load estimate regression were CNN returned lower results, approaching with SVM returned the lower

percentage errors. When estimating using FFT signal type, SVM was of better performance with lower percentage errors.

Above tables are based on best overall performance between the models, application wise. However there are certain scenarios where yet still they perform poorly. Results returned by accelerometer input for single class flow activity classification when subjected to shift related testing are low. A similar case can be said for multi class flow activity and load characterization classification using FFT signal type. In addition, feature fusion in shift related testing suffers with large reduction in accuracies. However under raw signal type, shift related accuracies drop with a significant margin across all inputs. For the application of load estimation, large percentage errors are reported under raw signal type for feature fusion during noisy shift testing. In addition, under FFT signal type large percentage errors are returned during shift testing for accelerometer input.

Results returned by the SVM model may be due to reliance of key support vector points in both FBG and Accelerometer data, relying on a just a subset of data points, converging to a global minimum. As opposed to, the main weakness of CNN being part of the neural network family, is that they require large amounts of data to return optimum accuracies since initializing is with unknown starting weights. The SVM and CNN results discussed in this thesis are to rather concur concept system effectiveness with high accuracies under both machine learning approaches. In addition, the results show how a conventionally used machine learning approach, SVM, and an innovative approach, CNN, tackles such application along with sensor data fusion.

## CHAPTER VII

### CONCLUSION AND FUTURE WORKS

To tackle the possible dangers that may arise from large crowd gatherings and high crowd density flow on a bridge structure, the thesis aims to offer an innovative crowd monitoring solution with instrumental and machine learning novelties. The conception is based on employing FOS FBG sensors through SHM systems along with wearable IMU sensors for intelligent crowd management. The utilization of multi modal sensors with machine learning along with multi modal sensor data fusion sheds light on a new scope and application, enhancing crowd and structure operation efficiency with warning capability from any safety risk.

Concept validation is performed on a model test bridge with group volunteers simulating crowd flow thus generating strain and accelerometer datasets. The presented approach allows for the estimation of crowd load and mobility parameters, hence providing a corner stone towards a comprehensive crowd management system with artificial intelligence-based decision making. In addition, results show how a conventionally used machine learning approach, SVM, and an innovative approach, CNN, tackles such application. Models with the aid from fusion techniques are generated to classify crowd flow activity and density load, in addition to crowd load estimation with greater robustness. The results showed that the monitoring solution to be highly effective with peak testing accuracy for single class flow

activity classification at 98%, multi-class flow and load characterization classification at 91%, and percentage error for load estimation regression reaching a minimum of 9%.

In summary, both SVM and CNN approach performed generally high for both classification and regression for our application of crowd monitoring. Classification of flow activity showed that CNN for both individual and fusion inputs, albeit slightly, outperformed SVM under both signals. However for multi-class classification, though reasonably still high, SVM within a small margin returned higher accuracies under certain testing cases. In terms of regression, with focus for Unmodified testing, returned percentage errors within same inputs are very similar with little margin between the two machine learning approaches.

#### Future Works

With a larger a dataset and further machine learning fine-tuning, better results with greater efficiency and robustness are possible. A study between the number of sensors in relation to cost and added efficiency could be used for future sensor choice and distribution. An added benefit of daisy chained FBG sensors in different configurations may be used for localization and better understanding of load distribution and flow, returning a heatmap of the bridge itself. Fusion techniques employed at the input and feature level could be further improved in terms of how the data is handled. More methods and approaches of fusion at the input and feature level is to be investigated. It is important to explore the aspect of variation of dimensionality between the two multi-modal sensors when applying fusion techniques at the employed levels. It is of interest to investigate other discriminant fusion approaches that may work appropriately for SVM. A further addition to the employment of fusion is at the decision



