Neutrosophic Sets and Systems

Volume 45 Article 8

8-28-2021

Multimodal Fusion: A Review, Taxonomy, Open Challenges, Research Roadmap and Future Directions

Mohd Anas Wajid

Aasim Zafar

NSS
Neutrosophic Sets and Systems, Vol. 45, 2021
University of New Mexico

Multimodal Fusion: A Review, Taxonomy, Open Challenges, Research Roadmap and Future Directions

Mohd Anas Wajid[1], Aasim Zafar[2]

[1] Department of Computer Science, Aligarh Muslim University, Aligarh, 202002; anaswajid.bbk@gmail.com
[2] Department of Computer Science, Aligarh Muslim University, Aligarh, 202002; aasimzafar@gmail.com

Abstract: The present work collects a plethora of previous research work in the field of multimodal fusion which despite a lot of research could not handle the imperfections. These imperfections could be at any stage initiating from the imperfections in data and its sources to imperfections in fusion strategies. Further, the work explores various applications of Neutrosophy in the field of handling imperfections along with description of previous work in this regard. These applications include the one which addresses the notion of imperfection and uncertainty among multimodal data which is being collected for fusion. In this way, the present work tries to incorporate neutrosophic logic and its applications in the field of computer vision including multimodal data fusion and information systems. It is assumed that if the notion of uncertainty is included in multimodal research, the development of newer algorithms for solving the problems of imperfections in multimodal systems will provide impetus to the existing research in this field.

Keywords: Multimodal Data; Multimodal Fusion; Imperfections; Fuzzy Logic; Neutrosophic Logic; Machine Learning.

1. Introduction

The present world is witnessing a change where the user is not only a consumer of information but a great producer of it. Earlier website owners were the main source of information production but in the current scenario, the social web has taken its position. This rapid development in the field of the web is termed Web 2.0. The repositories of multimedia content (Flicker, YouTube, Picasa and Twitter etc.) over the web is increasing at a faster pace than ever before. This plethora of content over the web as well as on personal computers has raised the issue of its effective storage, organization, indexing and retrieval. This multimedia content (image or video) has a multimodal (visual, textual) nature. These multimodalities are of utmost importance since the information conveyed by pixels covers only visual content which is totally different from tag information. These modalities must be combined in such a way that it gives more of the information needed on time. In order to combine the above-mentioned modalities it is important to consider the process of information fusion. This process at the initial level is carried forward in different ways. These may be data fusion (low level), feature fusion (intermediate level) and decision fusion (high level). When multiple sources of raw data are combined in such a way that the new source is more informative and synthetic than the

previous two, it is called data fusion. Feature fusion combines features extracted from different sources into a single stand-alone feature vector. Decision fusion is clearly based on classifiers when aid in giving unbiased and accurate results. One of the main characteristics of the fusion process is imperfection as explained by Bloch [2001]. These imperfections are the main reason for the fusion process to be carried out more effectively [22]. These imperfections could mainly be imprecision, uncertainty and incompleteness. These imperfections occur at a different level of fusion. In this paper we have reviewed work on multimodal fusion, also we have reviewed work on neutrosophic technologies which could be employed in the field of multimodal fusion and systems. The following figure shows the workflow:

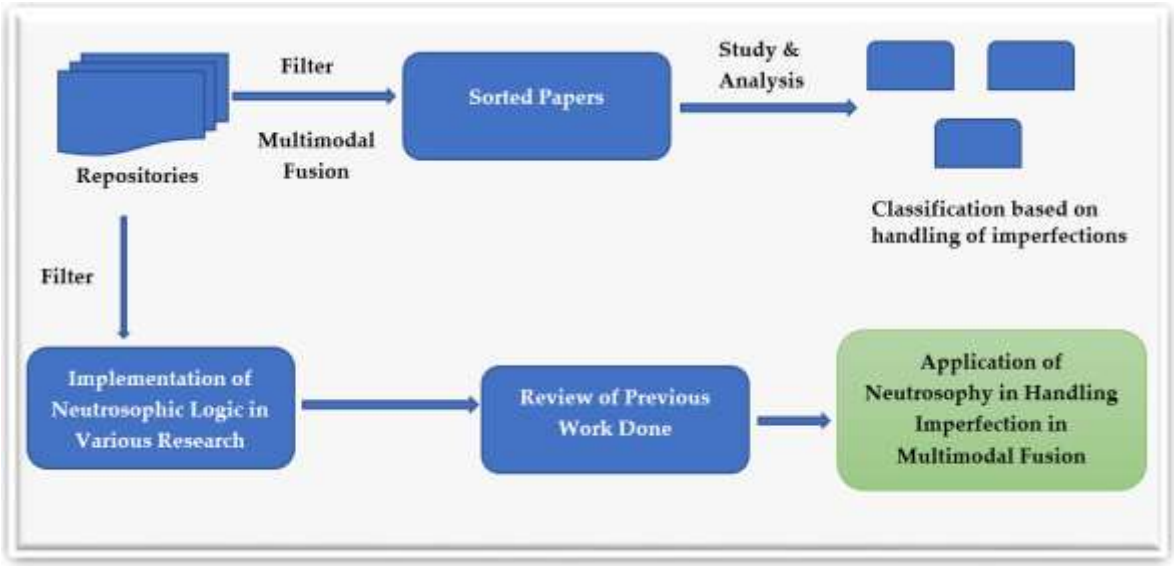

Figure 1 Block diagram for the process of research being adopted in the present manuscript

1.1 Background:

Multimodal fusion has been attaining exponential attention in multimodal information access and retrieval tasks and this has been well studied by Kludas and Marchand-Maillet [2011]; Souvannavong et al. [2005]; Marchand-Maillet et al. [2010] and Niaz and M´erialdo [2013] [23-26]. The facts related to this can be found in the study done by Atrey et al. [2010] [27]. The imperfections in textual modality are partially considered in the context of multimodal systems. Most of the state-of-the-art approaches which address the notion of textual imperfections always do so using relevancy e.g. imperfections in tags. The incompleteness issue has been well identified in the literature by Liu et al. [2009]; Tang et al. [2009] [28-29]; Wang et al. [2010] [40] but work on imprecision and uncertainty is left far behind e.g. noisy tags. Imperfection in different modalities has been studied by Bloch [2003] [41]. Though various research work has been done in the field of multimodal fusion but very little has focused on handling imperfections. Handling imperfections does not appear their primary goal while performing multimodal fusion tasks. Below Table 1 shows the work in above-mentioned field using different principles by various researchers around the globe.

