ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# MSFMT: MULTI-SPECTRAL FUSION SYSTEM BASED ON DEEP TRANSFER LEARNING AND DEMPSTER-SHAFER THEORY

**[1]DOAA MOHEY EL-DIN, [2] ABOUL ELLA HASSANIEN, [3] EHAB E. HASSANEIN**

[1]PhD Researcher, Information Systems Department, of Faculty of Computers and artificial intelligence, Cairo University, Egypt

[2]Professor, Information Technology Department of Faculty of Computers and artificial intelligence, Cairo University, Egypt

[3]Professor, Information Systems Department of Faculty of Computers and artificial intelligence, Cairo University, Egypt.

E-mail: [1]d.mohey@alumni.fci-cu.edu.eg, [2] aboitcairo@gmail.com, [3]e.ezat@Fci-cu.edu.eg

**ABSTRACT**

This paper presents a multi-spectral fusion system for improving the object detection and classification results in military and terrorism domains that entitled MSFMT. It relies on a combination between deep transfer learning and Dempster-Shafer statistical method in decision level fusion. It improves the classification results through fusing multiple sensory data that are extracted from multiple sources into two data types, images and videos, in night modes. It fuses multiple spectrums for showing the best vision for each object or action. These spectrums are Visual Intensified images (VIS), Near-infrared spectroscopy (NIR) images, thermal images, long wave infrared images (LIWR), DHV, and RGB). The neural network structure is constructed based on six neural networks. Each neural network is based on AlexNet pre-trained transfer neural networks for classifying spectrums. Each neural includes two neural networks for classifying objects and actions. MSFMT system improves the classification results through creating a new algorithm for multi-spectral fusion that depends on integration between machine learning (transfer learning model) and statistics methods (Dempster-Shafer evidential reasoning technique). MSFMT uses a purified dataset that consists of tuned six datasets for multiple spectrums in variant type format. It applies the data augmentation for enlarging the dataset that includes 875,970 number of images and video's frames. The fusion accuracy results reach 96 % that increases the classification results by 21% for NIR and VIS spectrums, 7% for Thermal spectrum, and 4 to 6% for LIWR, DHV, RGB, spectrums.

**Keywords:** Multi-Spectral, Sensor Fusion, Decision Level Fusion, Military, Terrorism, Transfer learning.

## 1. INTRODUCTION

Military is responsible for saving civilizations and making the guarantee of the safest countries. So, that is very important to determine any danger or threat by making fast a suitable decision in any terrorism's attack [1]. There is a big obstacle in automatic determining terrorism's objects and classifying their attack's actions [2]. Killing and intimidating the security, the attack on the army, or weapon robbery are the most frequent actions that are happened. Some weapons try to see strange objects or actions in various spectrums such as infrared missiles and thermal weapons. But there is still a challenge to see all objects or actions of terrorists due to their stealth tricks especially in multiple spectrums in night mode. So, there is a deep need to make an automated way for determining terrorism objects and actions in multiple spectrums in night mode [3, 4]. That can remove the images blurred simultaneously. For example, consider the problem of identifying an army tank or sniper in a surveillance environment. The army tank or sniper can be identified as either a friend hostile or neutral. Prior information about the environment may be available, such as the distribution of military missions that can help to determine if an army tank or people are friendly, hostile or neutral.

This paper proposes an automated multi-spectral fusion system, that is known as MSFMT, for improving object detection and classification results. The fusion process aims to reach the full vision of several spectrums and remove blurred images by getting some additional information that is hard to identify by the human eye with its receptors for red, green and blue. It increases reliability and reduces the uncertainty. It also minimizes the noise of the images, covers the wide spatial range and temporal resolution coverage. Multi-spectral imaging refers to the electromagnetic spectrum with various wavelengths [5, 6]. These spectrums are extracted from sensory data in variant types as videos and images. There are four abstractions levels of multi-sensor data fusion: signal, pixel, features, and decision levels. They differ in measurements, characteristics, and decisions. Decision level fusion is considered a high-level fusion that is interested in the fusion information to improve the decisions concurrently.

The proposed fusion system uses a combination of machine learning and statistical method in decision level fusion to improve the accuracy results. It uses deep transfer learning with Dempster-Shafer technique which uses a probability of a symbol representing a decision. It consists of three layers that are automated detection based on Alexnet pre-trained neural network, description layer through extracting several characteristics from the input data and creating the lexicon to improve the domain's classification, and the third layer is Dempster-Shafer technique for decision level fusion between a new purified dataset that is collected and tuned from six benchmark datasets.

The rest of this paper is constructed as the following. Section II examines the background and related works about object detection in multi-spectral in variant datatypes and statistics methods of data fusion. Section III presents the proposed solution for fusing multi-spectral images of terrorism object recognition. Section IV introduces the experiments that test the performance of the proposed method. At the end, the conclusion and future works of this work in Section. V.

## 2. BACKGROUND AND RELATED WORKS

This section presents a discussion of information fusion, multi-sensor data fusion, and many concepts and techniques to make a fusion. It also presents a comparative study for the previous motivations in multi-spectral fusion in variant types. They aim to improve the accuracy and remove the blurred images. Information fusion refers to a conformity problem for multi-source information in image processing, inference, and signal processing and knowledge representation [7].

### 2.1 Multi-Sensor Data Fusion

Multi-sensor data fusion refers to the integration of sensory data extracted from disparate sources [8]. Multi-spectral image refers to a type of sensor fusion in multiple bands of spectrums to reach the full vision of several spectrums and remove blurred images [9]. There are four abstractions levels of multi-sensor data fusion: signal, pixel, features, and decision levels [10, 11, 12, 13]. They differ in measurements, characteristics, and decisions. Decision level fusion is considered a high-level fusion that is interested in the fusion information to improve the decisions concurrently [14, 15]. In the following, a discussion for a comparison study between previous fusion researches in various spectrums, as shown in Table.1, and a comparison between their benefits and open challenges and their limitations as shown in Table.2.

