

ISSN - 0258-2724

DOI : 10.35741/issn.0258-2724.54.3.19

Research article

**USING ARTIFICIAL NEURAL NETWORK FOR MULTIMEDIA
INFORMATION RETRIEVAL**Maha Mahmood ^a, Wijdan Jaber AL-kubaisy ^b, Belal Al-Khateeb ^c^a College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq,
maha_mahmood@computer-college.org^b College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq,
wijdan@computer-college.org^c College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq,
belal@computer-college.org**Abstract**

Multimedia Information Retrieval (MIR) is an important field due to the great amount of information going through the Internet. Multimedia data can be considered as raw data or the features that compose it. Raw multimedia data consists of data structures with diverse characteristics such as image, audio, video, and text. The big challenge of MIR is a semantic gap, which is the difference between the human perception of a concept and how it can be represented using a machine-level language. The aim of this paper is to use different algorithms through two stages one for training and the other for testing. The first algorithm depends on the nature of the query language to retrieve the text document using two models, Vector Space Model (VSM) and Latent Semantic Index (LSI). The second algorithm is based on the extracted features using curvelet decomposition and the statistic parameters such as mean, standard deviation and energy of signals. The other algorithm is based on the discrete wavelet transform (DWT) and features of signals to retrieve audio signals, then the neural network is applied to describe the information retrieval model which retrieves the information from the multimedia. The neural network model, based on multiplayer perceptron and spreading activation network type, accepts the structure of conceptually and linguistically oriented model.

Keywords: Neural Networks, Information Retrieval, Natural Language Processing, Information Retrieval System, Multimedia Information Retrieval System.

摘要：多媒體信息檢索（MIR）是一個重要的領域，因為有大量的信息通過互聯網。多媒體數據可以被視為原始數據或組成它的功能。原始多媒體數據由具有不同特徵的數據結構組成，例如圖像、音頻、視頻和文本。MIR的一大挑戰是語義差距，即人類對概念的感知與使用機器級語言表達概念之間的差異。本文的目的是通過兩個階段使用不同的算法，一個用於訓練，另一個用於測試。第一種算法依賴於查詢語言的性質來使用兩種模型檢索文本文檔，即向量空間模型（VSM）和潛在語義索引（LSI）。第二種算法基於使用曲線分解的提取特徵和諸如平均值、標準偏差和信號能量之類的統計參數。另一種算法基於離散小波變換（DWT）和信號特徵來檢索音頻信號，然後應用神經網絡描述從多媒體中檢索信息的信息檢索模型。基於多玩

家感知器和擴散激活網絡類型的神經網絡模型接受概念和語言導向模型的結構。

关键词: 神經網絡, 信息檢索, 自然語言處理, 信息檢索系統, 多媒體信息檢索系統。

I. INTRODUCTION

There is huge information on the web at which users can utilize for creating and storing images. Which has posed the need for ways of managing and searching those images. This is why, finding sufficient multimedia retrieval approaches have become an important field of interest for scholars. Multimedia retrieval approach is a system for the search and retrieval of multimedia objects (texts, images, sounds and videos) from a large data-base of digital libraries [1]. The field of Information Retrieval (IR) is not new as early IR systems used simple tools of word matching for small texts. In the present scenario due to the large volume of availability of information sources and very different forms, there is a need of more efficient retrieval techniques which can retrieve only the relevant part of the information [2]. People need information frequently and this information must be stored in physical storage devices of computers. There are many algorithms used to retrieve information. MIR systems deal with different types of media (text, image, audio and video). Although of media is different but there is a common factor between them, the common factor is concerned with text. To illustrate this point, the user can search any information by introducing some keywords inside the search field such as in Google search engine [4].

II. MULTIMEDIA INFORMATION RETRIEVAL

MIR is a study area of computer science which has the aim of the extraction of semantic information from multimedia data sources, which include media which is directly perceivable like video, audio, and images, in addition to indirectly perceivable sources like text, bio signals, semantic descriptions, and non-perceivable sources like bio-information and stock prices. MIR system should be able to represent and store multimedia objects in a way that ensures their fast retrieval [3]. The system should be; therefore, able to deal with different kinds of media and with semi-structured data, i.e., data that has a structure which may not be matching or might only partially matching, the structure prescribed by the data scheme. In order to represent semi-structured data, the system must typically extract some features from the multimedia objects. The news is a particularly interesting for MIR since

many possible sources of indexing information are available: speech, audio, video and text obtained from television, radio and newspapers [7]. Due to the time constraints of continuous daily news coverage and the volume of data, automatic news library creation methods are required [6]. Figure 1 shows the general information retrieval system.