level. Output classification results from the two individual inputs are in return jointly voted and nominated to a final decisive and concurring output. This could be used in terms for flow activity and load characterization separately from different sensor modality. A main criterion in affecting CNNs' results are its architecture. Fine-tuning a CNN's architecture in terms of number and choice of layers with their corresponding parameters is very important to achieve a better and efficient result. For example, it is of interest to perceive how the convolution filter and their according sizes followed by the different types of pooling layer may handle the image from input level fusion data as compared to their individual counterparts. This concern rises from the fact that individually, Accelerometer and FBG image have different convolution filter sizes and a maximum pooling layer may be sufficient. However in the case of input level fusion, the image data is of two different sensors of different sizes were the possibility exists that either one of the sensor data range may not be fully comprehended. A recommended future work is employing an advanced hybrid machine learning model consisting of CNN's automatic feature extractor along with SVM's classifier. The hybrid model will rely partly on CNN convolution and corresponding layers to obtain discriminative features and weights. Rather than on continuing within the CNN's classifier, these are used as data employed in SVM classification. An important aspect in the joint structure and crowd monitoring is the ability to distinguish between overloading and damage. A path in data gathering and machine learning modelling from the test bridge with a damage with crowd replication is of interest. Supplementary testing of instrumentation at field on an actual pedestrian bridge could prove valuable insights on how an actual structure might behave and how the machine leaning models

might react with the gathered data. To better enhance the status overview of an individual in a crowd, the addition of other sensors of different modality within the wearable sensor could prove beneficiary. Monitoring a biological parameter such as blood pressure or heart rate could supplement current crowd status based on the individual's status and biological reactions. A Global Position System (GPS) receive could be embedded within for added tracking, in terms of speed and direction perceiving a better image of flow. Current accelerometer readings could be further trained to detect trips and falls of the individual within the crowd, warning monitors of potential stampede cases

Given the information available about the crowd load and behavior, it is essential to close the loop via a system that implements crowd management strategies. Based on the state of flow and the state of the structure under the crowd loading, strategies may include adapting routes and providing instructions to individuals via smartphone messages or loud speakers.

## REFERENCES

1. Mohandes, M.A., *Mobile technology for socio-religious events: a case study of NFC technology*. IEEE Technology and Society Magazine, 2015. **34**(1): p. 73-79.
2. Still G. K. *Crowd Safety and Risk Analysis - Crowd Disasters*. 2018 [cited 2017 12 December]; Available from: <http://www.gkstill.com/ExpertWitness/CrowdDisasters.html>.
3. Gayathri, H., P. Aparna, and A. Verma, *A review of studies on understanding crowd dynamics in the context of crowd safety in mass religious gatherings*. International Journal of Disaster Risk Reduction, 2017.
4. Helbing, D., A. Johansson, and H.Z. Al-Abideen, *Dynamics of crowd disasters: An empirical study*. Physical review E, 2007. **75**(4): p. 046109.
5. BBC News. *Many die in India stampede on way to Hindu religious event*. 2016 [cited 2017 15 October]; Available from: <http://www.bbc.com/news/world-asia-india-37667194>
6. BBC News. *India stampede 'kills 27 pilgrims' in Andhra Pradesh*. 2015 [cited 2017 14 July]; Available from: <http://www.bbc.com/news/world-asia-india-33518240>.
7. Cespedes, X. and S. Manon, *Dynamic analysis of the Bercy-Tolbiac footbride*. Proceedings of Footbridge, 2005.
8. Fournol, A., et al., *Dynamic behaviour of Ceramique footbridge (Maastricht, NL)*. Proceedings of Footbridge, 2005.
9. Moutinho, C., et al. *Dynamic behaviour of a long span stainless steel arch footbridge*. in *Proceedings of the International Conference on the Design and Dynamic Behaviour of Footbridges*. 2002.
10. Powell, D., A. De Donno, and A. Low, *Design of damping systems for footbridges. Experience from Gatwick and Ijburg*. Proceedings of Footbridge, 2005.
11. Dallard, P., et al., *The London millennium footbridge*. Structural Engineer, 2001. **79**(22): p. 17-21.
12. Dallard, P., et al., *London Millennium Bridge: pedestrian-induced lateral vibration*. Journal of Bridge Engineering, 2001. **6**(6): p. 412-417.
13. Hurriyet Daily News & Economic Review. *Swaying causes running wariness over Bosphorus Bridge*. 2010 [cited 2011 8 Aug.]; Available from: <http://www.hurriyetdailynews.com/n.php?n=swinging-bosporus-bridge-alerts-experts-2010-10-18>.
14. Wolmuth, B. and J. Surtees. *Crowd-related failure of bridges*. in *Proceedings of the Institution of Civil Engineers-Civil Engineering*. 2003. Thomas Telford Ltd.
15. Fujino, Y., et al., *Synchronization of human walking observed during lateral vibration of a congested pedestrian bridge*. Earthquake engineering & structural dynamics, 1993. **22**(9): p. 741-758.