*Table.1: A comparative study between previous researches in fusion*

| Paper No. | Target | Spectrums | Fusion level | Technique |
|---|---|---|---|---|
| [14] | Fusion images from Video surveillance in multi-spectral | VIS, NIR, mid-wave IR, LIWR | Pixel level | PCA (Principle component Analysis) |
| [15] | Creating framework over the varying aspects surrounding its implementation | VIS, Thermal in night vision | Decision level | Laplacian pyramid fusion |
| [16] | Fusing four grayscale images from different bands | Mid-infrared, NIR, Thermal-infrared | Pixel level | PCA (Principle component Analysis) And VTVA |
| [17] | framework via novel scale map construction | RGB, NIR | Feature level (Noisy-level) | Multi-spectral shadow detection |
| [18] | Classify objects in multi-spectral in urban area | long-wave infrared and visible images | Decision level | Dempster–Shafer classifier fusion |
| [19] | Detection of surface water in natural environment | NIR channels | Decision level | Dempster–Shafer classifier fusion |

Journal of Theoretical and Applied Information Technology
31st March 2020. Vol.98. No 06

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Researchers in [14], present an image fusion that is extracted from videos in surveillance for four spectrums VIS, NIR, mid-wave IR, and LIWR. They use PCA (Principle component Analysis) technique to improve accuracy by 95%-98%, however they find an obstacle in small scale data and noisy data features. Researchers in [15], introduce a framework for multi-spectral fusion based on two spectrums VIS and Thermal in night mode. They use the Laplacian pyramid fusion technqiue to increase the accuracy results 91% based on feature fusion level, however, they still face a problem in neglecting some features and information about each input data. Researchers in [16], present. Researchers in [16], show a fusion technqiue for four grayscale images from different bands based on Mid-infrared, NIR, thermal, and infrared spectrums. They use PCA (Principle component Analysis) And VTVA. VTVA algorithm improves results by 50% approximately. However, they still face a challenge in variant level between colors and grayscale images. Researchers in [17], show a framework for map construction based on RGB and NIR spectrums. The results of using Multi-spectral shadow detection technqiue increase the quality and accuracy by 91%, the limitation is appeared in that there is no guidance structure in the rectangle and making restoration less constrained. Researchers in [18], present an object classification in multi-spectral in urban area that includes two spectrums long-wave infrared and visible images. They increase accuracy by 93%. The accuracy is requiring improving. Researcher in [19], present an object Detection of surface water in natural environment. They use Dempster–Shafer classifier fusion to increase accuracy results by, 87.92% as water. Whatever the spectral water as well as unavailability of certain spectral band in our data, such as Shortwave Infrared (SWIR) band and ignorance several features are still challenges that are faced them. In the benefits and limitations challenges.

*Table.2: A comparative study of previous fusion researches*

| Paper No. | Benefits | Limitations |
|---|---|---|
| [14] | High accuracy | Sensor noise, real time challenges, and small-scale dataset |
| [15] | improve human perception | neglecting some features and information about each input data the accuracy and efficiency require improving |
| [16] | Clearer by VTVA | variant level between colors and grayscale images |
| [17] | Good quality results | There is no guidance structure in the rectangle of and making restoration less constrained. |
| [18] | Improve classification results More reliable results | Requires improving the accuracy results |
| [19] | very satisfying performances on detection of water bodies | the spectral water as well as unavailability of certain spectral band in our data, such as Shortwave Infrared (SWIR) band and ignorance data |

## 2.2 Multi-Sensor Data Fusion Techniques

Due to various levels of data fusion, there are many techniques that are suitable for each level. Multi-sensor data fusion covers many methods and algorithms, containing: Central limit theorem, Kalman filter, Bayesian networks, Dempster-Shafer, and Convolutional neural network.

The highest usage techniques in decision level fusion are statistical algorithms such as containing maximum a posteriori, maximum likelihood, Neyman-Pearson, minimax, and Bayes. [19]. Traditional statistics [20, 21] utilize Metrics from multiple sensors to work inferences about the objective's identity. An essential obstacle to traditional statistical measurement techniques is the absence of the use of prior information about the parameter being estimated. Traditional statistics only utilize measurements extracted from sensors, but if prior information is available about the target's identity, this information would be ignored.

Bayesian estimation techniques address this. Bayesian estimation techniques are based upon Bayes' rule,

$$p(Y_i|X) = \frac{P(X\}Y_i|X)\,P(Y_i)}{\Sigma_i P(X\}Y_i|X)\,P(Y_i)} \qquad (1)$$

Where the decision depends on both the a priori probabilities and the likelihoods as estimated by the sensors. Uncertainty in Bayesian analysis is represented by a probability which can take on values between 0 and 1 where a probability of 1 implies that an event always occurs and a probability of 0 implies that an event never occurs. The techniques of Bayesian have been implemented in the military domain to help consistency, condition assessment, and threat measurement.

Bayesian techniques [22] have also been utilized in robotics. The analysis of Bayesian has many advantages. Furthermore, it utilizes prior information, Bayesian techniques let one analyze an exhaustive group of causes and deal with single and multiple events. The disadvantages of Bayesian

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

analysis include the fact that it may be hard to identify the proper prior probabilities. If data is not available, the prior probabilities may tend to be subjective, defined by a panel of "experts". Bayesian analysis also does not allow for general uncertainty. Bayesian techniques may not be enough if the information from a sensor is incomplete or one sensor extends information at a different level than a second sensor.

The address of Dempster-Shafer methods refers to a shortage of Bayesian analysis in that it allows for general uncertainty [23, 24]. It utilizes evidential reasoning to find the identity. This permits each sensor to participate in data at its own level of detail. Uncertainty is represented by a degree of belief that may take on values between 0 and 1 where a degree of belief of 0 implies a complete lack of belief in a hypothesis whereas a degree of belief of 1 implies total belief in a hypothesis. It replaces the a priori probabilities used in Bayesian techniques with an ignorance factor. Ignorance can be reduced when supporting information becomes available. Dempster's rule of integration fits the mechanism to gather information via multiple sources. It relies on belief, also known by support function, and plausibility functions that provides bounds upper and lower of proposition A. a belief is defined by a lower bound on the probability of proposition A. the belief or support function is defined by equation (2).

$$Spt_i\,(A) = \sum A_j\; m_i\,(A_j), \qquad (2)$$

And the plausibility function is defined by an upper bound on the probability of proposition A, as shown in equation (3).

$$Pls\,(A) = 1\text{-}Spt\,(\sim A), \qquad (3)$$

A comparative study between traditional Bayesian technique and Dempster-Shafer technique is discussed in Table.3. It discusses the differences between usability target, advantages, and disadvantages.

