Integrating Heterogeneous Classifier Ensembles for EMG Signal Decomposition Based on Classifier Agreement

Sarbast Rasheed, *Member, IEEE*, Daniel W. Stashuk, *Member, IEEE*, and Mohamed S. Kamel, *Fellow, IEEE*

***Abstract*—In this paper, we present a design methodology for integrating heterogeneous classifier ensembles by employing a diversity-based hybrid classifier fusion approach, whose aggregator module consists of two classifier combiners, to achieve an improved classification performance for motor unit potential classification during electromyographic (EMG) signal decomposition. Following the so-called *overproduce and choose* strategy to classifier ensemble combination, the developed system allows the construction of a large set of base classifiers, and then automatically chooses subsets of classifiers to form candidate classifier ensembles for each combiner. The system exploits kappa statistic diversity measure to design classifier teams through estimating the level of agreement between base classifier outputs. The pool of base classifiers consists of different kinds of classifiers: the adaptive certainty-based, the adaptive fuzzy k-NN, and the adaptive matched template filter classifiers; and utilizes different types of features. Performance of the developed system was evaluated using real and simulated EMG signals, and was compared with the performance of the constituent base classifiers. Across the EMG signal datasets used, the developed system had better average classification performance overall, especially in terms of reducing classification errors. For simulated signals of varying intensity, the developed system had an average correct classification rate CC_r of 93.8% and an error rate E_r of 2.2% compared to 93.6% and 3.2%, respectively, for the best base classifier in the ensemble. For simulated signals with varying amounts of shape and/or firing pattern variability, the developed system had a CC_r of 89.1% with an E_r of 4.7% compared to 86.3% and 5.6%, respectively, for the best classifier. For real signals, the developed system had a CC_r of 89.4% with an E_r of 3.9% compared to 84.6% and 7.1%, respectively, for the best classifier.**

***Index Terms*—Base classifiers, classifier agreement, classifier ensemble, diversity measure, hybrid classifier fusion, κ statistic, motor unit potential classification.**

I. Introduction

MULTICLASSIFIER decision level fusion known as *classifier fusion* is the process of combining information from a predetermined set of previously constructed base classifiers after each classifier has given an opinion about the recognition of a pattern.

Combining a set of classifiers can be viewed as a way to manage and refine the recognition limitations of the individual classifiers with the following main motivations.

1) Improving the overall performance, where multiple classifier fusion systems try to exploit the local differences between base classifiers to enhance the accuracy and the reliability of the overall combined system. Each base classifier is known to make errors in such a way that the patterns that are misclassified by the different classifiers are not necessarily the same [1].This diversity in recognition between the base classifiers is used by the classifier fusion system to effectively lower the combined classification error and enhance the decision about the pattern to be classified.
2) Some base classifiers may be expected to fail in classifying some patterns, and in this situation the overall combined system can recover the error.

The set of base classifiers that work together to solve a recognition problem and whose decisions are fused to improve the performance of the overall system is called *classifier ensemble*.

The major problems that need to be solved in classifier fusion systems are the choice of an appropriate ensemble of base classifiers to be combined and the selection of the appropriate aggregation module. Most previous work focused on the development of fusion methods, and little attention has been devoted to the development of approaches for designing an effective classifier ensemble [2].

In this paper, we present a classifier fusion system designed to consider the selection of base classifiers based on their level of agreement and using the hybrid classifier fusion module [3], [4] for base classifier decisions aggregation. The hybrid classifier fusion module consists of two classifier combiners. For each combiner, there is a prestage classifier selection module responsible for choosing subset of base classifiers from a classifiers pool to form candidate classifier ensembles by exploiting a diversity measure for designing classifier teams. The output information from the multiple classifiers may agree or conflict with each other, and the task of the designed classifier fusion system becomes the search for subsets of classifiers based on their level of agreement.

The developed diversity-based hybrid classifier fusion system has been evaluated and applied for the motor unit potential classification task during electromyographic (EMG) signal

Manuscript received March 13, 2008; revised September 17, 2008. First published January 20, 2009; current version published June 3, 2010. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

S. Rasheed and D. W. Stashuk are with the Department of Systems Design Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mails: s.rasheed@ieee.org; stashuk@pami.uwaterloo.ca).