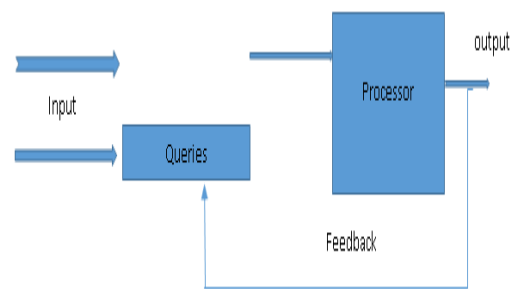


Figure 1: General Information Retrieval System.

The IR community has concentrated on the textual retrieval of mainly English documents. Content-based retrieval of textual documents is not limited to the plain text but is currently extended to specialized kinds of languages, such as formulas, tables, hand writing and special structures like hypertext [5]. Another line of research in text retrieval is the complementary application of Natural Language Processing (NLP) techniques. The use of (language specific) linguistic knowledge may improve the effectiveness of text retrieval, i.e., the quality of tasks such as summarization, extraction, filtering and categorization. NLP techniques in a combination with the use of structure of textual documents can even further improve the retrieval [8].

The interest in audio retrieval has been relatively low compared to the visual media types. Research in audio retrieval and extracting audio features is increasing resulting in different audio (speech and music) retrieval systems. To search for a certain sound or type of sound (such as speech of some speaker or music) can be a daunting task. Words are inadequate to convey the essence of sounds and there is no standard for sound classification. No two listeners will produce the same description for every sound. Humans tend to describe sounds by similar sounds, e.g. a buzzing sound for bees. The

feature based on matching applies where a sound is a single gestalt (short single sounds or longer recordings with a uniform texture, e.g. rain on roof) [10].

The convergence between the image processing and data-base communities produced many systems of content-based image retrieval. The semantic interpretation of visual information is of a considerably higher level of complexity compared to the structured text. In texts, each of the words has a limited number of meanings. Through paragraph or sentence analysis, it is possible to determine the precise meaning. Visual items that have the same semantic notion exhibit some great appearance varieties, for instance images of dogs in kids' books. The search of systems of image retrieval is approximate, which means that those systems utilize automatically obtained characteristics of an image for determining the visual similarity. The visual properties are derived from a process of computation that has been executed on the image object. Simple properties (like shape, color histograms, and texture) are calculated on the basis of pixel characteristics, such as position and color [9].

III. NEURAL NETWORKS IN INFORMATION RETRIEVAL

Neural networks (NNs) have been applied to IR in a variety of ways. The 3 main methods are:

- Transformation network has been proposed for the enhancement of queries. It includes a back-propagation NN with one or several hidden layers in which inputs and outputs are schemes of representation [15].
- The model of COSIMIR uses back-propagation NN for the match between the document representation and the query. Each of the document and the query play the role of an input to the network that computes their similarity in the output layer. The similarity represents of measurement for document relevance to the query. The training data must be gathered from a large amount of relevance judgements from users. Therefore, COSIMIR implements a function of cognitive similarity [12].
- Spreading models of activation are basically Hopfield networks that are customarily utilized with nodes for query terms and document nodes for the retrieval of the most activated documents. The links are weighted based on the term-matrix of the document which is determined using an algorithm of common indexing.

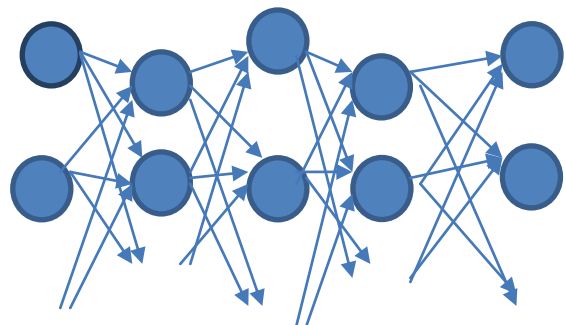
Spreading models of activation are applied for large amount of real world data and reached sufficient results compared with the results of the statistical models that dominate in IR development and research [14].