Table 1 A Summary of the state-of-the-art approaches for Multimodal Fusion

S No.	Work	Principle	Handling Imperfections
1.	Romberg et al. (2012)	Probabilistic latent semantic analysis on tags co-occurrence matrix.	No
2.	Zhang et al.(2012)	Semantic BOW based on the Tag-to-Tag similarity	The incomplete data problem.
3.	Xioufis et al. (2011) .	Binary BOW representation with feature selection	No
4.	Kawanabe et al. (2011)	Binary BOW representing the presence/absence of tags with random walks over tags.	The incomplete data problem.
5.	Guillaumin et al. (2010)	Binary BOW representation representing the presence/absence of tags.	No
6.	Liu et al. (2013)	Histogram of Textual Concepts based on the Tag-to-Concept similarity	The incomplete data problem.
7.	Nagel et al. (2011)	BOW based on the tf-idf values of tags.	No
8.	Li et al. (2010)	Compare Tag and annotation concepts expansion vectors.	No
9.	Gao et al. (2010)	Probability based on the tag-concept co-occurrence.	No
10.	Wang et al. (2010)	Semantic Fields based on the tag-concept co-occurrence.	No
11.	S. Poria et al. (2015)	Aggregate semantic and affective information associated with data	No
12.	S. Poria et al. (2015)	Decision level data fusion	No
13.	S. Poria et al. (2016)	Deep neural network & multiple kernel learning classifier	No
14.	Minghai Chen et al. (2017)	Modality fusion at word level	No

15.	Kyung-Min Kim et al. (2018)	Residual learning fusion	No
16.	Feiran Huang et al. 2019	internal correlation among features (textual & visual)for joint sentiment classification	No

Above mentioned approaches whether related to early fusion, late fusion or transmedia fusion, do not tackle the problem of imperfections at the feature level. Now after defining the imperfections or uncertainties at various levels of fusion, let us understand from Table 2 what are the terms being used to describe these data/information imperfections by prominent researchers in their work.

Table 2 A summary of terms used to describe tag imperfections in Multimodal Fusion

S No.	Work	Imperfections terms used
1.	Jin et al. (2005)	Noisy
2.	Weinberger et al. (2008)	Ambiguous
3.	Xu et al. (2009)	Ambiguous
4.	Wang et al. (2010)	Incomplete, Ambiguous
5.	Liu et al. (2009)	Incomplete, Imprecise, Noisy
6.	Kennedy et al. (2009)	Unreliable, Noisy
7.	Tang et al. (2009)	Incorrect, Noisy, Incomplete
8.	Liu et al. (2010)	Incomplete, Biased, Incorrect
9.	Zhu et al. (2010)	Noisy
10.	Yang et al. (2011)	Ambiguous, Noisy
11.	Wu et al. (2012)	Inconsistent, Noisy, Incomplete, Unreliable
12.	Valentin Vielzeuf et al. (2017)	Noisy
13.	Natalia Neverova et al. (2014)	Uncertain, Noisy

14.	A. Tamrakar (2012)	Noisy
15.	Kyung-Min Kim et al. (2018)	Ambiguous
16.	Feiran Huan et al. (2019)	Inconsistent, Noisy, Incomplete,
17.	Yagya Raj Pandeya and Joonwhoan Lee (2019)	Lack of labelling

The motivation behind carrying out present work is the negligence of research community towards addressing the problem of imprecision in data, which is used in designing multimodal systems. This imprecision arises due to dependence of classifiers on incomplete and uncertain data that leads to an imprecise decision function. This also happens when scores produced by different classifiers are combined, fusion faces the problem of imperfection. These imperfections could lead to imperfections in machine learning algorithms at the decision level. To achieve our goal we have described various work done by researchers in multimodal fusion at each stage i.e. early fusion, late fusion and transmedia fusion. We have also explained the terms that are used for showing imperfections and imprecision in data by various researchers. Further, we have explained the neutrosophic theory which provides a way to deal with uncertainty, imprecision and imperfections, with a detailed description of work carried out using this theory to address the imperfections at various levels. Our aim is to acknowledge the problem of uncertainty and imprecision in multimodal fusion tasks and introduce new researchers working in the concerned field to the notion of Neutrosophy and its applications in handling imperfections in multimodal fusion.

The taxonomy of research challenges and opportunities in multimodal fusion together with potential research challenges are summarized and highlighted in this article. The main objectives of this work include:

- To identify the problem of imperfections in multimodal fusion. Interpreting existing research conducted in this domain.
- To interpret current studies conducted in this area of research.
- To identify a research gap in the field that needs to be further investigated by the researchers in the field.
- To identify and introduce new researchers with the concept of neutrosophy and its applications in multimodal fusion.
- To identify roadmap that requires investigation in future by concerned researchers in the field of multimodal information systems.

The rest of the paper is divided into five sections, Section 2 explains research conducted in the field of multimodal fusion, including early fusion, late fusion and transmedia fusion. It also explains the problem of imperfection encountered in multimodal fusion. Section 3 introduces new researchers working in the field of multimodal information access and retrieval with the concept called Neutrosophy. It also describes the current research conducted in the field of neutrosophy which

could be employed in multimodal systems for handling imperfections. Section 4 summarizes and highlights the future research roadmap to multimodal fusion using Neutrosophy. Section 5 concludes the work.

2. Taxonomy of open issues and challenges in multimodal fusion approaches:

Multimodal fusion is one of the important steps in multimodal information access and retrieval. The accuracy of the framework depends on fusion strategies being adopted. Researchers dealing with multimodal fusion mainly use three strategies namely;

1. Early Fusion
2. Late Fusion
3. Transmedia Fusion

These strategies are well explained with their working principle in the work carried out by Mohd Anas Wajid and Aasim Zafar [2019] [83]. The above-mentioned fusion strategies could easily be understood by the following diagram.

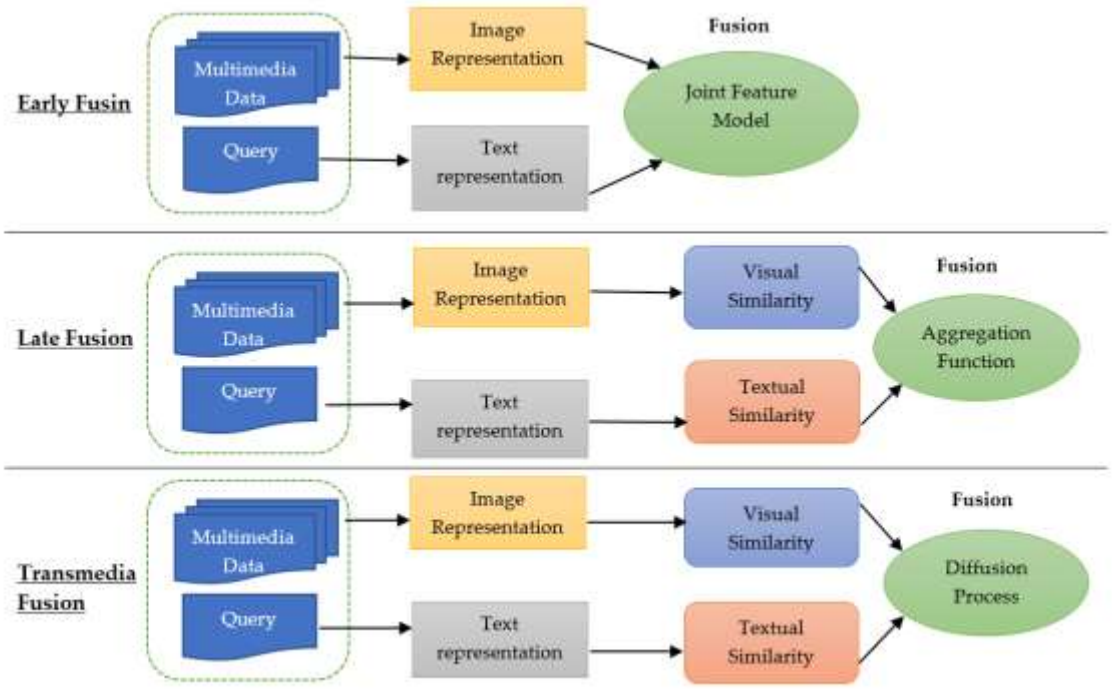

Figure 2 A summary of Fusion Approaches for Multimodal Fusion

2.1 Early Fusion:

The early fusion strategy has been adopted by a number of researchers around the globe. Though effective in many ways, it does not address the issue of imperfections while handling the data. Some of the prominent work using early fusion are explained below and later compared in the table on the basis of their fusion methodology being adopted.