M. S. Kamel is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: mkamel@pami.uwaterloo.ca).

Digital Object Identifier 10.1109/TITB.2008.2010552

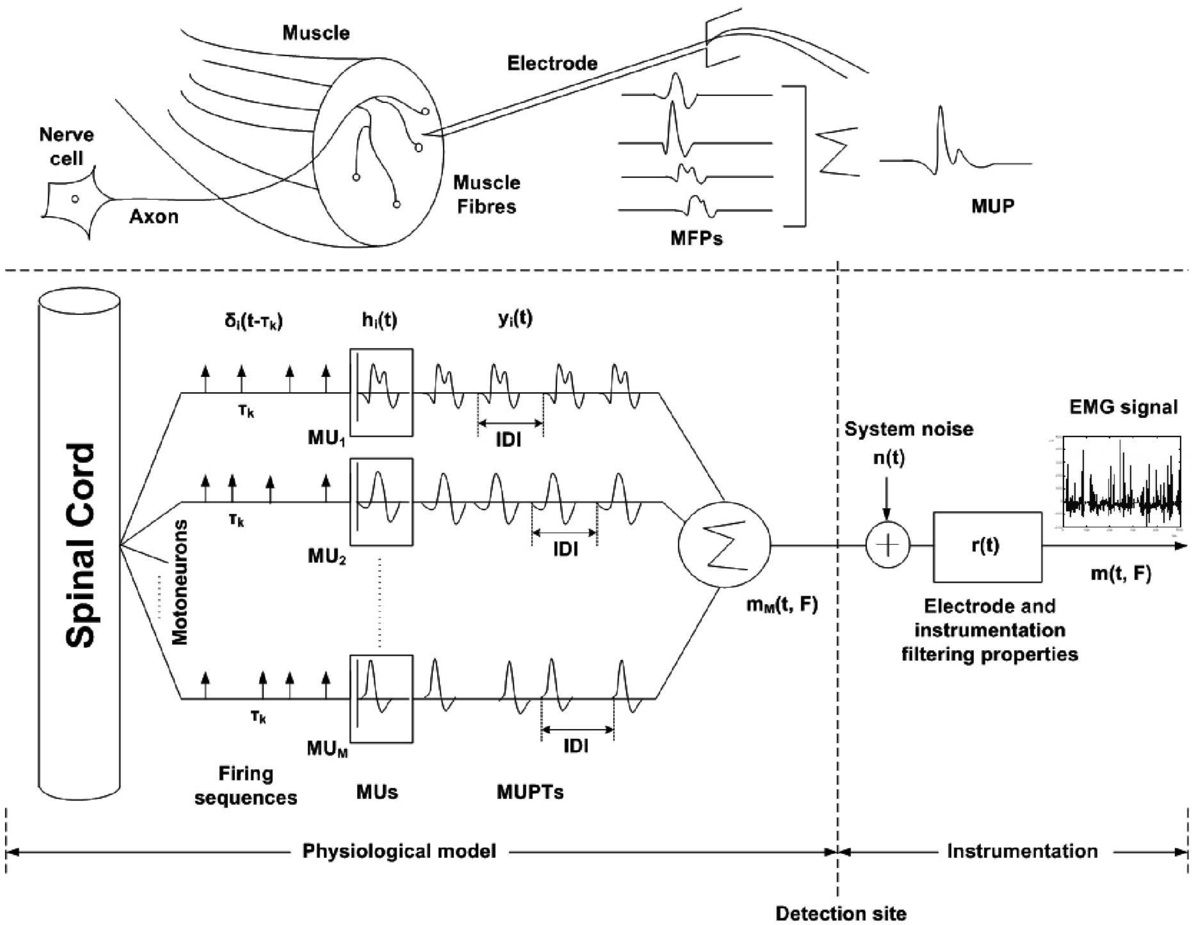


Fig. 1. Schematic representation of the model used for the composition of EMG signal.

decomposition process. Synthetic simulated EMG signals of known properties and real EMG signals had been used for the performance evaluation of the developed system, and then compared with the performance of the constituent base classifiers. Across the EMG signal datasets used, the diversity-based hybrid approach had better average classification performance overall, specially in terms of reducing the number of classification errors.

The paper is organized as follows. First, the composition and decomposition of EMG signals are discussed in Section II. Base classifiers used in the ensembles are outlined in Section III. The design process of classifier fusion systems is described in Section IV. The design of the classifier ensemble, the aggregation module, and the classification errors detection module are introduced in Sections V, VI, and VII, respectively. Methods employed for the evaluation of the developed system are explained in Section VIII. Results for two sets of simulated EMG signals and one set of real EMG signals are presented in Section IX. Finally, discussions of results and conclusions are presented in Sections X and XI, respectively.

II. Electromyographic (EMG) Signals

Electromyography is the detection of the electrical activity associated with muscle contraction. An EMG signal is obtained by measurement of the electrical activity of a muscle during contraction, and reflects the electrical depolarization of excitable muscle fiber membranes that create electrical signals called muscle fiber potentials (MFPs).

The forward problem in EMG is the composition of the electrical signal detected during a muscle contraction. In normal mammalian skeletal muscle, the fibers never contract as individuals. Instead, small groups of them contract in concert. All the fibers of each group of muscle fibers are controlled by the terminal branches of one nerve fiber or axon whose cell body is in the anterior horn of the spinal cord gray matter. These cells are the α-motoneurons and directly innervate skeletal muscle fibers.

α-motoneurons consisting of the nerve cell body, the long axon running down the motor nerve with its terminal branches, and all the muscle fibers controlled by these branches constitute a MU. The summation of all of a MU's spatially and temporally dispersed MFPs results in a signal called a motor unit potential (MUP). In order to sustain a muscle contraction, MUs must be repeatedly activated and each MU generates multiple MUPs. The collection of MUPs generated by one MU, positioned at their times of occurrence or separated by their interdischarge intervals (IDIs), is called a motor unit potential train (MUPT). The superposition of the MUPTs of all recruited MUs and background noise comprises an EMG signal.

Fig. 1 shows a muscle structure and a schematic representation of the physiological model and instrumentation for the generation of an EMG signal. α-motoneurons in the spinal cord send electrical pulses to the muscle fibers, which cause the depolarization and contraction of groups of muscle fibers. The superposition at the detection site forms the physiological EMG signal $m_M(t, F)$, which is a function of time (t) and force (F). The integer M represents the total number of MUPTs that contribute to the potential field at the detection site. When the signal is detected, an electrical noise $n(t)$ is introduced. The detected signal will also be affected by the filtering properties of the detection electrode $r(t)$, and possibly other instrumentation. The resulting signal $m(t, F)$ is the observable EMG signal.

The characteristics of an EMG signal are largely affected by anatomical and physiological properties of the muscle. For example, as the force of contraction increases, the number of MUs active and the rate at which they are active increases. The EMG signal therefore becomes more complex with increasing force of contraction. Furthermore, the fundamental structure of a muscle such as the size, distribution, and number of MUs and how they are controlled can also be reflected in the characteristics of an EMG signal. For these reasons, many researchers are interested in devising techniques for the quantitative analysis of EMG signals.

Clinically, EMG signal analysis, in the form of EMG signal decomposition and MUP classification into groups of similar shapes, is used to assist in the diagnosis of neuromuscular disorders, to analyze the neuromuscular system, and in biofeedback training.

A. EMG Signal Decomposition

The inverse problem in EMG consists of using the detected EMG signal to infer the MUPTs of the recruited MUs comprising the EMG signal and perform EMG quantification. The process of resolving a composite EMG signal into its constituent MUPTs is called EMG signal decomposition, and it can be configured as a classification problem with MUP waveforms representing the patterns to be recognized and the MUPTs as the classes into which the MUPs are to be grouped.

The objective of EMG signal decomposition is often the extraction of relevant clinical information from quantitative EMG (QEMG) analysis of individual MUP patterns and MU firing patterns [5]. The first task in EMG signal decomposition is the segmentation of the EMG signal and detection of possible MUP waveforms, which is then followed by the main task of MUP classification. The classification task involves dividing detected MUP patterns into groups such that each set of grouped MUP patterns represents the discharges of a single MU and through which the discharges of each active MU can be discriminated for subsequent processing. QEMG analysis then often involves the calculation, for each MUPT class, of a representative or template MUP waveform, which reflects information regarding individual MU morphology, and statistics related to the firing pattern of the MU.

The similarity criterion for grouping MUP patterns is usually based on a combination of MUP shape and statistics of the firing patterns of the MUs such that MUP patterns most likely belong to the same group if their shapes are closely similar and if their IDI intervals are consistent with the discharge pattern of the considered MU. This means that two kinds of information, the MUP shapes and the times of occurrences of MUP patterns should be considered for classification.

III. Base Classifiers

Each EMG signal is a classification problem in itself, and there is no a priori knowledge about the distribution of MUP data, where there are no a priori probabilities and no MUPT class conditional densities. Due to the lack of these information, base classifiers training is not possible, and we explored, for the purpose of MUP classification, nonparametric classification procedures based on a certainty measure, a nearest neighbor (NN) rule, and a similarity measure.