*Table.3: a comparison between Bayesian theory, and Dempster-Shafer.*

| | Bayesian | Dempster-Shafer |
|---|---|---|
| Target | Making fusion about the target identity | Solving the Bayesian problem |
| Advantages | the decision is based upon both the a priori probabilities and the likelihoods as estimated by the sensors. | Improve uncertainty of Bayesian theory. utilize for generally uncertainty level. |
| Disadvantages | The absence of the use of prior information about the parameter being estimated.<br>IT does not allow for general uncertainty<br><br>It may not be enough if the information from a sensor is incomplete or one sensor provides information at a different level than a second sensor. | It suffers from a possibility for weak decisions and can be computationally complicated as the number of classes and sensors raise. |

From prior motivations, finding the accuracy fusion requires improving. Multi-spectral fusion is an open research challenge due to the lack of datasets and lack of studying the spectrum's nature. A multi-spectral system usually prepares a combination of visible (0.4 to 0.7 µm), near-infrared (NIR; 0.7 to 1 µm), or long-wave infrared (LWIR; 8 to 12 µm) bands into a single system. Till now, no founded system can fuse six spectrums. Military domain has low motivations to improve decision making in classification and detection. Terrorism domain is one of hardest domains to find dataset for studying the multi- spectrums. So, this paper presents a solution for the previous challenges that are examined previously. The proposed solution is a multi-spectral fusion system that increases the visibility vision of detection and classification objects and action in night mode. Multi-spectral imaging merges two to six spectral imaging bands of relatively large bandwidth into a single optical system. It extracts the terrorism dataset from the military datasets. It aims to improve the accuracy results.

## 3. THE PROPOSED MULTI-SPECTRAL FUSION SYSTEM FOR IMPROVING MILITARY DETECTION AND TERRORISM CLASSIFICATION

The proposed system is entitled MSFMT, is a multi-spectral fusion system that uses for improving military detection and terrorism classification result. It fuses six spectrums which are extracted from multiple sensors in night vision. It improves the vision of detection objects and actions. It also extracts the terrorism objects and actions form the military dataset automatically. It proposes a new algorithm that is based on a combination of deep transfer learning and Dempster-Shafer technique for enhancing decision making and the classification's accuracy results. It

31st March 2020. Vol.98. No 06

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

fuses data from multiple data formats in decision level fusion.

The classification depends on supervised classification that uses for checking on similar signatures with respect to training samples. MSFMT aims to improve the classification objects and actions in military and terrorism domains with respect to two data type formats in multiple

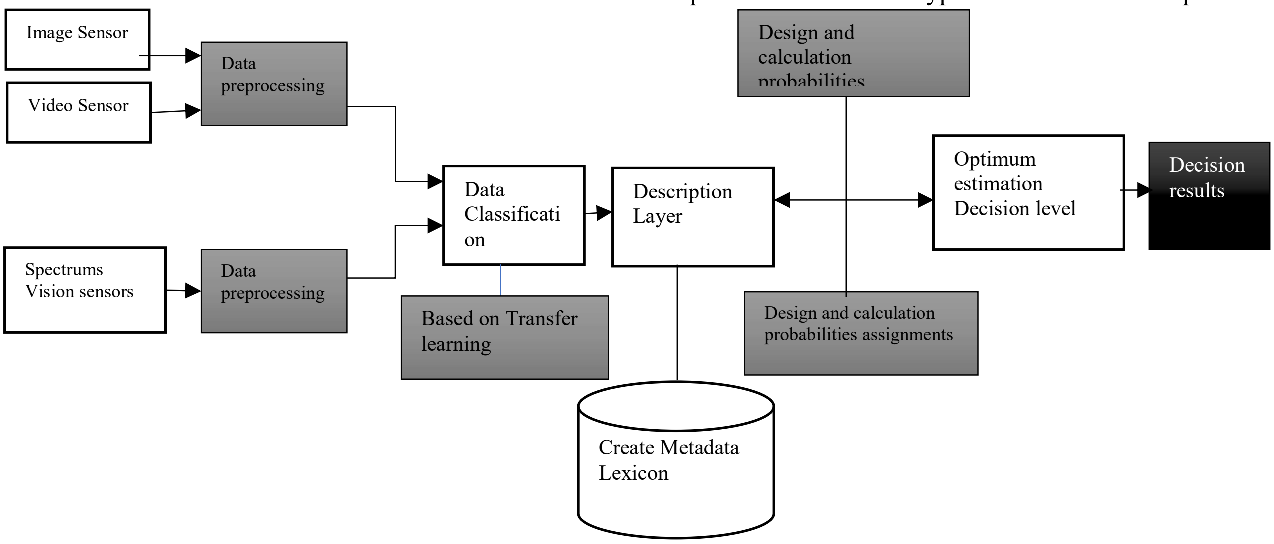


*Figure.1: the proposed structure of MSFMT: Multi-Spectral Fusion System based on a Combination of Deep Transfer Learning and Dempster-Shafer Technique for enhancing Military Detection and Terrorism Classification*

## 3.1 STRUCTURE

The structure of MSFMT consists of three layers (1) the first layer depends on twelve neural networks that refer to six main AlexNet pre-trained neural networks for classifying objects and actions on six spectrums as shown in the table.4. Each neural includes two neural networks for determining the domain classification whether military or terrorism. The power of using transfer neural network illustrates in improving results and cures the problem of lack of datasets. It supports making decisions fast.

*Table.4: the main structure of the neural network*

| Neural | Spectrums | Objects | | | Actions | | |
|---|---|---|---|---|---|---|---|
| N1 | Spectrum 1 | A | B | c | X | Y | Z |
| N2 | Spectrum 2 | A | B | C | X | Y | Z |
| N3 | Spectrum 3 | A | B | C | X | Y | Z |
| N4 | Spectrum 4 | A | B | C | X | Y | Z |
| N5 | Spectrum 5 | A | B | C | X | Y | Z |
| N6 | Spectrum 6 | A | B | C | X | Y | Z |
| | | **Classify objects** | | | **Classify actions** | | |
| **Classify domain** | | **Whether terrorism or military** | | | | | |

Spectrums. It can simulate the real image and avoid camouflage in vision. The output of classification illustrates in the following (Figure.2, 3).

The second layer generates a description layer for interpreting the features extraction and mentioning a description of the image holds to strengthen the ability of classification and fusion processes. This layer creates a lexicon for metadata.