Li et al. [2009] have used simple concatenation of visual aspects together with textual aspects of data for fusion [42]. Duygulu et al. [2002] have employed the correlation concept using an estimation maximization algorithm (EM). This is used for attaching words to the segmented image regions after the training phase is over. The model proposed is called the translation model [43].

Barnard et al. [2003] on one hand studied the joint distribution among the textual modality and the segmented image modality and on the other hand, used it to convert it in likelihood function between text and segmented image region [44].

According to Blei and Jordan [2003] Probabilistic Latent Semantic Analysis (pLSA) and Latent Dirichlet Allocation (LDA) can be used for correspondence between textual modality and the corresponding image modality. The model proposed by them is estimated using the EM algorithm [45].

Monay and Gatica-Perez [2003] in a similar fashion have used pLSA over the concatenated set of visual and textual modalities. The balance between the two modalities limits the size of visual representation [46]. A pLSA based model proposed by Lienhart et al. [2009] is again used to retrieve information in multimodal retrieval systems [47].

A. Tamrakar et al. [2012] have used BoW descriptors within Support Vector Machine (SVM). This is done for event detection and for this they have used many early and late fusion strategies [66]. Minghai Chen et al. [2017] have done multimodal fusion at the word level. They have emphasized Temporal Attention Layer for predicting sentiments in sentimental analysis. They have also described the noise that is present in data of different modality [39].

L. Morency et al. [2011] have proved the effectiveness of using different modalities for sentiment analysis. Though they have shown how the internet could be a source of information while using different modalities like audio, video and text but they have failed to address the imperfections present in data while carrying out the multimodal fusion [48].

How the error in sentiment classification is reduced while taking into consideration the combination of different modalities is studied by V. Pérez-Rosas [2013]. In their study authors have stressed that while using a single modality the error is 10% high as compared to using various modalities together [49]. A recent development in sentimental analysis and emotion recognition has been recorded by S. Poria et al. [2016]. Here authors have performed Emotion recognition and sentiment analysis using convolution MKL and SVM based approach [50].

The following table shows the work done by prominent researchers in the field of information retrieval using the early fusion strategy. Though all the approaches have their significance in fusing multimodal data yet they fail on the grounds of handling imperfections in data.

Table 3 A Summary of Approaches Based on Early Fusion

S No.	Work	Fusion Method	Handling Imperfections	Fusion Level
1.	Li et al. (2009)	concatenation of textual & visual modality	No	
2.	Duygulu et al. (2002)	Assigning words to segmented image regions using translation model.	No	
3.	Barnard et al. (2003)	joint distribution of text and segmented image	No	
4.	Blei and Jordan (2003)	LDA on visual and textual modality.	No	
5.	Monay and Gatica-Perez (2003)	pLSA for concatenating visual and textual features. .	No	
6.	Lienhart et al. (2009)	Multimodal pLSA multilayer model.	No	
7.	Chandrika and Jawahar (2010)	Multimodal pLSA	No	
8.	Nikolopoulos et al. (2013)	High Order pLSA.	No	**Early Fusion**
9.	Wang et al. (2009)	Visual tag dictionary using GMM.	No	
10.	A. Tamrakar (2012)	Event detection using BoW descriptors within SVM.	No	
11.	Minghai Chen et al. (2017)	Gated Multimodal Embedding LSTM with Temporal Attention	No	
12.	L. Morency et al. (2011)	Tri-model sentiment analysis using Gaussian mixtures and HMM	No	
13.	V. Pérez-Rosas et al. (2013)	BoW and OpenEAR an open source software with SVM	No	
14.	S. Poria (2016)	Emotion recognition using convolution MKL based approach	No	

***2.2 Late Fusion*:**

Now we describe multimodal fusion that is based on late fusion strategy. There exist a plethora of work using this strategy but they are all based on different methodologies which are discussed further.

Xioufis et al. [2011] in their approach worked in a different fashion. They introduced a multimodal fusion strategy based on late fusion. Their approach is totally based on predictions obtained by the classifiers from visual features. These predictions obtained from visual modality are averaged. Further, these are averaged with the predictions obtained from textual modality [82].

Wang et al. [2009a] worked in line with SVM where the scores from different classifiers are fed to SVM. The authors proposed to build two classifiers one for text modality and the other for visual modality. A third classifier is introduced to combine the confidence of the previous two and give final predictions [56].

Guillaumin et al. [2010] work with MKL framework which is considered to be a success for the feature fusion method. In the first step, their proposed semi-supervised method exploits both textual and visual features for learning a classifier. Later MKL framework is employed to predict text modality based on the visual content provided [33].

Kawanabe et al. [2011] have used a similar approach however it differs from the use of MKL. It deploys trained SVMs and uniform kernel weights and gives results approximately the same as MKL method [58]. Zhang et al. [2012a] have used the same method for combining kernels learned on textual and visual features [31].

Gao et al. [2010] have adopted a technique based on feature selection using Grouping Based Precision & Recall-Aided (GBPRA) in classifier combination which enriches the performance of classification [60]. Liu et al. [2011] have used Dempster's rule for combining classifiers predictions to achieve the best classification results [62]. Liu et al. (2013) worked on a fusion scheme termed Selective Weighted Late Fusion (SWLF). It works towards enhancing the mean average precision by selectively choosing the weights which in turn enhances the optimization [61].

Daeha Kim et al. (2017) have worked towards classifying human emotions using multimodal signals and neural networks. Though the data which they have used comprises of landmark, audio and image having various imperfections at fusion level but these are not been handled [63]. Moving towards a similar goal of emotion recognition, Valentin Vielzeuf et al. (2017) have explored several multimodal fusion strategies. They have used a supervised classifier to know emotion labels and later proposed 2D and 3D Convolution Neural Network approaches for better face descriptors [64].

Natalia Neverova et al. (2014) have worked towards gestures identification giving more stress on modality initialization and later on their fusion using late fusion strategy [65]. The work to use late fusion with dual attention mechanism has been mentioned by Kyung-Min Kim et al. (2018). This

approach is utilized in proposing an architecture that could be utilized in designing effective Question-Answering (Q & A) systems [67].

Feiran Huan et al. (2019) have proposed a Deep Multimodal Attentive Fusion (DMAF), for sentimental analysis using data from social media platforms. Authors have used late fusion strategy for an effective fusion of modalities like image and text but when it comes to handling imperfections their approach seems to be lacking on this ground [68]. The work done by Escalante et al. [2008] is totally based on predictions obtained from classifiers. These are learned on textual and visual modalities and later combined in a linear way [54].

Getting inspired by the music-video combination, Yagya Raj Pandeya and Joonwhoan Lee (2019) have prepared a dataset that could be effectively utilized for sentiment analysis. In their approach, they have extracted features of music and video separately, later characterized using long short-term memory (LSTM) and for evaluating the emotions various machine learning algorithms are used [69].