The base classifiers pool consists of classifiers belonging to three main types: certainty-based, fuzzy k-NN, and template matched filter classifiers. For each type, there are different configurations. There are classifier configurations based on MUP shapes only, MUP shape and passive use of firing pattern information with an adaptive train-wise setting of assignment threshold, MUP shape and active use of firing pattern information with/without adaptive train-wise setting of the assignment threshold, and others. The base classifiers similarity criterion is based on a combination of MUP shapes and a passive and an active use of MU firing patterns.

A. Certainty-Based Classification

A certainty-based classifier classifies a candidate MUP pattern to the MUPT class that produces the greatest estimated certainty, provided this maximal certainty is above a minimum certainty threshold. Classifiers that use this approach are the certainty classifier (CC) for which a complete description is given in [6]–[8] accompanied with testing and evaluation of its performance, and the adaptive CC (ACC), which is a modified version of the CC and for which a complete description is given in [4] and [9].

B. Nearest Neighbor Classification

Of the many nearest neighbor classification techniques, we chose the fuzzy k-NN classification rule, proposed by Keller *et al.* [10], as it deals with vagueness and uncertainty, and provides a confidence measure regarding the classification results.

The fuzzy k-NN classifier for MUP classification estimates a measure of assertion expressing confidence in the decision of classifying a MUP pattern to a particular MUPT class. Its adaptive counterpart, the adaptive fuzzy k-NN classifier (AFNNC), uses an adaptive assertion-based approach for assigning MUP patterns to MUPT classes. A complete description of the AFNNC is given in [4], [11], and [12].

C. Matched Template Filtering Classification

The basic MUP matched template filtering (MTF) classification approach consists of sliding MUPT class templates over the EMG signal detected MUP patterns, and for each candidate MUP pattern, a correlation measure was calculated estimating

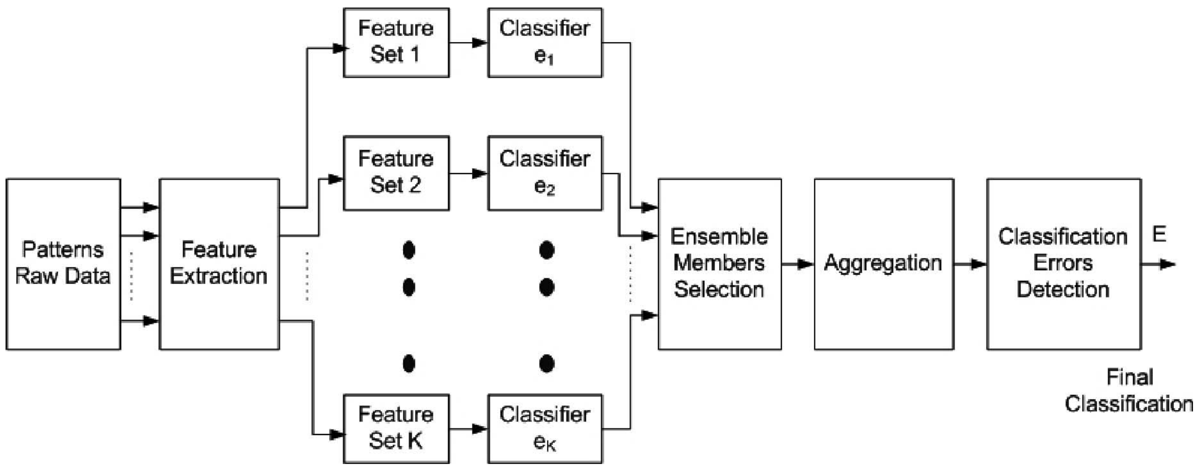


Fig. 2. Classifier fusion system basic architecture.

the degree of similarity between the template and the MUP pattern. Then, the maximum correlation position is taken to represent the instance of the template into the EMG signal under consideration with a threshold on the similarity measure allowing for rejection of poorly matched MUP patterns.

In this paper, we used two types of MTF classifiers for supervised MUP classification and associated with their adaptive variants [adaptive MTF (AMTF)]. One is based on the normalized cross correlation, which is the most widely used correlation measure [13], called the adaptive normalized cross-correlation classifier (ANCCC) and the other is based on a pseudocorrelation [14], [15] measure called the adaptive pseudocorrelation classifier (ApCC). A complete description of the ANCCC and the ApCC classifiers is given in [4].

IV. Design of a Classifier Fusion System

Classifier fusion system architecture belongs to the parallel category of combining classifiers. It consists of a set of base classifiers, invoked concurrently and independently, an aggregation module that fuses their output results, an ensemble members selection module, and a classification errors detection module, as shown in Fig. 2. The overall accuracy of the combined classifier system depends not only on the way the base classifiers are fused but also on the selection of the classifiers used in the fusion. Accordingly, the design of classifier fusion systems involves the following stages:

1) design of a classifier ensemble;
2) design of an aggregation module;
3) design of a classification error detection module.

To design the most appropriate classifier fusion system for MUP classification, we followed the so-called *overproduce and choose* [2], [16] paradigm (also called the *test and select* [17] approach) to classifier ensemble combination. The basic idea is to produce an initial large set of candidate classifier ensembles, and then select the ensemble that can be combined to achieve better accuracy.

Fig. 3 shows the main phases involved in the diversity-based hybrid classifier fusion system design cycle: the base classifier overproduction phase, the ensemble choice phase, the aggregator design phase, the classification errors detection phase, and the performance evaluation phase. The classifier ensembles or aggregation module must be redesigned when the output of the performance evaluation phase is not satisfactory, and in accordance to that, Fig. 3 shows feedbacks from later design phases to the earlier ones.

The base classifiers overproduction design phase produces a large set of base classifiers. The ensemble choice phase selects the subsets of classifiers that can be combined to achieve better accuracy. The subset giving the best accuracy could be obtained by exhaustive enumeration, such that, if K is the size of the base classifier set produced by the overproduction phase, the number of possible subsets is equal to $\sum_{i=2}^{K} \binom{K}{i}$ for classifier ensembles of two classifiers and up to K classifiers. Therefore, there is a need for a strategy to limit the computational complexity of the choice phase and follow techniques that choose an effective classifier ensemble without hypothesizing a specific combination rule [18]. For this purpose, we exploited ensemble diversity measure techniques to evaluate the error diversity of classifiers that make up an ensemble for classifier selection purposes. We chose the kappa statistic to select base classifiers having an excellent level of agreement to form ensembles giving satisfactory classification performance.

In the aggregation module design phase, the choice of the combination function should take into account the dependency among classifiers. In actual practice, a trial-and-error procedure is performed, because a clear model of the dependency among classifiers is difficult to obtain [18]. If all classifiers in an ensemble are totally positively dependent, i.e., they are identical, there will be no improvement in performance. However, if they are negatively dependent, i.e., the base classifiers commit mistakes on different MUP patterns, we could expect improvement in performance. This raises the issue of quantifying the degree of agreement among dependent classifiers.