The third layer applies a Dempster-Shafer theory of evidence for fusing multiple spectrums in night vision. This technique is also known as the "evidence theory" or the "belief function theory", is a formal framework for reasoning with partial, uncertain and imprecise information.

The basic concepts of evidence theory are belief structure, belief and plausibility functions. The experiment calculates accuracy measurement. MSFMT improves the accuracy results that reach to 96%. The experiment is applied to a purified dataset.

*Table.5: A Fusion Technique for Multi Spectral Fusion in Terrorism domain*

| A proposed Algorithm pseudo code |
|---|
| 1. Input (i) from (M, V) images and videos<br>2. Preprocessing input (get frames of video) |

3. Let ($v_n$) consists of some sequenced frames (f), so let ($v_n$) $=\{f_1, f_2, f_3, f_4, \dots, f_n\}$

4. Consequence frames for each action with key and time.

5. For each ($v_n$), each frame has key and time parameters (f), so let ($v_{nt}$) $=\{f_{t1}, f_{t2}, f_{t3}, f_{t4}, \dots, f_{nt}\}$, t: refers to time, key can be expressed by number.

6. Resize the images and frames, all images 227*227

7. Let ($m_s$), the same images with various spectrums (s), so let ($m_s$) $=\{m_{s1}, m_{s2}, \dots, m_{s6}\}$

Using pre-trained model approach for transfer learning

8. pre-trained model will act as a feature extractor
9. freeze the weights of all the other layers
10. Remove the last layer of the network and replace it with our classifier.
11. Train the network normally.
12. Classify objects and actions
13. Classify spectrums

Creating metadata layer

14. Create a description layer of extracted metadata of images and videos
15. Create Automatic Metadata Lexicon with weak supervision
16. Automatic similarity issues to support creating a new lexicon.
17. The output result of transfer learning is as an input for fusion.

Applying Fusion process

18. Let Θ represent the "frame of discernment". This is the set of all mutually and exhaustive propositions. Let 2 Θ represent the power set of Θ. For each proposition in $2^{\Theta}$, a probability mass m is assigned subject to the conditions that

$$m(\phi) = 0$$

And

$$\Sigma_{A_i \in 2^{\Theta}} m(A_i) = 1$$

19. The belief or support function provides a lower bound on the probability of proposition A and is defined by

$$Spt_i(A) = \Sigma_{A_j \subset A} m_i(A_j)$$

20. The plausibility function provides an upper bound on the probability of proposition A and is defined by

$$Pls(A) = 1 - Spt(\sim A)$$

21. The uncertainty interval of proposition A is [spt(A), pls(A)] and the uncertainty of proposition A is given by

$$u(A) = Pls(A) - Spt(A)$$

22. For each possible proposition (e.g., user-A), Dempster-Shafer theory gives a rule of combining sensor Si's observation mi and sensor Sj's observation mj

$$(m_i \oplus m_j)(A) = \frac{\Sigma_{(E_k \cap E_{k'}) = A}\, m_i(E_k) m_j(E_{K'})}{\Sigma_{(E_k \cap E_{k'}) = \emptyset}\, m_i(E_k) m_j(E_{K'})}$$

This combining rule can be generalized by iteration: if we treat mj not as sensor Sj's observation, but rather as the already combined (using Dempster-Shafer combining rule) observation of sensor Sk and sensor S

23. Output (fusion process and Improve accuracy results)

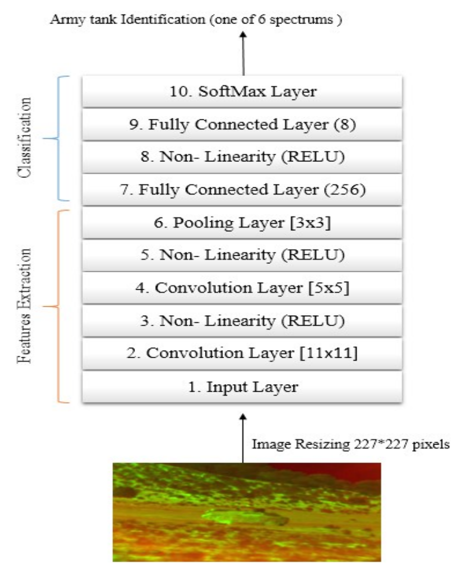


*Figure.2 An architecture of AlexNet pre-trained neural network for military detection and terrorism classification that depends on CNN*

The main obstacle of decision-level sensor fusion is that very little sensor information is utilized in the fusion process. The detection performance of a group of sensors can be increased by adding more information to the fusion process. MSFMT is based on one of the important techniques of Statistical inference that is the process of using data analysis to conclude characteristics of an underlying probability distribution. Inferential statistical analysis infers properties of a population. The used technique Dempster-Shafer theory is designed to overcome with changing levels of precision regarding the information and no further assumptions are needed to exemplify the information. It also enables the direct representation of uncertainty of system

Journal of Theoretical and Applied Information Technology
31st March 2020. Vol.98. No 06


ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

responses where a fuzzy input can be characterized by a set or an interval and the resulting output is a set or an interval.

The proposed algorithm is applied in experiment in the following section.

## 4. EXPERIMENT AND RESULTS

Many experiments are applied on the purified dataset that apply the proposed algorithm for improving the classification results.

### 4.1 The Dataset Description

The classification of MSFMT uses AlexNet which refers to a pre-trained neural network based on a self-built dataset having 3200 images about military objects collected from the Internet. AlexNet includes many pieces of trainings on previous images that reach to more than 1,000,000 images that divide into 1000 categories. There are several images that present military objects. Terrorism objects include some similar weapons, actions, and objects from the military domain. So, the AlexNet is an important to use in this domain. However, there is a big challenge due to the lack of images dataset in this domain that causes of problem in learning the specific features.

MSFMT creates a new purified dataset that is extracted and tuned from six datasets for six spectrums in two variant formats [33, 34, 35, 36, 37, and 38]. For examples actions as fire, and objects as army tanks, people, and weapons with respect to six types of spectral images.