Though all approaches have shown remarkable results in terms of the fusion of different modalities however they all lack on similar grounds i.e. handling imperfections. The Table 4 presents a summary of work using late fusion strategy compared on the basis of fusion method adopted.

Table 4 A Summary of Approaches Based on Late Fusion

S No.	Work	Fusion Method	Handling Imperfections	Fusion Level
1.	Escalante et al. (2008)	Prediction by different classifier is combined.	No	
2.	Xioufis et al. (2011)	Average rule used for late fusion.	No	
3.	Wang et al. (2009)	Predicted features are concatenated and SVM classifier is used.	No	
4.	Guillaumin et al. (2010)	Multiple Kernel Learning.	No	
5.	Kawanabe et al. (2011)	Multiple Kernel Learning.	No	
6.	Zhang et al. (2012)	Multiple Kernel Learning.	No	
7.	Gao et al. (2010)	Feature selection using Grouping Based Precision & Recall-Aided (GBPRA).	No	
8.	Liu et al. (2013)	Late fusion using selective weight	No	

9.	Liu et al. (2011)	Classifier predictions combined using Dempster's rule	Yes	
10.	Daeha Kim et al. (2017)	Semi supervised learning and neural network	No	**Late Fusion**
11.	Valentin Vielzeuf et al. (2017)	Temporal multimodal fusion	No	
12.	Natalia Neverova et al. (2014)	Multi-scale deep learning and localization	No	
13.	A. Tamrakar (2012)	BoW descriptors within SVM.	No	
14.	Kyung-Min Kim et al. (2018)	Residual learning fusion	No	
15.	Feiran Huan et al. (2019)	Internal correlation among features (textual & visual)for joint sentiment classification	No	
16.	Yagya Raj Pandeya and Joonwhoan Lee (2019)	Pre-trained neural networks	No	

2.3 Transmedia Fusion:

Transmedia fusion is also referred to as intermediate fusion or cross-media fusion. The basic notion of its functioning is to use visual features to accumulate image modality (Visually Nearest Neighbor) and later switch to the textual modality to collect features from the neighbors. All the approaches towards achieving transmedia fusion are listed in Table 5. It also mentions the fusion method being employed by the researchers. Though the results of the work are fully satisfying the goal of transmedia fusion; it does not handle imperfections present in different data modalities.

Table 5 A Summary of Approaches Based on Transmedia Fusion

S No.	Work	Fusion Method	Handling Imperfections	Fusion Level
1.	Makadia et al. (2008)	Nearest neighbors using Joint Equal Contribution.	No	
2.	Torralba et al. (2008)	Grasping texts from neighbors.	No	**Transmedia Fusion**

3.	Guillaumin et al. (2009)	Metric learning for text propagation.	No
4.	Li et al. (2009)	Votes are accumulated for tag relevance	No
5.	Feiran Huan et al. (2019)	internal correlation among features (textual & visual) for joint sentiment classification	No
6.	Daeha Kim et al. (2017)	Neural networks based on multimodal signals	No
7.	Valentin Vielzeuf et al. (2017)	Supervised classifier based on audio-visual signals	No
8.	Natalia Neverova et al. (2014)	Gesture detection using multimodal and multiscale deep learning	No
9.	A. Tamrakar (2012)	using BoW descriptors within an SVM approach for event detection	No

3. Taxonomy of Research Work Handling Imperfections Using Neutrosophy:

Neutrosophic logic has gained alarming attention since its inception. At present is has left no areas of research untouched. Researchers all around the globe are employing its tools and techniques for the computation of uncertainty and imprecision which was a problem since time immemorial [57] [85-87]. But with the advent of neutrosophic sets and theory, the days are not far for computational intelligence to achieve its verge with the address of uncertainty and indeterminacy in machine learning algorithms and models. This theory was proposed by Florentin Smarandache [2005] which is extensively used since then for handling imperfections at various levels in mathematics and computer science. It is also referred to as Smarandache's logic [84]. It states that a proposition could have values in the range of [T, I, F] where T refers to membership degrees of truth, I refers to membership degrees indeterminacy and F refers to membership degrees falsity. Bouzina Salah (2016) have compared operational fuzzy logic to that of neutrosophic logic. The authors have shown how in fuzzy logic the membership of truth and falsity gets changed into truth, falsity and indeterminacy in neutrosophic logic. The authors argue that how a change in principle changes the whole system of working [10].

Now we describe some of the work performed by researchers using neutrosophic sets and systems. These works show that if we employ their strategy at an early stage in multimodal fusion

then the problem of imperfections could easily be handled while carrying out this task. This would also enable us to remove imperfections in machine learning algorithms which in turn will not be transmitted to the modelling stage and our information access and retrieval will be more quick and accurate. Now let us understand how this work is carried out and what are strategies being followed by the researchers.

Ned Vito Quevedo Arnaiz et al. [2020] have proposed a method for dealing with unlabeled data. Their approach involves the usage of neutrosophic sets and systems. The treatment of unlabeled data is done by developing unsupervised Neutrosophic K-means algorithm. Their work is motivated due to the increasing amount of unlabeled data over the internet. The authors have taken data for experiments from a stored dataset of the City of Riobamba to show the effectiveness of their methodology [1].

Mouhammad Bakro et al. (2020) in their paper have adopted a neutrosophic approach to digital images. The elements of image modality are represented in the neutrosophic domain by dividing points of the image matrix into neutrosophic sets. The authors have also studied various methods and metrics for calculating similarity and dissimilarity between image modality. The authors have claimed that their approach would enable researchers in searching inside images and videos [2].

Abhijit Saha et al. (2020) have addressed the problem of incomplete data using neutrosophic soft sets taking in account various suitable examples. The authors have explained the inconsistent and consistent association among various parameters followed by definitions such as consistent association degree, consistent association number between the parameters, inconsistent association number between the parameters and inconsistent association degree to measure these associations. They have also proposed a data filling algorithm and proved its feasibility and validity [3].

Carmen Verónica Valenzuela-Chicaiza, et al. (2020) have done an analysis of emotional intelligence using Neutrosophic psychology. The experiment is carried out using 245 randomly selected students at the Autonomous University of Los Andes [4].

Ridvan Sahin (2014) have worked in achieving a Hierarchical clustering algorithm based on neutrosophy. This is achieved by extending algorithms proposed for Intuitionistic Fuzzy Set (IFS) and Interval Valued Intuitionistic Fuzzy Set (IVIFS) to Single Valued Neutrosophic Set (SVNS) and Interval Neutrosophic Set (INS). They have extended the algorithm for classifying neutrosophic data to show its effectiveness and applicability [5].

Yaman Akbulut et al. (2017) have worked towards enhancing the classification performance of k-Nearest Neighbour (k-NN) by the introduction of Neutrosophy. The authors have introduced Neutrosophic-k-NN. The authors have tested their approach on various datasets and have found good classification results as compared to k-NN [19]. Wen Ju et al. (2013) have introduced the Neutrosophic Support Vector Machine (N-SVM) [20].

A. A. Salama (2014) have done significant work in the domain of image processing by employing Neutrosophy in the field. They have proposed techniques to address imperfectly defined image modality. The authors have also worked towards similarity metrics for neutrosophic sets like Hamming distance and Euclidian distance. Possible applications to image processing are also touched upon [6]. The authors in the same year have worked extensively to introduce the researchers with neutrosophic linear regression and correlation [7].