The classification errors detection phase is responsible for reducing the number of erroneous assignments in the generated MUPT classes by the diversity-based hybrid classifier fusion system.

Performance evaluation is performed by assessing the classification accuracy, in terms of performance indices presented

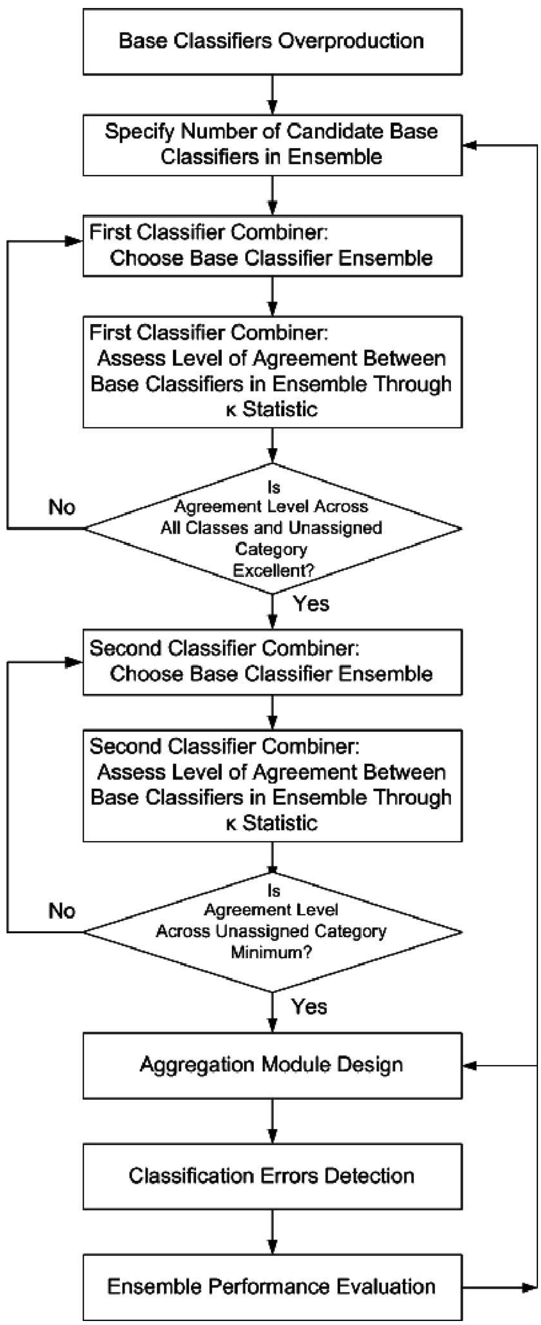


Fig. 3. Diversity-based hybrid classifier fusion system design cycle.

in Section VIII-C, of the selected classifier ensemble using the designed classifier aggregation module.

V. Design of the Classifier Ensemble

The improvement in performance of a classifier fusion system relies on properly choosing the base classifiers to be fused. As stated in Section IV, choosing base classifiers can be performed directly through exhaustive search with the performance of the fusion being the objective function. As the number of base classifiers increases, this approach becomes computationally too expensive.

Instead of the exhaustive search method of choosing base classifiers, we exploited a diversity measure for designing classifier teams. Diversity, as a measure, has been used for selecting ensembles in design of multiple classifier systems [19].

To evaluate the diversity in this paper, we chose the kappa measure κ to estimate the level of agreement between the base classifier outputs [20], i.e., to measure the degree of decision similarity between the base classifier outputs, for the following reasons:

1) κ statistics is a nonpairwise diversity measure used for measuring the degree of decision similarity between the base classifier outputs;
2) κ statistics expresses a special type of relationship between classifiers, and quantifies the level to which the classifiers agree in their decisions beyond any agreement that could occur due to chance. A value of 0.40 is considered to represent poor agreement beyond chance, values between 0.40 and 0.75 indicate fair agreement, and values beyond 0.75 indicate excellent agreement [20];
3) κ depends on the individual accuracies of the classifiers;
4) κ has a specific value 0 for statistically independent classifiers;
5) κ varies between -1 and $+1$. κ close to 1 indicates that the classifiers agree in the recognition of the same MUP patterns, and $\kappa = -1$ means that the assignment by the classifiers is different for the same MUP patterns.

The kappa statistic was first proposed by Cohen [21] and has the same interpretation as the *intraclass correlation coefficient* [20].

The kappa statistic corrects for chance expected agreement meaning that it should reflect the amount of agreement in excess of what would be expected by chance. It not only gives a measure of the degree of agreement, but it also has a test associated with it that can be employed to check if the apparent agreement cannot be attributed to chance only.

For an ensemble of K base classifiers $e_k, k = 1, 2, \ldots, K$, known to be dependent on each other as they work on the same data, used to classify a set of N MUP patterns into M MUPT classes and the unassigned category $\omega_i \in \Omega = \{\omega_1, \omega_2, \ldots, \omega_M, \omega_{M+1}\}$, we want to estimate the strength of the association among them by measuring the degree of agreement among dependent classifiers. For $j = 1, 2, \ldots, N$, and $i = 1, 2, \ldots, M + 1$, denote by d_{ji} the number of classifiers that assign candidate MUP pattern m_j to class ω_i, i.e.

$$d_{ji} = \sum_{k=1}^{K} T(e_k(m_j) = \omega_i) \tag{1}$$

where $T(e = \sigma)$ is a binary characteristic function and equals 1 if $e = \sigma$ and 0 otherwise. Note that $\sum_{i=1}^{M+1} d_{ji} = K$ for each MUP pattern m_j. Table I shows the per MUP pattern diversity matrix through d_{ji}.