The proposed dataset includes 145,660 images and video's frames in six multi-spectral images. It uses data augmentation for enlarge the data to reach 875,970 that is more powerful in training by avoiding the data overfitting. It is considered a purified scheme to reach the suitable dataset for the military and terrorism domains. This paper generates a new dataset that is designed from five datasets based on purified scheme due to the lack of critical data (images or videos) in these domains. The proposed dataset includes 875,970 images in six multi-spectral images [38,39,40, 41, and 42]. It uses data augmentation for enlarge the data to reach 875,970 images and image frames videos that is more powerful in training by avoiding the data overfitting.

*Table.6: The purified dataset is collected from six mentioned datasets i the following*

| Dataset | Ref | Description | Dataset Size |
|---|---|---|---|
| TNO Image Fusion | [38] | visual (0.4–0.7μm), near-infrared (NIR, 0.7–1.0μm) and long-wave infrared (LWIR, 8–14μm) | 579 images |
| Gun objects dataset | [39] | Images dataset in real images | 333 images |
| multi-spectral images | [40] | Dataset of seven objects into three spectrums: visible, near-infrared and thermal spectrum. | 420 images |
| Flir-starter thermal | [41] | Images Thermal Datasets based on LiDAR sensor | 119,491 images |
| Terravic We apon IR | [42] | Video dataset for Terravic Weapon Infrared (video sequences data) | 24837 images |

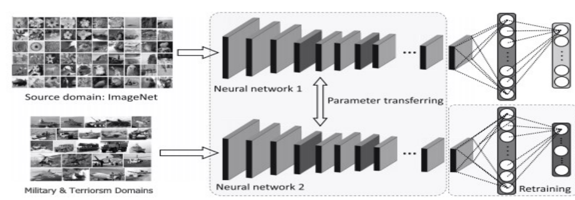


*Fig.3(a): Framework of the deep transferring network for military object recognition*

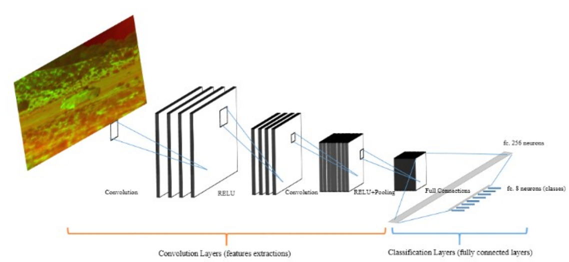

*Figure.3(b): an example of the deep learning model classification*

The new dataset makes a purified scheme to reach the suitable dataset for the military and terrorism domains. It consists of 875,970 images and videos for six spectrums.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

*Table.7: The purified dataset is based on purified scheme with respect to data processing in the five datasets.*

| Dataset | Ref | Tuning and Processing |
|---|---|---|
| TNO Image Fusion | [38] | Interpreting the video image frame, and detect military objects from these images |
| Gun objects dataset | [39] | Not all images are clear, so, the processing deletes 50 images. The tuning process applies on detecting the peoples attack with gun in some dimensions by focusing on 32 various position. That uses to support the action attacks detection from people and some gun types for objects detection. |
| Multi-spectral images | [40] | Delete 139 images for nine objects that are not included in this domain. |
| Flir-starter thermal | [41] | Delete 5902 images that includes several objects out of this domain such as dogs or bicycles |
| Terravic Weapon IR | [42] | The tuning processing applies some explanation in a generated lexicon in a description layer about images |

There are several examples from the dataset as shown in figure.4, 5.

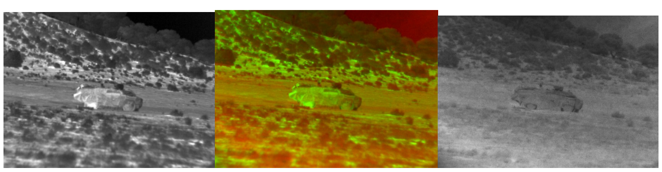

*Fig.4 (A): Examples of Army Tank in multi-spectral object detection*

*Fig.4 (B): Examples of helicopter in multi-spectral object detection*

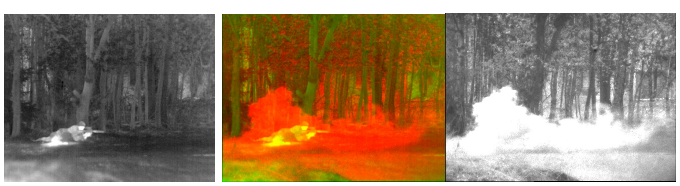

*Fig.4 (C): Examples of helicopter in multi-spectral object detection and fusion*

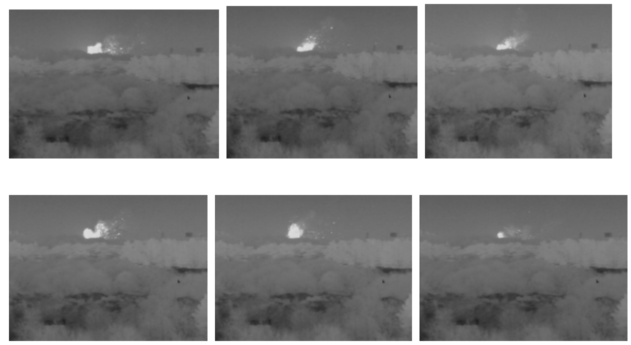

*Fig.4 (d) the terrorist Fire sequenced video object recognition test in Thermal spectrum*

*Fig.4: The network is trained with transfer learning, 98% of the samples are used as a training set*

## 4.2 A FUSION TRAINING TRACING EXAMPLE

A fusion technique is tracing in the following example sample as shown in table 5, 6. The input of fusion technique is extracted from the output of AlexNet pre-trained neural network classification. The classification output has six vectors from confusion metrics. Table.5 illustrates the relationship between 10 objects and finding in six spectrums according to the results of confusion matrices.