All of the models utilize the term "vector" where a document is characterized by some terms within a great deal of all terms that occur in the collection of documents. The models come across some issues that derive from the size of those sparsely coded vectors. Particularly, results of full text retrieval in large vectors that possibly contain every natural language word. Even manual indexing with the use of a controlled thesaurus typically produces big vectors. For instance, the Social Science Information Centre's thesaurus in Bonn has 22,000 entries. A NN that has similarly large number of nodes needs significant resources [11]. This is why, reduction of dimensions is an approach that has great perspectives for an efficient use of NNs in IR.

IV. NEURAL SYSTEM DESCRIPTION

The NN models that have been considered in the present paper are feed-forward networks. Input features are learned or extracted by neural networks with the use of numerous, stacked fully connected layers. Every one of those layers applies a linear transform to the vector output of the previous layer (conducting an affine transform). This is why, every one of the layers is associated with a matrix of parameters that will be estimated throughout the learning. Which precedes an element-wise application of a nonlinear activation function. Concerning the IR, the result of the whole net is usually some projected scores or a vector representation of the input [22]. Throughout the process of training, a loss function is produced via contrasting the prediction with the ground truth available for the training data, in which training alters the parameters of the network for minimizing loss [21]. Which is usually carried out through the classic algorithm of backpropagation, figure 2 shows the neural network IR model.

Query Documents Keywords



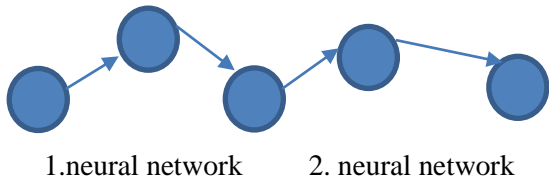


Figure 2: Information Retrieval Model of a Neural Network.

Retrieved multimedia are encoded on input of the first NN [20]. After that, it is determined by this NN whether or not the particular at multimedia in the document is the key-word. If it is, then the occurrence will be added to model matrix of the vector space [5]. Which results in normalized vector space model matrix. This matrix weights are matched with the second NN weights that has some key-words as an input and the base of multimedia as an output. Combining those 2 NNs, the system of information retrieval is developed in the manner which has been illustrated in figure. 3. After that, the input interface will be created, and that enables the user to enter a query and after that, the system of information retrieval enables finding the related multimedia [21]. Finally, the output interface will be created,

which is responsible for sorting the relevant multimedia and sending it to the user as results. In the case where there is a necessity for giving as a query that includes 2 words or more, there is a necessity of adding more groups of input neurons to the first NN where every one of the groups represents multimedia. After that, every type separately generated and there are created coherent key-words to every one of the queries. This structure offers the ability of giving more key-words as an input in addition to the ability of giving the multimedia connections as input to the NN.

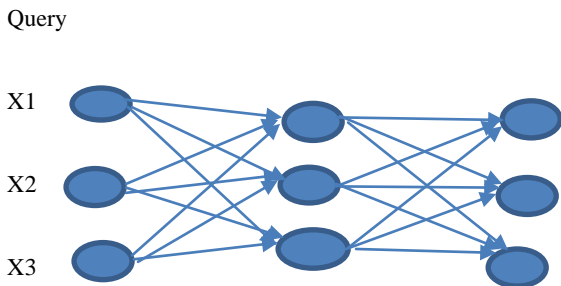


Figure 3: Neural Network for Determination of Multimedia Information.

The general steps of the proposed system for multimedia database are shown in figure 4 for three multimedia; text, image and audio. The

system consists of algorithms with two stages one for training called off-line stage and the other for testing called on-line stage. The first algorithm depends on nature of the query language to the retrieve text document using two models VSM and LSI. The second algorithm is based on the extracted features by Curvelet decomposition and the statistic parameters like the standard deviation, mean and Energy of signals. This algorithm is called Content Based Image Retrieval (CBIR). The third algorithm is based on the Discreet Wavelet Transform (DWT) and features of signal to retrieve audio.