16. Nakamura, S.-i. and Y. Fujino, *Lateral vibration on a pedestrian cable-stayed bridge*. Structural Engineering International, 2002. **12**(4): p. 295-300.
17. BBC News. *At least six dead in Florida university bridge collapse*. 2018 [cited 2018 16 March]; Available from: <https://www.bbc.com/news/world-us-canada-43418898>.
18. DW News. *Prague bridge collapse leaves 4 injured*. 2017 [cited 2017 2 December]; Available from: <https://p.dw.com/p/2ofeJ>.
19. Chertkoff, J.M. and R.H. Kushigian, *Don't panic: The psychology of emergency egress and ingress*. 1999: Praeger Westport, CT.
20. Helbing, D. and P. Mukerji, *Crowd disasters as systemic failures: analysis of the Love Parade disaster*. EPJ Data Science, 2012. **1**(1): p. 7.
21. Filingeri, V., et al., *Factors influencing experience in crowds—the participant perspective*. Applied ergonomics, 2017. **59**: p. 431-441.
22. Sharma, D., et al., *A review on technological advancements in crowd management*. Journal of Ambient Intelligence and Humanized Computing, 2016: p. 1-11.
23. Helbing, D. and P. Molnar, *Social force model for pedestrian dynamics*. Physical review E, 1995. **51**(5): p. 4282.
24. Zhan, B., et al., *Crowd analysis: a survey*. Machine Vision and Applications, 2008. **19**(5-6): p. 345-357.
25. Cassol, V.J., et al., *Simulating Crowds in Egress Scenarios*. 2017, Springer.
26. Sindagi, V.A. and V.M. Patel, *A survey of recent advances in cnn-based single image crowd counting and density estimation*. Pattern Recognition Letters, 2018. **107**: p. 3-16.
27. Wu, X., et al. *Crowd density estimation using texture analysis and learning*. in *Robotics and Biomimetics, 2006. ROBIO'06. IEEE International Conference on*. 2006. IEEE.
28. Zhang, Y., et al. *Single-image crowd counting via multi-column convolutional neural network*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
29. Sindagi, V.A. and V.M. Patel. *Cnn-based cascaded multi-task learning of high-level prior and density estimation for crowd counting*. in *Advanced Video and Signal Based Surveillance (AVSS), 2017 14th IEEE International Conference on*. 2017. IEEE.
30. Fu, M., et al., *Fast crowd density estimation with convolutional neural networks*. Engineering Applications of Artificial Intelligence, 2015. **43**: p. 81-88.
31. Shang, C., H. Ai, and B. Bai. *End-to-end crowd counting via joint learning local and global count*. in *Image Processing (ICIP), 2016 IEEE International Conference on*. 2016. IEEE.
32. Lau, B.P.L., et al., *Sensor fusion for public space utilization monitoring in a smart city*. IEEE Internet of Things Journal, 2018. **5**(2): p. 473-481.
33. Raykov, Y.P., et al. *Predicting room occupancy with a single passive infrared (PIR) sensor through behavior extraction*. in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2016. ACM.