*Table.8. the output of the six neural networks is considered an input for the Dempster-Shafer fusion technique. It consists of six vectors from six AlexNet pre-trained neural networks S1, S2, S3, S4, S5, S6.*

| | a | B | c | D | e | f | G | H | K | L | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| S1 | 16% | 15% | 14% | 23% | 3% | 10% | 7% | 5% | 1% | 6% | 100% |
| S2 | 28% | 31% | 3% | 1% | 15% | 0 | 2% | 0 | 1% | 1% | 100% |
| S3 | 1% | 11% | 2% | 59% | 7% | 2% | 1% | 6% | 9% | 2.1% | 100% |
| S4 | 18% | 31% | 12% | 9% | 0 | 1% | 14% | 6% | 2% | 7% | 100% |
| S5 | 0 | 4% | 17% | 18% | 23% | 28% | 8% | 3% | 5% | 2% | 100% |
| S6 | 4% | 3% | 0 | 0 | 7% | 3% | 39% | 12% | 19% | 13% | 100% |

1. Let Θ represent the "frame of discernment". This is the set of all mutually and exhaustive propositions. Let 2 Θ represent the power set of Θ. For each proposition in $2^\Theta$, a probability mass m is assigned subject to the conditions that

$m(\phi) = 0$

And

$$\sum_{A_i \in 2^\Theta} m(A_j) = 1 \quad (1)$$

Journal of Theoretical and Applied Information Technology
31st March 2020. Vol.98. No 06

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. The belief or support function provides a lower bound on the probability of proposition A and is defined by

$$Spt_i(A) = \sum_{A_j \subseteq A} m_i(A_j) \quad (2)$$

3. The plausibility function provides an upper bound on the probability of proposition A and is defined by

$$Pls(A) = 1 - Spt(\sim A) \quad (3)$$

Applying on the table ,

$Pls(A) = 1 - Spt(\sim 0.16)$=0.84,

$Pls(A) = 1 - Spt(\sim 0.28) = 0.72$

$Pls(A) = 1 - Spt(\sim 0.01)$=0.99

$Pls(A) = 1 - Spt(\sim 0.18) = 0.82$

$Pls(A) = 1 - Spt(\sim 0) = 0.90$

$Pls(A) = 1 - Spt(\sim 0.04) = 0.96$

Then, $Spt_i(A) = \sum_{A_j \subseteq A} m_i(A_j)$

(4)

M(s1-a)=
$\sum_{s1 \subseteq A} acc_i(A_j) = acc_{k1} * (1 - 0.16) = 0.75 * .84 = 0.63$

M(s1-a)=
$\sum_{s2 \subseteq A} acc_i(A_j) = acc_{k2} * (1 - 0.28) = 0.74 * 0.72 = 0.53$

M(s1-a)=
$\sum_{s3 \subseteq A} acc_i(A_j) = acc_{k3} * (1 - 0.01) = 0.89 * .99 = 0.88$

M(s1-a)=
$\sum_{s4 \subseteq A} acc_i(A_j) = acc_{k4} * (1 - 0.18) = 0.91 * .82 = 0.75$

M(s1-a)=
$\sum_{s5 \subseteq A} acc_i(A_j) = acc_{k5} * (1 - 0) = 0.90 * 1 = 0.90$

M(s1-a)=
$\sum_{s6 \subseteq A} acc_i(A_j) = acc_{k6} * (1 - 0.04) = 0.92 * .96 = 0.88$

Note: With respect to the pervious classification experiments on different spectrums,

Then $Spt_i(A) = \sum_{A_j \subseteq A} m_i(A_j)$

That equal the summation of all measures of all spectrums with respect to a, as the following equation:

$$Spt_i(A) = \sum_{A_j \subseteq A} (m(s1 - a) + m(s2 - a) + m(s3 - a) + m(s4 - a) + m(s4 - a) + m(s5 - a) + m(s6 - a)$$
= 4.57

For each possible proposition (e.g., user-A), Dempster-Shafer theory gives a rule of combining sensor Si's observation mi and sensor Sj's observation mj

$$(m_i \oplus m_j)(A) = \frac{\sum_{(E_k \cap E_{k'}) = A} m_i(E_k) m_j(E_{k'})}{\sum_{(E_k \cap E_{k'}) = \emptyset} m_i(E_k) m_j(E_{k'})}$$

This combining rule can be generalized by iteration: if we treat mj not as sensor Sj's observation, but rather as the already combined (using Dempster-Shafer combining rule) observation of sensor Sk and sensor $\sum m_k$=5.11

Output (fusion process and Improve accuracy results)

$$\frac{M(s1-a)}{5.11} = 0.12$$

$$\frac{M(s2-a)}{5.11} = 0.103$$

$$\frac{M(s3-a)}{5.11} = 0.172$$

$$\frac{M(s4-a)}{5.11} = 0.15$$

$$\frac{M(s5-a)}{5.11} = 0.18$$

$$\frac{M(s6-a)}{5.11} = 0.17$$

$acc_{k1} = 0.75, acc_{k2} = 0.74, acc_{k3} = 0.89, acc_{k4} = 0.91, acc_{k5} = 0.90, acc_{k6} = 0.92.3$

Journal of Theoretical and Applied Information Technology
31st March 2020. Vol.98. No 06

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Then the sum aggregation for measures with respect to the results of the previous equations

*Table.9: Supports and plausibility's associated with Table 5*

| | Sensor1 | | Sensor2 | | Sensosr3 | | Sensor4 | | Sensor5 | | Sensor6 | | Fused masses (mass m [1,2]) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Spt | Pls | Spt | Pls | Spt | Pls | spt | pls | Spt | pls | Spt | pls | 0.845 |
| A | 84% | 0.84 | 72% | 0.72 | 99% | 0 | 82% | | 90% | 1 | 96% | 0.96 | 0.894 |
| B | 64% | 0.85 | 51% | 0.69 | 79% | 0.89 | 63% | 0.69 | 86% | 0.95 | 89% | 0.97 | 0.83 |
| C | 64% | 0.86 | 72% | 0.97 | 87% | 0.98 | 80% | 0.88 | 89% | 0.99 | 92% | 1 | 0.93 |
| D | 58% | 0.77 | 73% | 0.99 | 36% | 0.41 | 83% | 0.91 | 74% | 0.82 | 92% | 1 | 0.80 |
| E | 73% | 0.97 | 63% | 0.85 | 83% | 0.93 | 91% | 1 | 69% | 0.77 | 86% | 0.93 | 0.89 |
| F | 68% | 0.90 | 74% | 1 | 87% | 0.98 | 90% | 0.99 | 65% | 0.72 | 89% | 0.97 | 0.914 |
| G | 69% | 0.93 | 73% | 0.98 | 88% | 0.99 | 55% | 0.59 | 82% | 0.92 | 56% | 0.61 | 0.99 |
| H | 71% | 0.95 | 7% | 1 | 84% | 0.94 | 87% | 0.96 | 87% | 0.97 | 75% | 0.81 | 0.92 |
| K | 74% | 0.99 | 73% | 0.99 | 81% | 0.91 | 86% | 0.94 | 87% | 0.97 | 81% | 0.88 | 0.94 |
| L | 71% | 0.94 | 73% | 0.99 | 87% | 0.98 | 85% | 0.93 | 88% | 0.98 | 80% | 0.87 | 0.93 |

$$\sum_{A_i \in 2^{\Theta}} 0.12 + 0.103 + 0.172 + 0.15 + 0.18 + 0.17 = 0.89.4$$

This number must be between interval [0,1]

Applying the Dempster-shafer equations on all numerical data in matrix for evaluating the support and plausibility for each sensor.