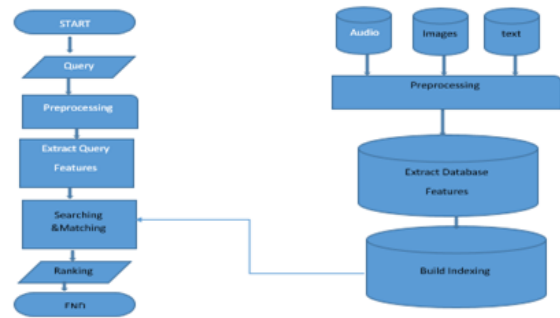


Figure 4: Flowchart of the Proposed System.

V. VECTOR SPACE MODEL

Vector Space Model (VSM) is a standard IR approach, where documents are characterized via the words they contain. It has been introduced by G. Salton in the early 60's for avoiding some issues of IR [5]. In the model of vector space, every one of the documents is represented by an N-dimensional term weight vector, every one of the elements of the vector which is the weight of every N terms in that document. In the case where a document group has M documents, then the group is represented in a form a matrix A of $M \times N$ dimension. Throughout the process of retrieval, the query is represented in an N-dimensional term weight vector as well. The similarity between the query and every stored document is computed as either the dot product or the cosine coefficient between the document vector and query vector and this is usually represented as in a table 1.

Table 1: Vector Space Model.

The magnitude of the vector for each document by using the Pythagorean Theorem but in this case there are more than two dimensions, so the formula would be in the equations below:

$$\|v_1\| = \sqrt{(a_{11})^2 + (a_{12})^2 + \dots + (a_{1n})^2} \quad (1)$$

$$\|v_2\| = \sqrt{(a_{21})^2 + (a_{22})^2 + \dots + (a_{2n})^2} \quad (2)$$

$$\|v_m\| = \sqrt{(a_{m1})^2 + (a_{m2})^2 + \dots + (a_{mn})^2} \quad (3)$$

Where:

V is the length of document vector

a (i, j) is the weight of term j in document i.

The inner product of query vector with every single document vector for every term, will compute the score via calculating the product of the weight for the query term by the weight of the document term and after that compute the summation of the term scores. Those calculations are referred to as the cosine correlation measurement which may directly be computed with the use of [13]:

$$\text{COSIN}(DOC_i, QUERY) = \frac{\sum_{k=1}^t (TERM_{ik} * QTERM_{jk})}{\sqrt{\sum_{k=1}^t (TERM_{ik})^2 * \sum_{k=1}^t (QTERM_{jk})^2}} \quad (4)$$

The cosine correlation measures the cosine of the angle between documents and query when those are viewed as vectors in the multi-dimensional term space of dimension t.

Algorithm1: Steps of Vector Space Model.

Input Text Database (DB).

Output Documents Features (DF).

Start

 Compute length DB

 I=1

 While (I<length (DB), I++)

 Compute length Doc (I)

 J=1

 While (J<length (Doc (I), J++)

 Read Term (J)

 If (Term (J) not Stop Word)

 Stem Term (J) using Porter stemmer

 Compute Term Frequency (TF)

(This is the frequency of a term inside a document. The frequency is usually normalized within the particular document)

Term Space	Term Count		
	Doc ₁	Doc ₂	...Doc _n
T ₁	a ₁₁	a ₁₂	a _{1n}
T ₂	a ₂₁	a ₂₂	a _{2n}
...
T _m	a _{m1}	a _{m2}	...a _{mn}

Else

 Remove Stop Word

 End while

End while

 Compute Documents Frequency (DF)

 Compute Inverse Document Frequency (IDF)

(IDF=log (total documents in database/documents containing the term))

 End.

VI. LATENT SEMANTIC INDEXING

Latent Semantic Analysis (LSA) or LSI is a theory and approach for the extraction and representation of the contextual-utilization meaning of words via statistical calculations that are applied to a considerable amount of text [18]. The initial phase of LSI is representing the text in a form of a matrix where every row represents a distinct word and every column represents a passage of the text or some other context. After that, the entries of the cells are subjected to every cell frequency that is weighted via a function which represents each of the word's importance in the certain passage and the level to which the type of the word carries data. Then, singular value decomposition (SVD) is applied to the matrix by the LSA. In SVD a rectangular matrix is divided to the result of the multiplication of 3 other matrices. One of the component matrices describes the entities of the original row as vectors of derived values of the orthogonal factor, another one similarly represents the entities of the original column, and the third one is a diagonal matrix which contains scaling values in a way that when those 3 components are matrix-multiplied, the initial matrix is rebuilt [17]. The theorem of SVD for any real matrix A with dimensions (m × n) can be expressed as shown in figure 5 [18]:

$$A=U*S*V^t \quad (5)$$

Where: U_{m*r} is a column-orthonormal (U*U=I), r is the rank of the A, S_{r*r} is a diagonal matrix, I is the identity matrix, V_{r*n} is a column-orthonormal matrix, U is column-orthonormal.