34. Tyndall, A., R. Cardell-Oliver, and A. Keating, *Occupancy estimation using a low-pixel count thermal imager*. IEEE Sensors Journal, 2016. **16**(10): p. 3784-3791.
35. Yaseen, S., et al., *Real-time crowd density mapping using a novel sensory fusion model of infrared and visual systems*. Safety science, 2013. **57**: p. 313-325.
36. Khozium, M.O., A.G. Abuarafah, and E. AbdRabou, *A proposed computer-based system architecture for crowd management of pilgrims using thermography*. Life Science Journal, 2012. **9**(2): p. 377-383.
37. Khaleghi, A.M., et al., *A DDDAMS-based planning and control framework for surveillance and crowd control via UAVs and UGVs*. Expert Systems with Applications, 2013. **40**(18): p. 7168-7183.
38. Ustev, Y.E., O. Durmaz Incel, and C. Ersoy. *User, device and orientation independent human activity recognition on mobile phones: Challenges and a proposal*. in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. 2013. ACM.
39. Kwapisz, J.R., G.M. Weiss, and S.A. Moore, *Activity recognition using cell phone accelerometers*. ACM SigKDD Explorations Newsletter, 2011. **12**(2): p. 74-82.
40. Wirz, M., et al., *Using mobile technology and a participatory sensing approach for crowd monitoring and management during large-scale mass gatherings*, in *Co-evolution of Intelligent Socio-technical Systems*. 2013, Springer. p. 61-77.
41. Gao, L., A.K. Bourke, and J. Nelson. *Activity recognition using dynamic multiple sensor fusion in body sensor networks*. in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2012. IEEE.
42. Ward, J.A., et al., *Activity recognition of assembly tasks using body-worn microphones and accelerometers*. IEEE transactions on pattern analysis and machine intelligence, 2006. **28**(10): p. 1553-1567.
43. Khan, A.M., et al., *Activity recognition on smartphones via sensor-fusion and kda-based svms*. International Journal of Distributed Sensor Networks, 2014. **10**(5): p. 503291.
44. Ronao, C.A. and S.-B. Cho, *Human activity recognition with smartphone sensors using deep learning neural networks*. Expert Systems with Applications, 2016. **59**: p. 235-244.
45. Jiao, L., et al., *Multi-sensor golf swing classification using deep CNN*. Procedia Computer Science, 2018. **129**: p. 59-65.
46. Noel, A.B., et al., *Structural health monitoring using wireless sensor networks: A comprehensive survey*. IEEE Communications Surveys & Tutorials, 2017. **19**(3): p. 1403-1423.
47. Ko, J. and Y. Ni, *Technology developments in structural health monitoring of large-scale bridges*. Engineering structures, 2005. **27**(12): p. 1715-1725.
48. Kahandawa, G.C., et al., *Use of FBG sensors for SHM in aerospace structures*. Photonic Sensors, 2012. **2**(3): p. 203-214.
49. Sierra-Pérez, J., A. Güemes, and L.E. Mujica, *Damage detection by using FBGs and strain field pattern recognition techniques*. Smart materials and structures, 2012. **22**(2): p. 025011.

