The main purpose of this research is to design a comprehensive algorithm which aims to aid the medical practitioners mainly the radiographers, radiologists and neurologists in primary brain tumors diagnosis. Detection of primary brain tumors is inspired by the necessity of high accuracy as it deals with human life. Presently, various imaging modalities techniques have incarnated as a tool for the doctors and radiologists to help them in diagnosis and treatment domain. While these are highly accurate and fast, they still require experienced and competent medical practitioners for the proper interpretation. Thus, the involvement of information technology is highly demanded in introducing reliable, simple and accurate computer systems. This study presents an algorithm for nosologic segmentation of primary brain tumors on Magnetic Resonance Imaging (MRI) brain images. The MRI technique has been chosen as the digital imaging modality since it provides clearer image for the tissue area as compared to the other techniques that focusing more on bone study such as Computed Tomography (CT) Scan and X-ray. The purpose of segmentation is to highlight the tumor areas, whereas classification is used to identify the type of the primary brain tumors. For this purpose, an algorithm which hybridized the Intensity Based Analysis (IBA), Grey Level Co-occurrence Matrices (GLCM), Adaptive Network-based Fuzzy Inference System (ANFIS) and Particle Swarm Optimization (PSO) Clustering Algorithm (CAPSOCA) is proposed. The combination of several computer vision techniques was presented which aims to deliver reproducible nosologic segmentation of primary brain tumors which are gliomas and meningiomas. A Rule Based Expert System (RBES) is used to preliminarily classify the various types of primary brain tumors. It is designed to incorporate with the CAPSOCA algorithm which intended to strengthen the classification outcomes. The performance of the proposed CAPSOCA algorithm is compared with the PSO and ANFIS algorithms separately. The performance of the algorithms is quantified by two measurements which are segmentation accuracy and classification accuracy. The segmentation accuracy is evaluated using two approaches which are tumor template matching and comparison with ground truth. On the other hand, the classification accuracy is quantified using truth table by comparing the classification outcomes with histopathology diagnosis collected from the patient’s record. The study revealed that very high pixel detection is noticeable in PSO algorithm which leads to over-segmentation in the tumor areas, and indirectly affected the segmentation accuracy. On the other hand, the ANFIS is found to have limited pixel detection in nosologic segmentation of primary brain tumors. The CAPSOCA was proven to be the best algorithm for nosologic segmentation of primary brain tumors from the MRI images data. It appeared to return the highest percentages of accuracy as compared to the PSO and ANFIS for both segmentation and classification at 78.79% of accuracy for gliomas segmentation, 76.68% of accuracy for meningiomas segmentation and 76.30% of accuracy for classification accuracy.

Huge reliance on computer usage in everyday life, leads to a continuous increase of large data applications in textual forms. The data are reposited to a secondary storage for future usage. Therefore, a relational database (RDB) is most commonly used as a backbone in most application software for organising such data into structured form. The RDB has robust and powerful structures for managing, organising, and retrieving the data. However, the database structure can still contain large amounts of unstructured textual data. Dealing with unstructured textual data leads to three basic issues; users encounter difficulties to find useful information, inaccurate information retrieval and insufficient performance of query processing. Attempts have been made to resolve all of these issues by using several methods such as: full text searching, text indexing, a database schema management, database data model, and query-based techniques. However, the front-end approach, in the form of software applications, are still needed to organise the unstructured textual information in the RDB. This study proposes a Textual Virtual Schema Model (TVSM) as the back-end approach to reorganising textual data inside relational databases, while performing automatic semantic linking and clustering assignments. Upon storing any new unstructured textual data into a database, all words are extracted to uncover the underlying meaning of such data. Their name entities and top most frequent terms are selected for the factors used in a cluster assignment. The model is tested and evaluated by embedding it in a component-based package of a relational database’s internal structure. Three experiments have been conducted on textual Reuters corpus, Classic and WAP dataset. The clustering results have been validated using the F-measure, Entropy and Purity methods of measurement and compared with two common methods, which are information extraction and textual document clustering, for example, K-means, Frequent Item-Set, Hierarchical Clustering Algorithms and Oracle Text. The results show that there are linkages between structured textual data and unstructured information, high performance of query processing and time improvement in document clustering with accurate clusters. Thus, the proposed technique can increase retrieval performance and produce high accuracy textual data clusters. This model envisages a beneficial and useful approach for various domains that involve big textual data such as document clustering, topic detecting and tracking, information integration, personal data management and information retrieval.

- This research work published in eight international proceeding indexed by ISI and Scopus and one international journal.
- This research work has patent pending under serial number PI2013002636 from MYiPO Malaysia.