When S elements are arranged in a descending order, the above decomposition is unique. In text document retrieval context, the rank r of A equals the number of concepts. U is considered as the similarity matrix of document-to-concept, whereas V represents the term-to-concept similarity matrix. For instance, $U_{2, 3} = 0.6$ indicates the fact that concept 3 has weight 0.6 in document 2, and $V_{1, 2} = 0.4$ indicates the fact that the similarity between concept 2 and term 1 is 0.4. The LSI strong points are strong formal framework completely automatic without the need for stemming, can be utilized for multi-lingual search and conceptual IR recall improvement. Weaknesses points of LSI are calculating LSI which is expensive, continuous normal-distribution- based approaches not so suitable for count data and usually enhancing accuracy is of a higher importance: require query and word sense disambiguation [18].

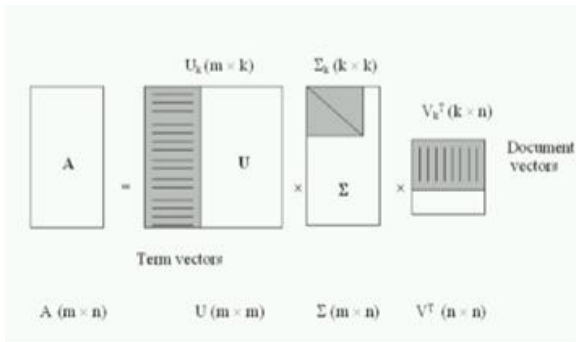


Figure 5: Latent Semantic Indexing.

VII. DISCRETE WAVELET TRANSFORM

The wavelet transform is an important computational tool for many applications of signal and image processing. For instance, it is useful for compressing digital media files (images or audio). In 1D signal, for every one of the levels the signal is decomposed to 2 frequency sub-bands High H and Low L where L represents the low frequency and H represents the high frequency. As illustrated in figure. 6. The calculation of the wavelet transform for a two-dimensional signal is involved with sub-sampling and recursive filtering. At every level, the signal is divided to 4 frequency sub-bands, which are: HL , HH , LL , and LH [19].



Figure 6: 1D Discrete Wavelet Transform with Different Levels.

Feature extraction requires that voice should be clean from all noise therefore the Discrete Wavelet Transform used here so that feature extracted per frame of a voice command should have a value within the range of the class the voice command belongs. Audio features computed by the following audio features which are Zero Crossing Rate(ZCR), Spectral Flux(SF), Spectral Roll off (SR), Spectral Centroid (SC), Energy(E), and Energy Entropy(EE), then Standard Deviation (SD) for each these features to create feature vector.

VIII. RESULTS

This paper implements MIR system using three algorithms through 2 phases, one of them for training and the other one for testing. The first algorithm depends on the nature of the query language to retrieve the text document using two models, VSM and LSI. The second algorithm is based on the extracted features using curvelet decomposition and the statistic parameters like mean, standard deviation and Energy of signals. This algorithm is called CBIR. The third algorithm is based on DWT and features of signal to retrieve audio signals. Neural networks are tested by selecting 1000 images. The data-set of tiny images on which all of the experiments are based on, has been gathered by colleagues at NYU and MIT over 6 months; it has been thoroughly described in [14], which are categorized to 10 categories; every one of which has 100 images as depicted figure 7.

category	random people	beach	building	buses	dinosaurs	elephants	flowers	horses	mountains and glaciers	food
number of image	0-99	100-199	200-299	300-399	400-499	500-599	600-699	700-799	800-899	900-999

Figure 7: Image Categories.