50. Sohn, H., et al., *Structural health monitoring using statistical pattern recognition techniques*. Journal of dynamic systems, measurement, and control, 2001. **123**(4): p. 706-711.
51. Hasni, H., et al., *Detection of fatigue cracking in steel bridge girders: a support vector machine approach*. Archives of Civil and Mechanical Engineering, 2017. **17**(3): p. 609-622.
52. Tang, H., G. Tang, and L. Meng. *Prediction of the bridge monitoring data based on support vector machine*. in *Natural Computation (ICNC), 2015 11th International Conference on*. 2015. IEEE.
53. Loutas, T.H., et al., *Intelligent health monitoring of aerospace composite structures based on dynamic strain measurements*. Expert Systems with Applications, 2012. **39**(9): p. 8412-8422.
54. Panopoulou, A., et al., *Health monitoring of aerospace structures using fibre Bragg gratings combined with advanced signal processing and pattern recognition techniques*. Strain, 2012. **48**(3): p. 267-277.
55. Abdeljaber, O., et al., *Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks*. Journal of Sound and Vibration, 2017. **388**: p. 154-170.
56. Martella, C., et al., *On current crowd management practices and the need for increased situation awareness, prediction, and intervention*. Safety science, 2017. **91**: p. 381-393.
57. Jain, A.K., R.P.W. Duin, and J. Mao, *Statistical pattern recognition: A review*. IEEE Transactions on pattern analysis and machine intelligence, 2000. **22**(1): p. 4-37.
58. Devroye, L., L.a.a. Györfi, and G. Lugosi, *A probabilistic theory of pattern recognition*. 1996: Springer.
59. Duda, R.O., P.E. Hart, and D.G. Stork, *Pattern classification*. 2012: John Wiley & Sons.
60. Vapnik, V.N. and V. Vapnik, *Statistical learning theory*. Vol. 1. 1998: Wiley New York.
61. Cortes, C. and V. Vapnik, *Support-vector networks*. Machine learning, 1995. **20**(3): p. 273-297.
62. Burges, C.J., *A tutorial on support vector machines for pattern recognition*. Data mining and knowledge discovery, 1998. **2**(2): p. 121-167.
63. Byun, H. and S.-W. Lee, *A survey on pattern recognition applications of support vector machines*. International Journal of Pattern Recognition and Artificial Intelligence, 2003. **17**(03): p. 459-486.
64. Basak, D., S. Pal, and D.C. Patranabis, *Support vector regression*. Neural Information Processing-Letters and Reviews, 2007. **11**(10): p. 203-224.
65. Smola, A.J. and B. Schölkopf, *A tutorial on support vector regression*. Statistics and computing, 2004. **14**(3): p. 199-222.
66. Chang, C.-C. and C.-J. Lin, *LIBSVM: a library for support vector machines*. ACM transactions on intelligent systems and technology (TIST), 2011. **2**(3): p. 27.
67. Drucker, H., et al. *Support vector regression machines*. in *Advances in neural information processing systems*. 1997.

68. Gunn, S.R., *Support vector machines for classification and regression*. ISIS technical report, 1998. **14**: p. 85-86.
69. Rasmussen, C.E. and C.K. Williams, *Gaussian processes in machine learning*. Lecture notes in computer science, 2004. **3176**: p. 63-71.
70. Rasmussen, C.E. and C.K. Williams, *Gaussian processes for machine learning*. Vol. 1. 2006: MIT press Cambridge.
71. Chan, A.B. and N. Vasconcelos, *Counting people with low-level features and Bayesian regression*. IEEE Transactions on Image Processing, 2012. **21**(4): p. 2160-2177.
72. Gu, J., et al., *Recent advances in convolutional neural networks*. arXiv preprint arXiv:1512.07108, 2015.
73. López-Higuera, J.M., et al., *Fiber optic sensors in structural health monitoring*. Journal of lightwave technology, 2011. **29**(4): p. 587-608.
74. Gholamzadeh, B. and H. Nabovati, *Fiber optic sensors*. World Academy of Science, Engineering and Technology, 2008. **42**(3): p. 335-340.
75. Glisic, B. and D. Inaudi, *Fibre optic methods for structural health monitoring*. 2008: John Wiley & Sons.
76. Majumder, M., et al., *Fibre Bragg gratings in structural health monitoring—Present status and applications*. Sensors and Actuators A: Physical, 2008. **147**(1): p. 150-164.
77. Haghghat, M., M. Abdel-Mottaleb, and W. Alhalabi, *Discriminant correlation analysis: Real-time feature level fusion for multimodal biometric recognition*. IEEE Transactions on Information Forensics and Security, 2016. **11**(9): p. 1984-1996.