Based on the per MUP pattern diversity matrix of K classifiers, the degree of agreement among dependent classifiers $e_k(m_j), k = 1, 2, \ldots, K$, in classifying MUP pattern m_j is measured using the following kappa hat statistic formula for

TABLE I
PER MUP PATTERN DIVERSITY MATRIX OF K CLASSIFIERS

MUP Pattern	MUPT ω_1	MUPT ω_2	...	MUPT ω_M	MUPT ω_{M+1}	$\sum_{i=1}^{M+1} d_{ji}^2$
m_1	d_{11}	d_{12}	...	d_{1M}	$d_{1(M+1)}$	$\sum_{i=1}^{M+1} d_{1i}^2$
m_2	d_{21}	d_{22}	...	d_{2M}	$d_{2(M+1)}$	$\sum_{i=1}^{M+1} d_{2i}^2$
.	.	.	...	.	.	.
.	.	.	...	.	.	.
.	.	.	...	.	.	.
m_N	d_{N1}	d_{N2}	...	d_{NM}	$d_{N(M+1)}$	$\sum_{i=1}^{M+1} d_{Ni}^2$
Total	$\sum_{j=1}^{N} d_{j1}$	$\sum_{j=1}^{N} d_{j2}$	...	$\sum_{j=1}^{N} d_{jM}$	$\sum_{j=1}^{N} d_{j(M+1)}$	$\sum_{j=1}^{N}\sum_{i=1}^{M+1} d_{ji}^2$

multiple outcomes (classes) and multiple classifiers [20]:

$$\hat{\kappa} = 1 - \frac{NK^2 - \sum_{j=1}^{N}\sum_{i=1}^{M+1} d_{ji}^2}{KN(K-1)\sum_{i=1}^{M+1} \bar{p}_i \bar{q}_i} \tag{2}$$

where $\bar{p}_i = (\sum_{j=1}^{N} d_{ji}/NK)$ represents the overall proportions of outputs of classifiers in MUPT class ω_i, and $\bar{q}_i = 1 - \bar{p}_i$. The value of the kappa hat statistic $\hat{\kappa}_i$ for MUPT class $\omega_i, i = 1, 2, \ldots, M$, and the unassigned category ω_{M+1} may be measured using the formula [20]

$$\hat{\kappa}_i = 1 - \frac{\sum_{j=1}^{N} d_{ji}(K - d_{ji})}{KN(K-1)\bar{p}_i \bar{q}_i}. \tag{3}$$

The proposed diversity-based hybrid classifier fusion approach uses formulas (2) and (3) for selecting the base classifiers comprising the ensembles, as described in Section VI-A.

VI. Design of the Aggregation Module

The next step after selecting the individual classifiers to form a team in an ensemble is combining the classifier outputs by the aggregator module. Classifiers used in this paper express their decisions and provide information about identifying an MUP pattern at two levels.

1) *Abstract level:* The classifier output is a unique MUPT class label or several MUPT class labels, in which case the MUPT classes are equally identified without any qualifying information. A well-known fusion method for this level is the *majority voting* [22], [23].
2) *Measurement level:* The classifier attributes to each MUPT class label a confidence measure value representing the degree to which the MUP pattern has that MUPT class label. *Simple averaging* [23], *weighted averaging* [24], and *fuzzy integral* [25], [26] fusion methods are a few techniques for this level.

Consider a decision space P with M mutually exclusive sets $\omega_i \in \Omega = \{\omega_1, \omega_2, \ldots, \omega_M\}$. Each set ω_i represents a class or category into which MUP patterns will be grouped or classified. The decision space may be written as

$$P = \omega_1 \bigcup \omega_2 \bigcup \cdots \bigcup \omega_M. \tag{4}$$

The decision space P is the set of all possible MUP patterns from all MUPT classes. The set of the corresponding integer labels Ω is defined such that $\Omega = \{\omega_1 = 1, \omega_2 = 2, \ldots, \omega_M = M\}$, and provides all possible integer labels for the valid MUPT classes. As some of the MUP patterns may not be assigned to any of the valid MUPT classes, the decision space set P can then be extended to include $\Omega \bigcup \{\omega_{M+1}\}$, where ω_{M+1} designates the unassigned category that by some established criteria the classifier decides not to assign the input MUP pattern.

For an ensemble of K classifiers $e_1, e_2, \ldots, e_K$, each recognition engine in the system may be simply regarded as a functional box that receives an input MUP pattern $\{m_j, j = 1, 2, \ldots, N\}$ and outputs an MUPT class label $\{\omega_i, i = 1, 2, \ldots, M, M+1\}$ denoted by $e_k(m_j) = \omega_i$ and associated with decision function values that can be used as measure of confidences in the decision of classifying a MUP pattern to a particular MUPT class.

When combining classifier outputs based on the abstract level, we use the MUPT class label provided by the respective classifiers, whereas when combining classifier outputs based on the measurement level, we use the real valued outputs for each MUPT class provided by the classifiers in the ensemble. Assuming when classifying a MUP pattern m_j, each classifier produces output values, in the interval [0, 1], interpreted as a confidence $Cf_i(m_j)$ in the decision of classifying MUP pattern m_j with respect to a particular MUPT class ω_i $(i = 1, 2, \ldots, M)$. One can think of these outputs as *a posteriori* probabilities such that the confidence $Cf_i(m_j)$ is defined as

$$Cf_i(m_j) = P(\omega_i | m_j) \tag{5}$$

and in relation with a specific classifier $e_k(m_j)$, the confidence depends on the outcome $e_{ik}(m_j)$ of this classifier for MUPT class ω_i

$$Cf_{ik}(m_j) = P(\omega_i | e_{ik}(m_j)). \tag{6}$$

The confidences in (5) and (6) are defined over all the detected MUP patterns in an EMG signal.

Once a set of decision confidences $\{Cf_{ik}(m_j), i = 1, 2, \ldots, M; k = 1, 2, \ldots, K\}$ for M MUPT classes and K classifiers is computed for a MUP pattern m_j, they can be combined with a classifier fusion module into a new set of decision confidences that can be used, by maximum selection, for the final classification.

Fixed combination rules [1], [27] do not require training to fuse classifier outputs. These schemes are based on the *product*, *sum*, *max*, *min*, *median*, and *average* rules. For a candidate MUP pattern m_j, they are used for combining the set of decision confidences $\{Cf_{ik}(m_j), i = 1, 2, \ldots, M; k = 1, 2, \ldots, K\}$ for M MUPT classes and K base classifiers

$\{e_k(m_j), k = 1, 2, \ldots, K\}$ into combined classifier decision confidences $\{Q_i(m_j), i = 1, 2, \ldots, M\}$

$$Q_i(m_j) = \text{rule}\,\{Cf_{ik}(m_j)\}\,. \tag{7}$$

We used the average fixed combination rule as it shows better classification performance than the other fixed rules. The combined decision confidence $Q_i(m_j)$ for MUPT class ω_i is computed by

$$Q_i(m_j) = \frac{\sum_{k=1}^{K} Cf_{ik}(m_j)}{K}. \tag{8}$$

The final classification is made by

$$\omega(m_j) = \arg\max_{i=1}^{M}(Q_i(m_j)). \tag{9}$$

The proposed diversity-based hybrid aggregation module is a hybrid combination [3] of two classifier fusion schemes: the first is based on the abstract level and the second is based on the measurement level, and with each combiner, there is a prestage classifier selection module responsible for selecting the base classifiers comprising the ensemble. Both combiners may be either data independent or the first combiner data independent and the second data dependent. For the purpose of experimentation of this paper, we used as the first combiner the majority voting scheme, while we used as the second combiner one of the fixed combination rule behaving as a data-independent combiner or the fuzzy integral with the λ-fuzzy measure as an implicit data-dependent combiner. A complete description of the data independent, data dependent, implicit data dependent, majority voting, fixed combination rules, and the fuzzy integral one-stage classifier fusion scheme is presented in [3], [4], and [28].