Feature extraction is a fundamental component in Multimedia Information Retrieval systems. These occur in off-line pre-processing phase and when users request a system with an image query. The purpose of feature extraction is determining a group of features for describing

every image. In that step, the features of images data are obtained from images. These features are used to compare to query feature. After matching the results are ranking based on similarity level. Efficiency evaluations are carried out via applying 2 metrics which are referred to as the precision and the recall. The training phase results have shown that the neural network based curvlet has greater elapsed time than the neural networks based on DWT or Histogram as shown in figure 8.

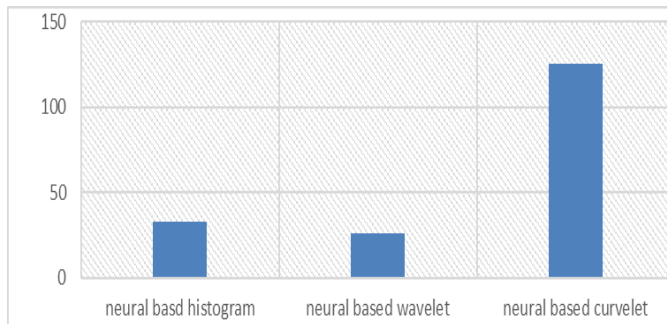


Figure 8: Proposed System Training Time.

In the online stage multimedia have been arbitrarily chosen from every set in images, text and audio data-base then the multimedia were tested by the CBIR system based on Curvlet, Histogram, and Wavelet. The training stage outcomes show that the neural networks system when based on Curvlet, but the search duration is short, in comparison with CBIR when based on Wavelet and Histogram. The curves of recall and precision are shown in the figure 9.

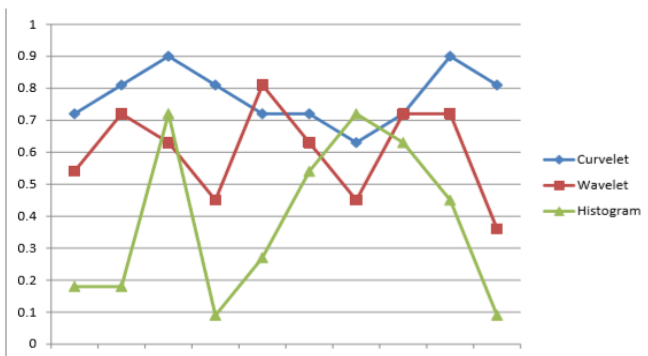


Figure 9: The Recall and Precision curves based on Histogram, Wavelet and Curvlet.

IX. CONCLUSIONS

This work presents a neural network for the retrieval of multimedia information. The classification information is so important to reduce the search time. This point appears when

the user search in clustering database therefore the proposed system gives good results in a with low retrieving time. A neural network learning is modelled and trained on a collection of different multimedia. The learned features have been utilized for presenting a highly sufficient system of multimedia retrieval which operates for a large set of multi-modal data-set.

X. FUTURE WORK

The latest technology that addresses, there will be an investigation on more sophisticated approaches of deep learning and assess more other different data-sets for more in-depth experiential researches in order to provide more knowledge for reducing the semantic gap of the retrieval of multimedia information in the long term. This field of study utilizes enhanced statistical analysis and ML approaches for uncovering patterns and correlations in information multimedia and conventional databases which classical approaches might not be able to discover.

References

- [1] Chowdhary, K.R. and Bansal V.S. (2001). Current Trends in Information Retrieval. Prof. Emeritus, at 4th International Conference of Asian Digital Libraries, at University of Mysore, (Held at Hotel Le Meridian, Bangalore), 10-12 Dec.
- [2] Carkacioglu A. (2003) Texture Descriptions for Content Based Image Retrieval. PhD. Dissertation: The Graduate School of Natural and Applied Sciences of the Middle East Technical University.
- [3] Kuper, J. et al. (2003) Intelligent Multimedia Indexing and Retrieval through Multi-Source Information Extraction and Merging. International Joint Conference on Artificial Intelligence, IJCAI.
- [4] Zhang, D. et al, (2004) Content Based Image Retrieval Using Gabor Texture Features. Proc.1st IEEE Pacific Ri.
- [5] Ibrahiem, M. and Atwan, J. (2005) Designing and Building an Automatic Information Retrieval System for Handling the Arabic Data. American Journal of Applied Sciences, Volume 2, Issue 11, pp. 1520-1525.
- [6] Negoia, C. and Monica, V. (2006) Querying and Information Retrieval in Multimedia Databases. Buletinul Universității Petrol – Gaze din Ploiești Vol. LVIII No. 2/ 73 - 78 Serial Mathematica - Informatics – Fizică.