*Table.10: Mass assignments for the various terrorism objects*

| Target type | Sensor 1 | Sensor 2 | Sensosr 3 | Sensor 4 | Sensor 5 | Sensor 6 |
|---|---|---|---|---|---|---|
| A | 84% | 72% | 99% | 82% | 90% | 96% |
| B | 64% | 51% | 79% | 63% | 86% | 89% |
| C | 64% | 72% | 87% | 80% | 89% | 92% |
| D | 58% | 73% | 36% | 83% | 74% | 92% |
| E | 73% | 63% | 83% | 91% | 69% | 86% |
| F | 68% | 74% | 87% | 90% | 65% | 89% |
| G | 69% | 73% | 88% | 55% | 82% | 56% |
| H | 71% | 7% | 84% | 87% | 87% | 75% |
| K | 74% | 73% | 81% | 86% | 87% | 81% |
| L | 71% | 73% | 87% | 85% | 88% | 80% |

The theory of Dempster-Shafer [41, 42] includes two new concepts that are foreign to Bayes theory. These are the notions of support and plausibility. For example, the support for the target being "fast" is defined to be the total mass of all states implying the "fast" state, which is illustrated in equations (3).

The support is a type of loose lower limit to the uncertainty. On the other hand, a loose upper limit to the uncertainty is the plausibility. This is identified, for the "infrared" state, as the total mass of all states that don't contradict the "infrared" state, which is explained in equation (4). The supports and plausibilities for the masses of Table 2 are given in Table 3. Interpreting:

For measuring the accuracy [39] [40], that requires computing all cases in the previous example to measure precision, recall, and f1-measure.

### 4.3 RESULT DISCUSSION

The main proof relies on the brightness of statistics method integration. Images noticed through various channels suffer variance because of the sensor mechanism, resolution, quantization error, error presented to numerical reconstruction algorithm, etc. Typically, images have characteristics that are likely to seem different from one sensor image to another but are usually closely related. It is significant for observing that the relationships between the image characteristics are in nature. Furthermore, several fusion algorithms are designed effectively due to the Idealism of integrating various images relies on the local relationship between sensor imagery. For example, fusion based on averaging works well for images that are roughly the same except for the collective noise. The integration evidence of the dynamic components and static components circumvents the disparity problem as well as provides a human cognitive favored image evidential representation. An integration and decision scheme are described in detail at the end.

A simple probabilistic classification based on the transfer neural network algorithm that gives roughly 88% of good classification (fig.2.a). The contribution of Dempster-Shafer fusion initialized by the neural network classification algorithm with the pre-trained neural entitled Alexnet and description layer with generating a lexicon for enhancing the classification quality for terrorism domain that extracting from military datasets such as (objects and classes). The average rate of classification is 75% for sensor 1, 74% for sensor2, 89% for sensor 3, 91% for sensor4, 90% for sensor 5, and 92% or sensor 6, that can be interpreted in sum of them approximately 85 % (fig.2 a, b). The contribution of slope information (fig.3) improves the full picture of data based on fusing and combining full vision of multi-spectral images and videos. In fact, there is missing and a loss of data in pictures that can be an obstacle to classify objects automatically because of the types of spectrums especially in night mode. The Dempster-Shafer fusion technique removes this kind of error by decreasing the credibility associated with this decision. The average rate of classification after Dempster-Shafer fusion is 94%. The improvement rate can increase the results by 4% totally and 15%,20% partially for some of the classifications' types

## 5 CONCLUSION AND FUTURE WORK

MSMFT is a multi-spectral fusion model that is constructed to enhance the object detection and classification results. It applies to military and terrorism domains to classify objects and actions. It depends on the combination of deep transfer learning and Dempster-Shafer techniques in decision level fusion. It fuses multiple sensory data into images and videos in night modes with respect to six spectrums (Intensified visual images, Near-infrared spectroscopy (NIR) images, thermal images, LIWR (long-wave infrared images), DHV, and RGB). This system is designed based on an AlexNet pre-trained transfer neural network model. The used neural network consists of six neural networks for six spectrums. This system creates a new algorithm for multi-spectral fusion that introduces integration between machine learning and statistical methods. It solves the lack of datasets problem by generating a purified dataset that consists of tuned six datasets for multiple spectrums in variant type format. It applies the data augmentation for enlarging the dataset that includes 875,970 number of images and video's frames. The fusion accuracy results reach 96 % that increases the classification results by 21% for NIR and VIS spectrums, 7% for Thermal spectrum, and 4 to 6% for LIWR, DHV, RGB spectrums.

**REFRENCES:**

[1] Dong Jiang, et al., Chapter: Survey of Multispectral Image Fusion Techniques in Remote Sensing Applications, In book: Image Fusion and Its Applications, 2011.

[2] Philip N Salter, Survey of multispectral imaging systems for earth observations, Remote Sensing of Environment Vol. 17(1), pp. 85-102, 1985.

[3] Bahador Khaleghi, et al., Multisensor data fusion: A review of the state-of-the-art, information Fusion, Vol. 14 (1), 2013.

[4] A. Karlsson, Dependable and Generic High-level Information Fusion – Methods and Algorithms for Uncertainty Management, School of Humanities and Informatics, University of Skövde, Sweden, Technical Report, pp. 1-23, 2007.

[5] Federico Castanedo, A Review of Data Fusion Techniques, The Scientific World Journal Vol. 2013, pp. 1-20, 2013.

[6] Pratiksha Rashinkar and V. S. Krushnasamy, An overview of data fusion techniques, International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), 2017.

[6] Adnan Akbar, et al., Real-Time Probabilistic Data Fusion for Large-Scale IoT Applications, IEEE Access, Vol. 6, 2018.