Anjan Mukherjee et al. (2015) has studied Neutrosophy and its application in the field of pattern recognition. The authors have proposed a weighted similarity measure between two neutrosophic soft sets and verified its application in recognizing patterns in computer vision problems by taking some suitable examples [8].

A. A. Salama et al. (2016) have represented image modality features in the neutrosophic domain. For this purpose authors have stressed on textual modality. The authors have used these features extensively in training the model so that it could easily be used in image processing tasks [9].

Nguyen Xuan Thao et al. (2017) in their work mentioned various applications of Soft Computing. The authors have introduced a new concept of Support Neutrosophic Set (SNS) which is a combination of fuzzy set and neutrosophic set. They have also described the operations of these sets together with their properties [11].

Okpako Abugor Ejaita et al. (2017) have studied the uncertainties in medical diagnosis. Authors have stressed how negligence of uncertainty at the initial stage of diagnosis could lead to fatal problems in patients at a later stage. To overcome these authors have introduced a framework based on Neutrosophic Neural Network for diagnosis of confusable disease [12].

A. A. Salama et al. (2018) have worked towards enhancing the quality of image modality. For this reason authors have converted the image in the neutrosophic domain so that their contrast could be enhanced. This approach to the neutrosophic grayscale image domain would enable image processing to yield good results while performing information retrieval [13]. To achieve the same goal Ming Zhang et al. (2010) have proposed an image segmentation approach based on Neutrosophy [16]. Abdulkadir Sengur and Yanhui Guo (2011) have done colour, texture image segmentation based on neutrosophic set and wavelet transform [17].

D. Vitalio Ponce Ruiz et al. (2019) have introduced a new concept of linguistic modelling in Neutrosophy. This is done to remove the uncertainty which seems to be a big hurdle while modelling linguistic terms while performing information retrieval. The modelling is performed using LOWA operator. Their work seems to be a milestone achieved for modelling linguistic modality in multimodal systems [14].

Elyas Rashno et al. (2019) have worked towards recognizing noisy speech. Their approach of recognition employs Convolution Neural network (CNN) model based on Neutrosophy. They have

proposed Neutrosophic Convolution Neural Network (NCNN) claiming that this would ease the task of classification [18].

G. Jayaparthasarathy et al. (2019) have discussed various applications of Neutrosophy in data mining. To illustrate their objective, authors have taken the medical domain as their field of research [15]. A survey of machine learning in neutrosophic environment is presented by Azeddine Elhassouny et al. (2019) [59].

Kritika Mishra et al. (2020) have performed sentiment analysis using neutrosophy. Their proposed framework works with audio files and calculates their Single-Valued Neutrosophic Sets (SVNS) and clusters them into positive-neutral-negative. Later, obtained results from the above tasks are combined with sentiment analysis results obtained from textual files of the same audio file. Their approach seems to yield good results [21]. Table 6 presents a summary of work using neutrosophy giving more stress on author's contribution in the field for handling imperfections.

Table 6 A Summary of Research Work for Handling Imperfections Using Neutrosophy

S No.	Author & Year	Primary Contribution	Handling Imperfections
1.	Ned Vito Quevedo Arnaiz et al. (2020)	Developing Neutrosophic K-means based method for treatment of unlabelled data.	Yes
2.	Mouhammad Bakro et al. (2020)	• Neutrosophic representation of digital image. • Points of digital picture matrix converted into neutrosophic sets.	Yes
3.	Abhijit Saha et al. (2020)	• Described neutrosophic soft sets having incomplete data. • Described consistent and inconsistent association between parameters.	Yes
4.	Carmen Verónica Valenzuela-Chicaiza, et al. (2020)	Classical statistical inference tools for emotional intelligence.	Yes

5.	Ridvan Sahin (2014)	Hierarchical clustering algorithm based on Neutrosophy.	Yes
6.	A. A. Salama et al. (2014)	Image modality processing using Neutrosophy.	Yes
7.	A. Salama et al. (2014)	Introduced neutrosophic simple regression and correlation.	Yes
8.	Anjan Mukherjee et al. (2015)	• Application of Neutrosophy in pattern recognition. • Proposed weighted similarity measure between two neutrosophic soft sets.	Yes
9.	A. A. Salama et al. (2016)	Representing features of image modality in neutrosophic domain.	Yes
10.	Bouzina, Salah (2016)	Compared fuzzy logic with neutrosophic logic.	Yes
11.	Nguyen Xuan Thao et al. (2017)	Introduces Support Neutrosophic Set (SNS).	Yes
12.	Okpako Abugor Ejaita et al. (2017)	• Addressed uncertainties in medical diagnosis using Neutrosophy. • Introduced a framework based on Neutrosophic Neural Network.	Yes
13.	A. A. Salama et al. (2018)	Introduced an approach to grayscale image in neutrosophic domain.	Yes

14.	D. Vitalio Ponce Ruiz et al. (2019)	• Treatment of uncertainty while retrieving information. • Linguistic modelling using Neutrosophy.	Yes
15.	G. Jayaparthasarathy et al. (2019)	Applications of Neutrosophy in data mining.	Yes
16.	Azeddine Elhassouny et al. (2019)	Presented a survey of machine learning in neutrosophic environment.	Yes
17.	Ming Zhang et al. (2010)	A neutrosophic approach to image segmentation.	Yes
18.	Abdulkadir Sengur &Yanhui Guo (2011)	Color, texture image segmentation based on neutrosophic set and wavelet transform.	Yes
19.	Elyas Rashno et al. (2019)	• Worked to recognize noisy speech. • A Convolution Neural Network model based on Neutrosophy.	Yes
20.	Yaman Akbulut et al. (2017)	• Enhanced classification performance of k-NN by the introduction of Neutrosophy. • Introduced Neutrosophic-k-NN.	Yes
21.	Wen Ju et al. (2013)	Introduced Neutrosophic Support Vector Machine.	Yes
22.	Kritika Mishra et al. (2020)	Sentiment analysis using Neutrosophy.	Yes

4. Future Research Trends and Directions for Handling Imperfections in Multimodal Fusion:

Based on our literature investigation and analysis of more than 80 articles, various research trends, research directions, and potential research topics are drawn for handling imperfections in multimodal fusion research and development. Though the direct handling of imperfection at the fusion stage will not yield fruitful results, we recommend handling imperfections at each stage starting from selecting data sources and collection of data in different modalities to fusing the features together. The procedure involved in this is summarized in the following Figure 3.

Figure 3 Potential Trends and Future Directions for Multimodal Fusion Research and Development Using Neutrosophy

5. Conclusion

The plethora of digital data over the internet has surged the need of on-time accurate information using various intelligent information systems. This need has enabled researchers to design multimodal information systems which mainly depend on multimodal fusion. As the data which is collected for modelling these systems is in no way free from imperfections so is the multimodal fusion which deals with such data. The motivation behind the current study is to introduce researchers working in this field of multimodal fusion with the notion of indeterminacy, uncertainty and imprecision (imperfections) present in existing approaches. This work also enables researchers to understand the field of Neutrosophic Sets and Systems by illustrating various work which are conducted using this theory to handle imperfections. The present works clearly mention how the imperfections could be handled using neutrosophy in multimodal systems. Though the work has explained well the applicability of neutrosophy in multimodal information access and retrieval systems for handling imperfections, it has not implemented the concepts in present work. The future

work in this regard would include the use of neutrosophic logic, neutrosophic algorithms and converting modalities in the neutrosophic domain so that multimodal fusion is achieved addressing the notion of imperfection. If this work is performed as explained in present paper, it would enable the design of multimodal systems more effectively so that it could be used in other areas such as medical diagnosis, financial market information, robotics, security, information fusion system, expert system and bioinformatics.