- [7] Magallanes, J., Overell, S. and Rüger, S. (2007) A Semantic Vector Space for Query by Image Example.
- [8] Chung, Y.Y. et al. (2007) Design of a Content Based Multimedia Retrieval System. WSEAS Transactions on Computers. Issue3, Vol. 6, March.
- [9] Wong, Y., Hoi, C. and Michael, R. (2007) An Empirical Study on LargeScale Content-Based Image Retrieval. in Proceedings of the 2007 IEEE International Conference on Multimedia and Expo (ICME'2007); Beijing.
- [10] Magalhães, J. (2008) Statistical Models for Semantic Multimedia Information Retrieval. PhD. Dissertation: Department of Computing, Imperial College of Science, Technology and Medicine, University of London.
- [11] Mostafa, S.A., Mustapha, A., Mohammed, M.A., Hamed, R.I., Arunkumar, N., Ghani, M.K.A., Jaber, M.M. and Khaleefah, S.H., 2019. Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease. Cognitive Systems Research, 54, pp.90-99.
- [12] Siewt, T. C. (2008) Feature Selection for Content Based Image Retrieval Using Statistical Discriminate Analysis. M.Sc. Thesis: University of Technology Malaysia, Faculty of Computer Science and Information System.
- [13] Mohammed, M.A., Gunasekaran, S.S., Mostafa, S.A., Mustafa, A. and Ghani, M.K.A., 2018, August. Implementing an Agent-based Multi-Natural Language Anti-Spam Model. In 2018 International Symposium on Agent, Multi-Agent Systems and Robotics (ISAMSR) (pp. 1-5). IEEE.
- [14] Torralba, A., Fergus, R. and Freeman, W. T. (2008) 80 million tiny images: a large dataset for nonparametric object and scene recognition.
- [15] Büttcher, S. Clarke, L. and Cormack, G. (2010) Information Retrieval. Implementing and Evaluating Search Engines"; Ch14; MIT Press.
- [16] Benjamin Durakovic, Muris Torlak, "Simulation and experimental validation of phase change material and water used as heat storage medium in window applications", J. of Mater. and Environ. Sci., Vol. 8, No. 5, (2017), ISSN: 2028-2508, pp 1837-1846.
- [17] Ogudo, K.A.; Muwawa Jean Nestor, D.; Ibrahim Khalaf, O.; Daei Kasmaei, H. A Device Performance and Data Analytics Concept for Smartphones' IoT Services and Machine-Type Communication in Cellular Networks. *Symmetry* **2019**, *11*, 593
- [18] Prasannakumari, V. (2010) Contextual Information Retrieval for Multimedia Database with Learning by Feedback Using Vector Space Model. Asian journal of information Management 4(1); 12-18.
- [19] Fadhil, A. F. (2010) Image Steganography Based Curvelet Transform. Al-Rafidain Engineering Journal Volume 18 No.5 October.
- [20] Al-Khateeb, B. and Mahmood, M. (2013) A Framework for an Automatic Generation of Neural Networks. IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 3, No 2.
- [21] Mahmood, M. and Al-Khateeb, B. (2017) Towards an Automatic Generation of Neural Networks", Journal of Theoretical and Applied Information Technology, Vol.95. Issue 23.
- [22] Mahmood, M. and Al-Khateeb, B. and Alwash W.M. (2018) Review of Neural Networks Contribution in Network Security. Journal of Advanced Research in Dynamical and Control Systems (SCOPUS Q4), Volume 10, Issue 13, 2139-2145, December.

