

ABSTRACT

Pattern Recognition (PR) is the science of making inferences from perceptual data, using tools from statistics, probability, computational geometry, machine learning, signal processing, and algorithm design. Pattern recognition aims to classify patterns based on either a priori knowledge or on statistical information extracted from the patterns. A pattern could be a fingerprint image, a handwritten cursive word, a human face, or a speech signal.

Intelligent systems (IS) provide a methodological approach to solve fairly complex problems to obtain reliable results, emulating some aspects of intelligence exhibited by nature. IS comprises of Artificial Intelligence (AI) and Computational Intelligence (CI). AI is restricted to symbolic representations and manipulations in a top-down way as well as a sound logical inference mechanism. CI is defined as a methodology involving computing that exhibits an ability to learn and / or deal with new situations, comprises of soft computing whose principal constituents have extended from fuzzy logic, neurocomputing, and evolutionary computing to probabilistic reasoning and rough sets. Type 2 Fuzzy Set (FS) is an extension of FS and is highly suitable for approximate reasoning. The Interval-valued fuzzy sets (IVFS), the Intuitionistic FS (IFS), vague set and grey set are all extensions of Type 2 FS. Each extension is different from the other, and each of them is originated to solve a different problem.

The main objective of this thesis is to study and analyze the applicability of intelligent systems to certain phases in pattern recognition. The phases that are studied are instance selection phase, preprocessing phase, feature subset selection phase, logical abductive inference and logical deductive inference in the classification phase. The following sections outline the chapters of the thesis.

Offline character recognition is a PR problem that refers to the classification problem of identifying handwritten characters. The difficulty of identifying the correct character increases as the style and method of writing varies from person to person. The performance of a character recognition system depends on the selection of features, instances and the classifier. A new algorithm for instance selection is proposed to identify the nature of the instances. The key parameters that characterize the instances are obtained using reduct soft sets and an interval estimate for categorizing the instances through statistical approach is defined. The proposed algorithm is tested on hand-printed isolated character database consisting of characters of an Indian language (Tamil). The unlabeled instances are used in the testing phase in a multi-layered neural network to validate the instance selection efficiency using reduct soft sets.

Signature verification system is a widely used biometric system in places like passport, payroll, driving license, credit cards for identity verification. Signatures are a behavioral biometric that change over a period of time and are influenced by physical and emotional conditions of the signatories. Offline verification lacks any form of dynamic information so

that it has to rely on the features that are extracted from the static signature image and hence much more difficult to validate. The methodology proposed uses fuzzy sets / intuitionistic fuzzy sets based contrast enhancement in preprocessing, gradient direction histogram combined with one, two and three levels of equimass in feature extraction and a similarity score based on IVFS combined with adaptive thresholding in verification.

The difficulty of signature verification lies in the variability of signing. A second methodology is proposed which uses fuzzy rule based contrast enhancement in preprocessing, gradient direction histogram combined with one level of equimass in feature extraction, fuzzy rough approach in feature selection and a simplified fuzzy ARTMAP in verification.

Corruption is defined as the misuse of public office for private gain which remains a serious problem in both developing and developed countries. Corruption perceptions index (CPI) is a poll-of-polls index computed from multiple independent sources of data, consolidated through surveys for all countries. In such a case, identifying factors that contribute to corruption is a highly challenging task. Abductive inference combined with previous literature study is used to identify an initial list of features, a subset is selected from them based on information gain ranking and correlation analysis and verified using fuzzy C-means clustering.

Bone mineral density (BMD) is measured by dual-energy X-ray absorptiometry equipment (DXA). The DXA equipments from different manufacturers might not give identical results, because of differences in calibration and bone edge detection algorithms. Also, BMD values vary

depending on race, age, gender and other health conditions. An affordable and reliable system is proposed depending on single X-ray absorptiometry using detective inference and supervised learning method. The methodology proposed uses application of image filters and H-domes image slicing in preprocessing, computing texture and fractal dimension in feature extraction, fuzzy rough approach in feature subset selection, and a supervised learning algorithm to classify patients into Normal (Normal mineral density), Medium (osteopenia condition) and Low (osteoporosis condition).

Newer solutions to the already existing applications in PR domain are proposed in this research by way of exploring the phases of instance selection using reduct soft sets, image enhancement in preprocessing using fuzzy, intuitionistic fuzzy and fuzzy rule based methods, feature selection using fuzzy rough method, classification using MLP, similarity measure based IVFS, SFAM and Adaboost methods and clustering using FCM. The experiments conducted show the applicability of the proposed methods and enhances the knowledge acquired in the domain.