[7] Fuyuan Xiao, Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy, Information Fusion, Vol. 46, 2019.

[8] F.Ruzzo and Giovanni Ramponi, Fuzzy methods for multisensor data fusion, IEEE Transactions on Instrumentation and Measurement, Vol. 43(2), pp. 676– 680, 1993.

[9] R.Johnson Suthakar, et al., Study of Image Fusion- Techniques, Method and Applications, International Journal of Computer Science and Mobile Computing (IJCSMC), Vol. 3(11), pp.469 – 476, 2014.

[10] Pham, I., and Polasek, M., Algorithm for military object detection using image data, IEEE/AIAA 33rd Digital Avionics Systems Conference (DASC), 2014.

[11] Sang-Il Oh and Hang-Bong Kang, Object Detection and Classification by Decision-Level Fusion for Intelligent Vehicle Systems, Sensors, 2017.

[12] Xue-Bo Jin, et al., Advances in Multi-Sensor Information Fusion: Theory and Applications 2017, sensors, 2018.

[13] Yunqian Ma and Michael E. Bazakos, Multi-spectral fusion for video surveillance, Proceedings of the Fourth IEEE, Knowledge-based multi-spectral pixel level fusion for surveillance, 2006.

[14] Jennifer L. Bittner, et al., The effect of multispectral image fusion, Cognitive Research: principles and Implications, 2017.

[15] Preema Mole and M Mathurakani, Colour Image representation of Multispectral Image Fusion, Advanced Computing: An International Journal (ACIJ), Vol.7 (3), 2016.

[16] Xiaoyong Shen, et al., Multispectral Joint Image Restoration via optimizing a scale map, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 37 (12), 2015.

[17] Abdi, G., Samadzadegan, F., and Reinartz, P., A decision-based multi-sensor classification system using thermal hyperspectral and visible data in urban area. European Journal of Remote Sensing, Vol. 50 (1), pp. 414–427, 2017.

[18] Na Li, Arnaud Martin, and Remi Estival, an automatic water detection approach based on Dempster-Shafer theory for multi-spectral images, arXiv:1708.02747v2 [cs.AI], 2017.

[19] R.P.S. Mahler, Statistical Multisource-Multitarget Information Fusion, Artech House, 2007.

[20] Zuhe Zhang, Benjamin I.p. Rubinstein, and Christos Dimitrakakis, On the differential privacy of Bayesian inference, AAAI'16 Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, PP. 2365-2371, 2016.

[21] Goodman, I,R,, Mahler, R.p., and Hung T.Nguyen, Book: Mathematics of Data Fusionm, Springer, 1997.

[22] Christophe Simon, Philippe Weber, and Alexandre Evsukoff., Bayesian networks inference algorithm to implement Dempster Shafer theory in reliability analysis. Reliability Engineering and System Safety, Vol. 93 (7), pp.950-963, 2008.

[23] Subhash Challa and Don Koks, Bayesian and Dempster–Shafer fusion, Vol. 29 (2), pp. 145–176, 2004.

[24] Zhi Yang, et al., Deep Transfer Learning for Military Object Recognition under Small Training Set Condition, Neural Computing and Applications, Vol. 31, pp. 6469–6478, 2019.

[25] Wu Zhihuan, et al., Rapid Target Detection in High Resolution Remote Sensing Images Using YOLO Model, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, ISPRS TC III Mid-term Symposium "Developments, Technologies and Applications in Remote Sensing", Vol. XLII-3, China, 2018.

[26] Krebs, S., Duraisamy, B., and Flohr, F., A survey on leveraging deep neural networks for object tracking. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 2017.

[27] Chuanqi Tan, et al., A Survey on Deep Transfer Learning, arXiv:1808.01974v1 [cs.LG], 2018.

[28] Romain Mormont, Pierre Geurts, and Raphael Maree, Comparison of deep transfer learning strategies for digital pathology, IEEE, 2018.

[29] Zhi Yang, et al., Deep transfer learning for military object recognition under small training set condition, place, Vol.xx (x), 2018.

[30] Ali abd ALmisreb, Nursuriati Jamil, and Norashidah Md din, Utilizing AlexNet Deep Transfer Learning for Ear Recognition, Fourth International Conference on Information Retrieval and Knowledge Management (CAMP), 2018.

[31] Shuihua Wang, et al., Alcoholism Identification Based on an AlexNet Transfer Learning Model, Frontiers in Psychiatry 10, 2019.

[32] Alexander Toet, Alan R Pinkus, and Maarten A.Hogervost, The TRICLOBS Dynamic Multi-Band Image Data Set for the development and evaluation of image fusion, Plos One, 2016.

[33] Bilodeau, G.-A., et al., Thermal-Visible Registration of Human Silhouettes: a Similarity Measure Performance Evaluation, Infrared Physics and Technology, Vol. 64, pp. 79-86, 2014.

[34] Justin Lai and Syndeny Maples, Developing a Real-Time Gun Detection Classifier, ACM, 2017.

[35] ZUKAL, M.; MEKYSKA, J.; ČÍKA, P.; SMÉKAL, Z. Interest Points as a Focu s Measure in Multi-Spectral Imaging. Radio engineering, Vol. 22 (1), pp. 68-81, 2013.

[36] FLIR Systems, Inc. FLIR ONE is Lightweight, Easy to Connect and Easy to Use, 2015.

Retrieved from FLIR: https://www.flir.eu/oem/adas/adas-dataset-form/

[37] IEEE OTCBVS WS Series Bench; Roland Miezianko, Terravic Research Infrared Database, 2013.

[38] Khalifa, N. E. M., et al., "Deep Iris: Deep Learning for Gender Classification Through Iris Patterns", Acta Informatica Medica, Vol. 27 (2), pp. 96-102, 2019.

[39] Chuanqi Tan, et al., A Survey on Deep Transfer Learning, arXiv:1808.01974v1 [cs.LG], 2018.

[40] Valentin Radu, et al., Multimodal Deep Learning for Activity and Context Recognition, Proceeding ACM Interact. Mob. Wearable Ubiquitous Technol. 1, 4, Article 157, pp.1-27, 2017.

[41] Ofir, N., Silberstein, et al., Registration and Fusion of Multi-Spectral Images Using a Novel Edge Descriptor, 25th IEEE International Conference on Image Processing (ICIP), 2018.