參考文:

- [1] Chowdhary, K.R. 和 Bansal V.S. (2001 年)。信息檢索的當前趨勢。名譽教授，邁索爾大學第四屆亞洲數字圖書館國際會議 (12 月 10 日至 12 日在班加羅爾 Le Meridian 酒店舉行)。
- [2] Carkacioglu A. (2003) 基於內容的圖像檢索的紋理描述。博士學位。學位論文：中東技術大學自然與應用科學研究生院。
- [3] Kuper, J. 等。 (2003) 通過多源信息提取和合併的智能多媒體索引和檢索。IJCAI 國際人工智能加入會議。
- [4] Zhang, D. 等, (2004) 基於內容的圖像檢索使用 Gabor 紋理特徵。 Proc.1st IEEE Pacific Ri.
- [5] Ibrahiem, M. 和 Atwan, J. (2005) 設計和建立一個處理阿拉伯數據的自動信息檢索系統。 American Journal of Applied Sciences, 第 2 卷, 第 11 期, 第 1520-1525 頁。

- [6] Negoită, C. 和 Monica, V. (2006) 多媒體數據庫中的查詢和信息檢索。Buletinul Universității Petrol - Gazedin Ploiești Vol. LVIII No. 2/73 - 78 Serial Mathematica - Informatics - Fizică.
- [7] Magallanes, J., Overell, S. 和 Rüger, S. (2007) 用於查詢圖像的語義向量空間示例。
- [8] Chung, Y.Y. 等。 (2007) 基於內容的多媒體檢索系統的設計。WSEAS 計算機上的事務。第 3 期, Vol. 3 月 6 日。
- [9] Wong, Y., Hoi, C. 和 Michael, R. (2007) “基於大規模內容的圖像檢索的實證研究”。在 2007 年 IEEE 多媒體和博覽會國際會議論文集 (ICME'2007); 北京。
- [10] Magalhães, J. (2008) 語義多媒體信息檢索的統計模型。博士學位。學位論文：倫敦大學帝國理工學院計算系。
- [11] Mostafa, S.A., Mustapha, A., Mohammed, M.A., Hamed, R.I., Arunkumar, N., Ghani, M.K.A., Jaber, M.M. 和 Khaleefah, S.H., 2019。檢查多功能評估和分類方法，以改善帕金森病的診斷。Cognitive Systems Research, 54, pp.90-99。
- [12] Siewt, T. C. (2008) 使用統計判別分析進行基於內容的圖像檢索的特徵選擇。碩士論文：馬來西亞科技大學計算機科學與信息系統學院。
- [13] Mohammed, M.A., Gunasekaran, S.S., Mostafa, S.A., Mustafa, A. 和 Ghani, M.K.A., 2018, August。實現基於代理的多自然語言反垃圾郵件模型。2018 年代理，多代理系統和機器人國際研討會 (ISAMSR) (第 1-5 頁)。IEEE。
- [14] Torralba, A., Fergus, R. 和 Freeman, W. T. (2008) 8000 萬個微小圖像：用於非參數對象和場景識別的大型數據集。
- [15] Büttcher, S. Clarke, L. 和 Cormack, G. (2010) Information Retrieval。實施和評估搜索引擎“；第 14 章；麻省理工學院出版社。
- [16] Benjamin Durakovic, Muris Torlak, “用於窗戶應用中作為儲熱介質的相變材料 and 水的模擬和實驗驗證”，J.of Mater. 和環境。Sci., Vol. 8, No. 5, (2017), ISSN : 2028-2508, pp 1837-1846。
- [17] Ogudo, K.A.; Muwawa Jean Nestor, D.; Ibrahim Khalaf, O.; Daei Kasmaei, H. 智能手機在蜂窩網絡中的物聯網服務和機器類通信的設備性能和數據分析概念。Symmetry 2019,11,593
- [18] Prasannakumari, V. (2010) 使用向量空間模型通過反饋學習的多媒體數據庫的上下文信息檢索。亞洲信息管理雜誌 4 (1) ;12-18。
- [19] Fadhil, A. F. (2010) Image Steganography Based Curvelet Transform. Al-Rafidain Engineering Journal 第 18 卷第 10 期。
- [20] Al-Khateeb, B. 和 Mahmood, M. (2013) “自動生成神經網絡的框架”。IJCSI 國際計算機科學期刊, Vol. 10, 第 3 期, 第 2 期。
- [21] Mahmood, M. 和 Al-Khateeb, B. (2017) “自動生成神經網絡”，“理論與應用信息技術期刊”，第 95 卷。第 23 期。
- [22] Mahmood, M. 和 Al-Khateeb, B. 和 Alwash W.M. (2018) 神經網絡在網絡安全中的貢獻回顧。動力與控制系統高級研究期刊 (SCOPUS Q4), 第 10 卷, 第 13 期, 2139-2145, 12 月。