A. Diversity-Based Hybrid Classifier Fusion Scheme

The hybrid classifier fusion scheme described in [3] uses as a classifier ensemble a fixed set of classifiers, and consequently, both combiners act on the outputs of the same classifiers. The drawback of this scheme is apparent when following the *overproduce and choose* [2], [16] paradigm for the choice phase in the classifier fusion system design cycle, as shown in Fig. 3, where there is a need to perform an exhaustive search for the best accurate classifier ensemble. For example, if the pool of base classifiers contains 16 base classifiers and the intention is to choose an ensemble of six base classifiers for fusion, so there is a need to compare the performance of $\binom{16}{6} = 8008$ ensembles. Therefore, to limit the computational complexity encountered, we modified the hybrid classifier fusion scheme [3] so that the candidate classifiers chosen for fusion by both combiners are based on a diversity measure.

The diversity-based hybrid classifier fusion scheme is a two-stage process and consists of two combiners with a prestage classifier selection module for each combiner. Once the pool of base classifiers has been constructed, it is not clear which subset of base classifiers when combined gives the best performance. The ensemble candidate classifiers selected for combination are decided by assessing the degree of agreement by exploiting the kappa statistic diversity measure.

The MUP pattern dataset to be classified may contain MUP patterns with different complexities. There may be single MUP patterns each belonging to one MU, partially superimposed MUP patterns each representing the overlap of more than one MUP pattern without their peaks being obscured, completely superimposed MUP patterns representing the overlap of more than one MUP pattern and in which the peaks of the MUP patterns combine to make one large peak, and destructively superimposed MUP patterns representing the overlap of more than one MUP pattern and in which the MUP patterns are superimposed in such a way that their out-of-phase peaks are summed together and cancel each other [29].

Based on the previous MUP pattern dataset, the stage of designing the classifier ensemble classifier fusion system is involved with the prestage classifier selection module for each combiner, such that the first combiner deals with combining the outputs of the classifier ensemble having the excellent level of agreement throughout all the MUPT classes and the unassigned category using (2). This means that we want the first combiner to assign a large number of single MUP patterns correctly based on the highest level of consensus among the members of the ensemble and leave the classification of the more complex MUP patterns to a different classifier ensemble whose outputs are to be combined by the second combiner. At the second combiner, the team's classifiers selection is based on assessing the classifier agreement considering only the unassigned category using (3) and choosing the ensemble having the least level of agreement. This is due to the fact that most classifiers face difficulties when classifying the more complex MUP patterns and do not assign them to any valid MUPT classes, and also due to the lack of consensus among the members of the ensemble.

The stage of designing the aggregation module for the diversity-based hybrid classifier fusion system is involved with the choice of the combiners themselves. We chose the first combiner at the abstract level of classifier fusion represented by the majority voting data-independent aggregator, whereas we chose the second combiner at the measurement level of classifier fusion represented by either one of the fixed combination rules or the fuzzy integral.

The diversity-based hybrid fusion scheme works as follows.

1) *First stage:* The ensemble candidate classifiers selected for combination by the first combiner are those having the maximum degree of agreement, i.e., having the maximum value of kappa statistics $\hat{\bar{\kappa}}$ evaluated using (2). The outputs of the ensemble candidate classifiers are presented to the majority voting combiner. If all the classifiers state a decision that a MUP pattern is left unassigned, then there is no chance to reassign that MUP pattern to a valid MUPT class and it stays unassigned. If over half of the classifiers assign a MUP pattern to the same MUPT class, then that MUP pattern is allocated to that MUPT class and no further assignment is processed. For these MUP patterns, an overall confidence value is calculated by averaging the confidence values given by the ensemble classifiers who contributed in the decision of assigning the MUP pattern. In all other situations, i.e., when half or less than half of the classifiers specify a decision for a MUP pattern to

be assigned to the same MUPT class, the measurement level combination scheme is used in the second stage to specify to which MUPT class the MUP pattern should be assigned based on which MUPT class has the largest combined confidence value. From the first stage, a set of incomplete MUPT classes is generated missing those MUP patterns that need to be assigned to a valid MUPT class in the second stage.

2) *Second stage:* This stage is used for those MUP patterns for which only half or less than half of the ensemble classifiers in the first stage specify a decision for a MUP pattern to be assigned to the same MUPT class. The ensemble candidate classifiers selected for combination at the second combiner are those having a minimum degree of agreement considering only the unassigned category, i.e., having the minimum value of kappa statistics $\hat{\kappa}_i$ evaluated using (3) for $i = M + 1$. The outputs of the ensemble classifiers are presented to one of the fixed rule combiners or one of the trainable combiners represented by a Sugeno [26] or Choquet [30] fuzzy integral combiner. For each MUP pattern, overall combined confidence values representing the degree of membership in each MUPT class are determined, and accordingly, the MUP pattern is assigned to the MUPT class for which its determined overall combined confidence is the largest and if it is above the specified combiner confidence threshold set for that MUPT class, thus forming a more complete set of MUPT classes.

VII. Design of Classification Errors Detection Module

To reduce the number of erroneous assignments in the generated MUPT classes by the diversity-based hybrid classifier fusion approach, we adopted the adaptive setting of MUPT class assignment threshold used for base classifiers described in [9] and [12].

A. Motor Unit Firing Pattern Statistics

The classification of MUP patterns to MUPT classes cannot be considered independent of MU firing pattern constraints. These constraints include MU physiological limitations, the likelihood of a particular MU firing in a given epoch of time, and the expectation that the classification algorithm might not assign each detected MUP. For these reasons, the generated MUPT classes may not represent the correct discharges of the active MUs and the need will arise to reclassify some MUP patterns.

Two modes of use of MU firing pattern information are possible: passive and active. Passive mode use refers to the use of firing pattern information in the adjustment of the assignment threshold of a MUP to a MUPT class based on its firing pattern consistency statistics that adjust MUP assignments in general but not the assignment of any specific MUP, whereas active mode use refers to the situation in which firing pattern information is used to determine to which class each specific MUP should be assigned.

To take MU firing pattern constraints into consideration during MUP classification task, we specified a set of firing time consistency statistics that assists in the detection of MUP misclassifications.

B. IDI Statistics for the Detection of MUP Misclassifications

Errors in the determination of the activation pattern of a MU, caused by erroneous MUP classifications, can be detected by analyzing IDI statistics. Specific erroneous IDIs can be identified, and determinations regarding the number of erroneous IDIs can be made. The following firing pattern consistency statistics can be used to determine if a significant number of IDI errors exist in a MUPT class.