Conflicts of Interest: "The authors declare no conflict of interest."

References

1. Arnaiz, Ned Vito Quevedo and Nemis Garcia Arias Leny Cecilia Campaña Muñoz. "Neutrosophic K-means Based Method for Handling Unlabeled Data." Neutrosophic Sets and Systems 37, 1 (2020). https://digitalrepository.unm.edu/nss_journal/vol37/iss1/37
2. Bakro, Mouhammad; Reema Al-Kamha; and Qosai Kanafani. "A Neutrosophic Approach to Digital Images." Neutrosophic Sets and Systems 36, 1 (2020). https://digitalrepository.unm.edu/nss_journal/vol36/iss1/2
3. Saha, Abhijit; Said Broumi; and Florentin Smarandache. "Neutrosophic Soft Sets Applied on Incomplete Data." Neutrosophic Sets and Systems 32, 1 (2020). https://digitalrepository.unm.edu/nss_journal/vol32/iss1/17
4. Valenzuela-Chicaiza, Carmen Verónica; Olga Germania Arciniegas-Paspuel; Paola Yesenia Carrera-Cuesta; and Sary Del Rocío Álvarez-Hernández. "Neutrosophic Psychology for Emotional Intelligence Analysis in Students of the Autonomous University of Los Andes, Ecuador." Neutrosophic Sets and Systems 34, 1 (2020). https://digitalrepository.unm.edu/nss_journal/vol34/iss1/1
5. Sahin, Ridvan. "Neutrosophic Hierarchical Clustering Algoritms." Neutrosophic Sets and Systems 2, 1 (2014). https://digitalrepository.unm.edu/nss_journal/vol2/iss1/4
6. Salama, A. A.; Florentin Smarandache; and Mohamed Eisa. "Introduction to Image Processing via Neutrosophic Techniques." Neutrosophic Sets and Systems 5, 1 (2014). https://digitalrepository.unm.edu/nss_journal/vol5/iss1/9
7. Salama, A.; O. M. Khaled; and K. M. Mahfouz. "Neutrosophic Correlation and Simple Linear Regression." Neutrosophic Sets and Systems 5, 1 (2014). https://digitalrepository.unm.edu/nss_journal/vol5/iss1/2
8. Mukherjee, Anjan and Sadhan Sarkar. "A new method of measuring similarity between two neutrosophic soft sets and its application in pattern recognition problems." Neutrosophic Sets and Systems 8, 1 (2015). https://digitalrepository.unm.edu/nss_journal/vol8/iss1/11
9. Salama, A. A.; Mohamed Eisa; Hewayda ElGhawalby; and A.E. Fawzy. "Neutrosophic Features for Image Retrieval." Neutrosophic Sets and Systems 13, 1 (2016). https://digitalrepository.unm.edu/nss_journal/vol13/iss1/7
10. Bouzina, Salah. "Fuzzy Logic vs Neutrosophic Logic: Operations Logic." Neutrosophic Sets and Systems 14, 1 (2016). https://digitalrepository.unm.edu/nss_journal/vol14/iss1/6

11. Thao, Nguyen Xuan; Florentin Smarandache; and Nguyen Van Dinh. "Support-Neutrosophic Set: A New Concept in Soft Computing." Neutrosophic Sets and Systems 16, 1 (2017). https://digitalrepository.unm.edu/nss_journal/vol16/iss1/16
12. Ejaita, Okpako Abugor and Asagba P.O.. "An Improved Framework for Diagnosing Confusable Diseases Using Neutrosophic Based Neural Network." Neutrosophic Sets and Systems 16, 1 (2017). https://digitalrepository.unm.edu/nss_journal/vol16/iss1/7
13. Salama, A. A.; Florentin Smarandache; and Hewayda ElGhawalby. "Neutrosophic Approach to Grayscale Images Domain." Neutrosophic Sets and Systems 21, 1 (2018). https://digitalrepository.unm.edu/nss_journal/vol21/iss1/3
14. Ponce Ruiz, D. Vitalio; J. Carlos Albarracín Matute; E. José Jalón Arias; and L. Orlando Albarracín Zambrano. "Softcomputing in neutrosophic linguistic modeling for the treatment of uncertainty in information retrieval." Neutrosophic Sets and Systems 26, 1 (2019). https://digitalrepository.unm.edu/nss_journal/vol26/iss1/11
15. Jayaparthasarathy, G.; V. F. Little Flower; and M. Arockia Dasan. "Neutrosophic Supra Topological Applications in Data Mining Process." Neutrosophic Sets and Systems 27, 1 (2019). https://digitalrepository.unm.edu/nss_journal/vol27/iss1/8
16. Zhang, Ming, Ling Zhang, and Heng-Da Cheng. "A neutrosophic approach to image segmentation based on watershed method." Signal Processing 90, no. 5:1510-1517. (2010).
17. Sengur, Abdulkadir, and Yanhui Guo. "Color texture image segmentation based on neutrosophic set and wavelet transformation." Computer Vision and Image Understanding 115, no. 8:1134-1144. (2011).
18. Rashno, Elyas, Ahmad Akbari, and Babak Nasersharif. "A convolutional neural network model based on neutrosophy for noisy speech recognition." In 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA), pp. 87-92. IEEE, (2019).
19. Akbulut, Yaman, Abdulkadir Sengur, Yanhui Guo, and Florentin Smarandache. "NS-k-NN: Neutrosophic set-based k-nearest neighbors classifier." Symmetry 9, no. 9: 179. (2017).
20. Ju, Wen, and H. D. Cheng. "A novel neutrosophic logic svm (n-svm) and its application to image categorization." New Mathematics and Natural Computation 9, no. 01:27-42. (2013).
21. Mishra, Kritika, Ilanthenral Kandasamy, Vasantha Kandasamy WB, and Florentin Smarandache. "A Novel Framework Using Neutrosophy for Integrated Speech and Text Sentiment Analysis."Symmetry 12, no. 10: 1715. (2020).
22. Bloch, Isabelle. "Fusion of image information under imprecision and uncertainty: numerical methods." In Data Fusion and Perception, pp. 135-168. Springer, Vienna, (2001).
23. Kludas, Jana, and Stéphane Marchand-Maillet. "Effective multimodal information fusion by structure learning." In 14th International Conference on Information Fusion, pp. 1-8. IEEE, (2011).
24. Souvannavong, Fabrice, Bernard Merialdo, and Benoit Huet. "Multi-modal classifier fusion for video shot content retrieval." In Proceedings of WIAMIS. (2005).
25. Marchand-Maillet, Stéphane, Donn Morrison, Enikö Szekely, and Eric Bruno. "Interactive representations of multimodal databases." Multimodal Signal Processing.279-307. (2010).