1) *Percentage of inconsistent IDIs:* An IDI inconsistency in a MUPT class can be defined as those IDIs less than $\mu - 2\sigma$ and any IDI less than 15 ms.
2) IDIs coefficient of variation (CV): defined as
$$\mathrm{CV} = \frac{\sigma}{\mu}. \tag{10}$$
3) IDIs lower coefficient of variation CV_l: defined as the ratio of lower standard deviation σ_l of the distribution of the MUPT class IDIs and the MUPT class mean IDI μ, and is given by
$$\mathrm{CV}_l = \frac{\sigma_l}{\mu} \tag{11}$$
where σ_l is calculated using the IDIs whose values are less than the mean value.
4) Lower IDI ratio: defined as the ratio of the count of IDIs whose values are less than 0.5 of the MUPT class mean IDI μ to the count of the IDIs whose values are less than the MUPT class mean IDI μ.

To get accurate estimates of the IDIs μ and σ, the error-filtered estimation (EFE) [31], [32] algorithm was used.

C. Adaptive Setting of MUPT Class Assignment Threshold

The adaptive nature of MUP classification is related to the adjustment of the minimal assignment threshold for each MUPT class based on firing pattern statistics. We start with a very low assignment threshold and test all MUPT classes against firing pattern constrains. The assignment threshold of a MUPT class might be increased to exclude MUP patterns causing firing pattern inconsistencies. The occurrence of a significant number of MUP classification errors is detected by the use of the firing pattern consistency statistics described in Section VII-B.

The MUP classification using the diversity-based hybrid classifier fusion approach is based on the outputs of the base classifiers and does not take into consideration the MU firing patterns. Therefore, following the generation of MUPT classes, the MUPT classes should be tested for any firing pattern inconsistencies, and if, based on firing pattern statistics, it is expected that a MUPT class has too many erroneous assignments, its minimal assignment threshold is increased or otherwise it is kept constant. MUPT classes to which MUP patterns can be confidently assigned will have a lower minimal assignment threshold and have a higher MUP identification (ID) rate. Alternatively,

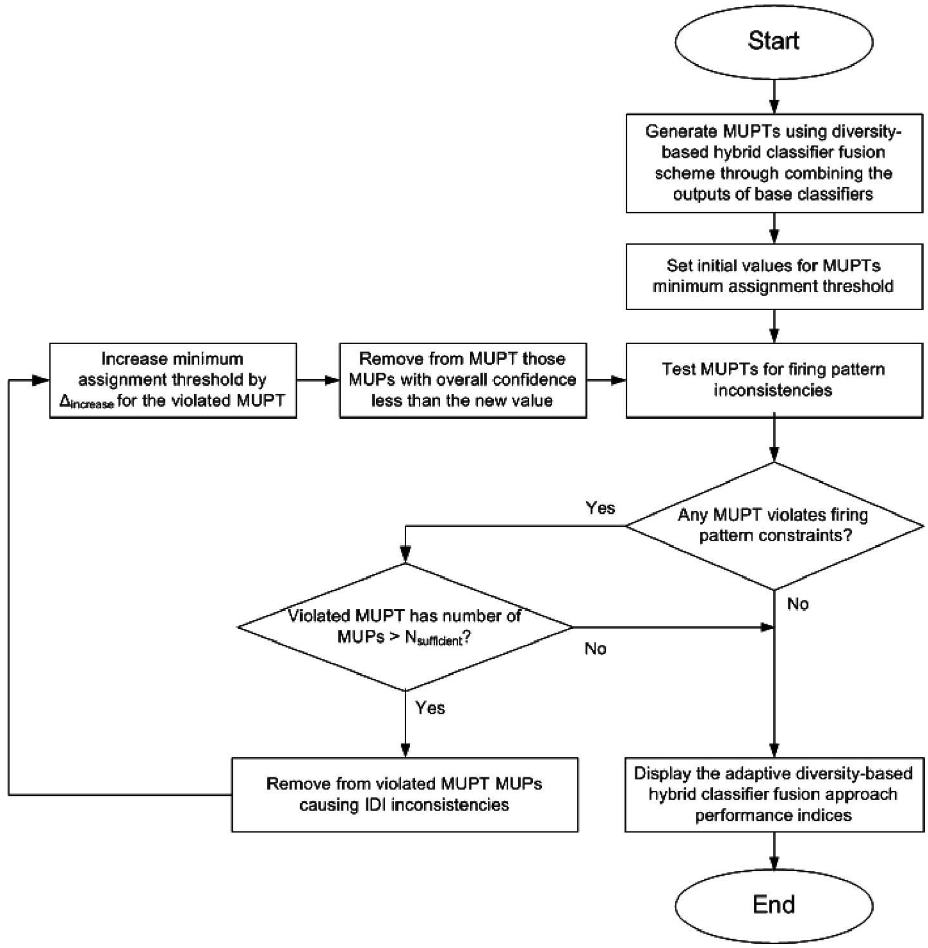


Fig. 4. Flowchart of the adaptive classification using the diversity-based hybrid classifier fusion approach.

MUPT classes to which MUP patterns cannot be confidently assigned, possibly because of shape similarity with the MUP patterns of another MUPT class or MUPT class shape and firing pattern variability, will have a higher minimal assignment threshold and have a lower MUP ID rate.

The adaptive MUP classification consists of only one stage and is based on the passive use of firing pattern information to remove possible erroneous classifications. Fig. 4 shows the steps involved in this process. Each candidate MUP m_j is assigned to the MUPT that has the highest confidence of belonging to it expressed in terms of the maximum assignment value and if it is higher than the minimum assignment threshold of that MUPT.

Once the set of MUPT classes is created by the diversity-based hybrid classifier fusion system, the firing pattern consistency statistics for each MUPT class with $N_{\text{sufficient}}$ or more IDIs are calculated. A small percent of those MUP patterns causing IDI inconsistencies is allowed. The firing pattern consistency statistics of each MUPT class with $N_{\text{sufficient}}$ or more IDIs are then recalculated and checked against the following constraints:

1) percentage number of IDI inconsistencies $\leq 2.5\%$;
2) absolute value of $\text{CV}_l \leq 0.35$, representing an upper bound of physiological expectation;
3) ratio of CV_l to $\text{CV} \leq 1.25$, empirically determined value;
4) lower IDI ratio ≤ 0.175, empirically determined value.