26. Niaz, Usman, and Bernard Merialdo. "Fusion methods for multi-modal indexing of web data." In 2013 14th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), pp. 1-4. IEEE, (2013).
27. Atrey, P. K., Hossain, M. A., El Saddik, A., and Kankanhalli, M. S. Multimodal fusion for multimedia analysis: a survey. Multimedia Systems, 16:345–379. (2010).
28. Liu, Dong, Meng Wang, Linjun Yang, Xian-Sheng Hua, and Hong Jiang Zhang. "Tag quality improvement for social images." In 2009 IEEE International Conference on Multimedia and Expo, pp. 350-353. IEEE, (2009).
29. Tang, Jinhui, Shuicheng Yan, Richang Hong, Guo-Jun Qi, and Tat-Seng Chua. "Inferring semantic concepts from community-contributed images and noisy tags." In Proceedings of the 17th ACM international conference on Multimedia, pp. 223-232. (2009).
30. Romberg, Stefan, Rainer Lienhart, and Eva Hörster. "Multimodal image retrieval." International Journal of Multimedia Information Retrieval 1, no. 1: 31-44. (2012).
31. Zhang, Yu, Stephane Bres, and Liming Chen. "Semantic bag-of-words models for visual concept detection and annotation." In 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems, pp. 289-295. IEEE, (2012a).
32. Nowak, Stefanie, Karolin Nagel, and Judith Liebetrau. "The CLEF 2011 Photo Annotation and Concept-based Retrieval Tasks." In CLEF (Notebook Papers/Labs/Workshop), pp. 1-25. (2011).
33. Guillaumin, Matthieu, Jakob Verbeek, and Cordelia Schmid. "Multimodal semi-supervised learning for image classification." In 2010 IEEE Computer society conference on computer vision and pattern recognition, pp. 902-909. IEEE, (2010).
34. Nagel, Karolin, Stefanie Nowak, Uwe Kühhirt, and Kay Wolter. "The Fraunhofer IDMT at ImageCLEF 2011 Photo Annotation Task." In CLEF (Notebook Papers/Labs/Workshop). (2011).
35. Li, Wei B., Jinming Min, and Gareth JF Jones. "A text-based approach to the imageclef 2010 photo annotation task." (2010).
36. Poria, Soujanya, Erik Cambria, Amir Hussain, and Guang-Bin Huang. "Towards an intelligent framework for multimodal affective data analysis." Neural Networks 63: 104-116. (2015).
37. Poria, Soujanya, Erik Cambria, and Alexander Gelbukh. "Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis." In Proceedings of the 2015 conference on empirical methods in natural language processing, pp. 2539-2544. (2015).
38. Poria, Soujanya, Iti Chaturvedi, Erik Cambria, and Amir Hussain. "Convolutional MKL based multimodal emotion recognition and sentiment analysis." In 2016 IEEE 16th international conference on data mining (ICDM), pp. 439-448. IEEE, (2016).
39. Chen, M., Wang, S., Liang, P.P., Baltruˇsaitis, T., Zadeh, A., Morency, L.P.: Multimodal sentiment analysis with word level fusion and reinforcement learning. In: Proceedings of the 19th ACM International Conference on Multimodal Interaction, ACM: 163–171. (2017)

40. Wang, Gang, Tat-Seng Chua, Chong-Wah Ngo, and YongCheng Wang. "Automatic generation of semantic fields for annotating web images." In Coling 2010: Posters, pp. 1301-1309. (2010).
41. Bloch, Isabelle. "Fusion d'informations en traitement du signal et des images." Hermes Science Publication 2 (2003).
42. Li, Yunpeng, David J. Crandall, and Daniel P. Huttenlocher. "Landmark classification in large-scale image collections." In 2009 IEEE 12th international conference on computer vision, pp. 1957-1964. IEEE, (2009).
43. Duygulu, Pinar, Kobus Barnard, Joao FG de Freitas, and David A. Forsyth. "Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary." In European conference on computer vision, pp. 97-112. Springer, Berlin, Heidelberg, (2002).
44. Barnard, K., Duygulu, P., Forsyth, D., de Freitas, N., Blei, D. M., and Jordan, M. I. Matching words and pictures. J. Mach. Learn. Res., 3:1107–1135. (2003).
45. Blei, David M., and Michael I. Jordan. "Modeling annotated data." In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, pp. 127-134. (2003).
46. Monay, Florent, and Daniel Gatica-Perez. "On image auto-annotation with latent space models." In Proceedings of the eleventh ACM international conference on Multimedia, pp. 275-278. (2003).
47. Lienhart, Rainer, Stefan Romberg, and Eva Hörster. "Multilayer pLSA for multimodal image retrieval." In Proceedings of the ACM international conference on image and video retrieval, pp. 1-8. (2009).
48. L. Morency, R. Mihalcea, P. Doshi, Towards multimodal sentiment analysis: Harvesting opinions from the web, in: H. Bourlard, T.S. Huang, E. Vidal, D. Gatica Perez, L. Morency, N. Sebe (Eds.), Proceedings of the 13th International Conference on Multimodal Interfaces, ICMI 2011, Alicante, Spain, November 14–18, ACM, 2011, pp. 169–176. (2011). http://dx.doi.org/10.1145/2070481.2070509.
49. V. Pérez-Rosas, R. Mihalcea, L. Morency, Utterance-Level multimodal sentiment analysis, in: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, ACL 2013, 4–9 August 2013, Sofia, Bulgaria, in: Long Papers, vol. 1, The Association for Computer Linguistics, pp. 973–982. (2013), URL http://aclweb.org/anthology/P/P13/P13-1096.pdf.
50. S. Poria, I. Chaturvedi, E. Cambria, A. Hussain, Convolutional MKL based multimodal emotion recognition and sentiment analysis, in: F. Bonchi, J. Domingo- Ferrer, R.A. Baeza-Yates, Z. Zhou, X.Wu(Eds.), IEEE 16th International Conference on Data Mining, ICDM 2016, December 12–15, 2016, Barcelona, Spain, IEEE, pp. 439–448, (2016). http://dx.doi.org/10.1109/ICDM.2016.0055
51. Chandrika, Pulla, and C. V. Jawahar. "Multi modal semantic indexing for image retrieval." In Proceedings of the ACM International Conference on Image and Video Retrieval, pp. 342-349. (2010).
52. Nikolopoulos, Spiros, Stefanos Zafeiriou, Ioannis Patras, and Ioannis Kompatsiaris. "High order pLSA for indexing tagged images." Signal Processing 93, no. 8: 2212-2228. (2013).
53. Wang, Meng, Kuiyuan Yang, Xian-Sheng Hua, and Hong-Jiang Zhang. "Visual tag dictionary: interpreting tags with visual words." In Proceedings of the 1st Workshop on Web-scale Multimedia Corpus, pp. 1-8. (2009).

54. Escalante, Hugo Jair, Carlos A. Hérnadez, Luis Enrique Sucar, and Manuel Montes. "Late fusion of heterogeneous methods for multimedia image retrieval." In Proceedings of the 1st ACM international conference on Multimedia information retrieval, pp. 172-179. (2008).
55. Nowak, Stefanie, Karolin Nagel, and Judith Liebetrau. "The CLEF 2011 Photo Annotation and Concept-based Retrieval Tasks." In CLEF (Notebook Papers/Labs/Workshop), pp. 1-25. (2011).
56. Wang, Gang, Derek Hoiem, and David Forsyth. "Building text features for object image classification." In 2009 IEEE conference on computer vision and pattern recognition, pp. 1367-1374. IEEE, (2009a).
57. Mohd. Saif Wajid, Mohd Anas Wajid, The Importance of Indeterminate and Unknown Factors in Nourishing Crime: A Case Study of South Africa Using Neutrosophy, Neutrosophic Sets and Systems, vol. 41, pp. 15-29. DOI: 10.5281/zenodo.4625669. (2021).
58. Kawanabe, Motoaki, Alexander Binder, Christina Müller, and Wojciech Wojcikiewicz. "Multi-modal visual concept classification of images via Markov random walk over tags." In 2011 IEEE Workshop on Applications of Computer Vision (WACV), pp. 396-401. IEEE, (2011).
59. Elhassouny, Azeddine; Soufiane Idbrahim; and Florentin Smarandache. "Machine learning in Neutrosophic Environment: A Survey." Neutrosophic Sets and Systems 28, 1 (2019). https://digitalrepository.unm.edu/nss_journal/vol28/iss1/7
60. Gao, Shenghua, Liang-Tien Chia, and Xiangang Cheng. "Web image concept annotation with better understanding of tags and visual features." Journal of Visual Communication and Image Representation 21, no. 8: 806-814. (2010).
61. Liu, Ningning, Emmanuel Dellandréa, Liming Chen, Chao Zhu, Yu Zhang, Charles-Edmond Bichot, Stéphane Bres, and Bruno Tellez. "Multimodal recognition of visual concepts using histograms of textual concepts and selective weighted late fusion scheme." Computer Vision and Image Understanding 117, no. 5: 493-512. (2013).
62. Liu, Ningning, Emmanuel Dellandrea, Bruno Tellez, and Liming Chen. "Associating textual features with visual ones to improve affective image classification." In International Conference on Affective Computing and Intelligent Interaction, pp. 195-204. Springer, Berlin, Heidelberg, (2011).
63. Kim, D.H., Lee, M.K., Choi, D.Y., Song, B.C.: Multi-modal emotion recognition using semi-supervised learning and multiple neural networks in the wild. In: Proceedings of the 19th ACM International Conference on Multimodal Interaction, ACM, 529–535. (2017).
64. Vielzeuf, V., Pateux, S., Jurie, F.: Temporal multimodal fusion for video emotion classification in the wild. In: Proceedings of the 19th ACM International Conference on Multimodal Interaction, ACM, 569–576. (2017).
65. Neverova, N., Wolf, C., Taylor, G.W., Nebout, F.: Multi-scale deep learning for gesture detection and localization. In: Workshop at the European conference on computer vision, Springer, 474–490. (2014).
66. A. Tamrakar, S. Ali, Q. Yu, J. Liu, O. Javed, A. Divakaran, H. Cheng, and H. Sawhney. Evaluation of low-level features and their combinations for complex event detection in open source videos. In CVPR, (2012).

67. Kim, Kyung-Min, et al. "Multimodal dual attention memory for video story question answering." Proceedings of the European Conference on Computer Vision (ECCV). (2018).
68. Huang, Feiran, et al. "Image–text sentiment analysis via deep multimodal attentive fusion." Knowledge-Based Systems 167: 26-37. (2019).
69. Pandeya, Yagya Raj, and Lee Joonwhoan. "Music-video emotion analysis using late fusion of multimodal." DEStech Transactions on Computer Science and Engineering iteee (2019).
70. Makadia, Ameesh, Vladimir Pavlovic, and Sanjiv Kumar. "A new baseline for image annotation." In European conference on computer vision, pp. 316-329. Springer, Berlin, Heidelberg, (2008).
71. Torralba, Antonio, Rob Fergus, and William T. Freeman. "80 million tiny images: A large data set for nonparametric object and scene recognition." IEEE transactions on pattern analysis and machine intelligence 30, no. 11: 1958-1970. (2008).
72. Guillaumin, Matthieu, Thomas Mensink, Jakob Verbeek, and Cordelia Schmid. "Tagprop: Discriminative metric learning in nearest neighbor models for image auto-annotation." In 2009 IEEE 12th international conference on computer vision, pp. 309-316. IEEE, (2009).
73. Li, Xirong, Cees GM Snoek, and Marcel Worring. "Learning social tag relevance by neighbor voting." IEEE Transactions on Multimedia 11, no. 7: 1310-1322. (2009).
74. Jin, Yohan, Latifur Khan, Lei Wang, and Mamoun Awad. "Image annotations by combining multiple evidence & wordnet." In Proceedings of the 13th annual ACM international conference on Multimedia, pp. 706-715. (2005).
75. Weinberger, Kilian Quirin, Malcolm Slaney, and Roelof Van Zwol. "Resolving tag ambiguity." In Proceedings of the 16th ACM international conference on Multimedia, pp. 111-120. (2008).
76. Xu, Hao, Jingdong Wang, Xian-Sheng Hua, and Shipeng Li. "Tag refinement by regularized LDA." In Proceedings of the 17th ACM international conference on Multimedia, pp. 573-576. (2009).
77. Kennedy, Lyndon, Malcolm Slaney, and Kilian Weinberger. "Reliable tags using image similarity: mining specificity and expertise from large-scale multimedia databases." In Proceedings of the 1st workshop on Web-scale multimedia corpus, pp. 17-24. (2009).
78. Liu, Dong, Xian-Sheng Hua, Meng Wang, and Hong-Jiang Zhang. "Image retagging." In Proceedings of the 18th ACM international conference on Multimedia, pp. 491-500. (2010).
79. Zhu, Guangyu, Shuicheng Yan, and Yi Ma. "Image tag refinement towards low-rank, content-tag prior and error sparsity." In Proceedings of the 18th ACM international conference on Multimedia, pp. 461-470. (2010).
80. Yang, Kuiyuan, Xian-Sheng Hua, Meng Wang, and Hong-Jiang Zhang. "Tag tagging: Towards more descriptive keywords of image content." IEEE Transactions on Multimedia 13, no. 4: 662-673. (2011).
81. Wu, Lei, Rong Jin, and Anil K. Jain. "Tag completion for image retrieval." IEEE transactions on pattern analysis and machine intelligence 35, no. 3: 716-727. (2012).
82. Xioufis, E. S., Sechidis, K., Tsoumakas, G., and Vlahavas, I. P. Mlkd's participation at the clef 2011 photo annotation and concept-based retrieval tasks. In CLEF (Notebook Papers/Labs/Workshop). (2011).

83. Wajid, Mohd Anas, and Aasim Zafar. "Multimodal Information Access and Retrieval Notable Work and Milestones." In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-6. IEEE, (2019).
84. Smarandache, Florentin. "Neutrosophic set-a generalization of the intuitionistic fuzzy set." International journal of pure and applied mathematics 24, no. 3: 287. (2005).
85. Zafar, Aasim, and Mohd Anas Wajid. Neutrosophic cognitive maps for situation analysis. Infinite Study, (2020).
86. Zafar, Aasim, and Mohd Anas Wajid. A Mathematical Model to Analyze the Role of Uncertain and Indeterminate Factors in the Spread of Pandemics like COVID-19 Using Neutrosophy: A Case Study of India. Vol. 38. Infinite Study, (2020).
87. Wajid, Mohd Anas, and Aasim Zafar. "PESTEL Analysis to Identify Key Barriers to Smart Cities Development in India." Neutrosophic Sets and Systems 42: 39-48, (2021).

Received: May 2, 2021. Accepted: August 10, 2021