If a MUPT class meets all the previous constraints, it keeps its MUP patterns and its minimum assignment threshold is unchanged. Otherwise, the minimum assignment threshold for the MUPT is increased by Δ_{increase} and its MUP patterns confidence values are checked against the new value. Those MUP patterns causing IDI inconsistencies in the firing pattern of a MUPT and MUP patterns with confidence values less than the minimum assignment threshold are designated unassigned and removed from the MUPT. This process is repeated until all the imposed firing pattern constraints for all generated MUPT classes are satisfied. A MUPT's minimum assignment threshold is not increased above an extreme value if it has fewer than a minimum number of MUP patterns $N_{\text{sufficient}}$. Used value for $N_{\text{sufficient}}$ was 50 MUP patterns and for Δ_{increase} was 0.001.

VIII. Evaluation Methodology

The testing process and evaluation of performance of the base classifiers and the diversity-based hybrid classifier fusion system require a reference decomposition result. For this purpose, different approaches can be followed [33]–[35]. Two of these approaches were used here:

1) using simulated EMG signals with specific properties as a reference;

2) using real EMG signals decomposed manually by a human expert as a reference.

The methods employed for the evaluation process are described in the following sections.

A. MUP Representation and Feature Extraction

The classification task completed by the base classifiers uses different domain feature sets: MUP time-domain and wavelet-domain features.

The first task in EMG signal decomposition is segmentation and MUP detection. During segmentation, an EMG signal is divided into segments containing significant MUP waveforms, and segments assumed to contain MUP patterns from distant MUs or signal baseline. A segment can contain either a single MUP or superimposed MUP patterns. In this paper, MUP occurrence times were provided either by the simulator output, for simulated data, or by the manual decomposition results, for real data. MUP patterns were represented using 80 time-domain data points centered on the MUP occurrence time provided. All MUP data were created with a 31.25-kHz sampling rate. Isolated and superimposed MUP patterns were included.

For each EMG signal investigated, we extracted from the MUP raw data feature space the first-order discrete derivative features by applying a low-pass differentiation filter to accentuate the MUP's rapidly rising edges (high-frequency components) and shorten the duration of MUP patterns, which reduces their temporal overlap, and therefore reduces the number of superimposed waveforms. This filtering also suppresses low-frequency background noise. The effect is to transform the MUP patterns into narrow spikes that stand out clearly from a flat baseline.

Denoting the discrete samples of the raw data as $m_j[i]$ and the first-order discrete derivative data as $y^{(1)}[i]$, then to perform first-order differentiation we used the filter given by (12) and (13) [36], [37]

$$y^{(1)}[i] = \left(\frac{1}{2}\right) \sum_{k=1}^{W} C_k \; (m_j[i+k] - m_j[i-k])$$
$$\text{with} \quad \sum_{k=1}^{W} k\, C_k = 1 \tag{12}$$

where W is the window width for adjusting the cutoff frequency. Taking the window width $W = 3$, the optimum coefficients C_k are $C_1 = C_2 = 0,\ C_3 = 1/3$. For these values, (12) becomes

$$y^{(1)}[i] = \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) \; (m_j[i+3] - m_j[i-3]). \tag{13}$$

The wavelet-domain features are extracted by taking the discrete-time wavelet transform (DTWT) of the EMG signal MUP patterns. The DTWT is computed using the wavelet decomposition algorithm of the fast wavelet transform (FWT) through multiresolution analysis, and is implemented using a filter bank structure consisting of only the analysis bank.

The wavelet-domain full MUP pattern feature vector after the full wavelet decomposition with M scale levels would be the concatenation of the detail coefficients at all scale levels and the approximation coefficients at the last scale level.

B. Classifier Seeding

For each EMG signal decomposed, a single reference set of correctly classified MUP patterns was chosen using information provided by the EMG simulator or from manually decomposed real data to seed the supervised base classifiers evaluated. For each MUPT class, 20 isolated MUP patterns were selected. The isolated MUP patterns had occurrence times that were separated by more than 3 ms from any other MUP occurrence times to avoid choosing superimposed MUP patterns. The selected reference set was used to calculate initial MUPT templates for the certainty-based classifiers and template matched filtering classifiers and to establish core membership values for the fuzzy k-NN classifiers.

C. Classifier Performance Indices

The performance of each individual base classifier and the diversity-based hybrid classifier fusion system were evaluated in terms of the numbers of assigned and rejected MUP patterns and the numbers of correctly and erroneously classified MUP patterns and from which a set of related performance indices was determined.

The assignment rate $A_r\%$ is defined as the ratio of the total number of assigned MUP patterns, which is equal to the total number of MUP patterns detected minus the number of MUP patterns unassigned, to the total number of MUP patterns detected

$$A_r\% = \frac{\text{number of MUPs assigned}}{\text{total number of MUPs detected}} \times 100. \tag{14}$$

The error rate $E_r\%$ is defined as the ratio of the number of MUP patterns erroneously classified to any valid MUPT class to the number of MUP patterns assigned

$$E_r\% = \frac{\text{number of MUPs erroneously classified}}{\text{number of MUPs assigned}} \times 100. \tag{15}$$

The correct classification rate $CC_r\%$ is defined as the ratio of the number of correctly classified MUP patterns, which is equal to the number of MUP patterns assigned minus the number of MUP patterns erroneously classified, to the total number of MUP patterns detected

$$CC_r\% = \frac{\text{number of MUPs correctly classified}}{\text{total number of MUPs detected}} \times 100. \tag{16}$$

Classifier performances are compared in terms of the previous set of performance indices such that the classifier with the *better performance* is the one with the highest correct classification rate $CC_r\%$ and lowest error rate $E_r\%$. In situations when the highest correct classification rate $CC_r\%$ and lowest error rate $E_r\%$ does not judge the differentiation between two classifiers, we take the difference between the correct classification rate $CC_r\%$ and error rate $E_r\%$ for each classifier and consider the classifier with the higher difference as the one with the better performance.

D. EMG Signal Data Sets Used

The effectiveness of using the diversity-based hybrid classifier fusion system for EMG signal decomposition was demonstrated through the analysis of simulated and real EMG signals. For comparison with the performance of the constituent base classifiers, we used the same data sets used for evaluating them presented in [9] and [12]. The EMG signal data used consisted of two sets of simulated EMG signals: an independent set and a related set, each of 10 s length, and a set of real EMG signals.

The simulated data were generated from an EMG signal simulator based on a physiologically and morphologically accurate muscle model [38]. The simulator enables us to generate EMG signals of different complexities with knowledge of the signal intensity represented by the average number of MUP patterns per second (pps), the numbers of MUPT classes, and which MU created each MUP pattern. Furthermore, the amount of MUP shape variability represented by jitter and/or IDI variability can be adjusted. Jitter represents the standard deviation of the initiation times of each MFP [38].

The EMG signals within the set of independent simulated signals have different levels of intensity and each have unique MUPT classes and MUP distributions. The EMG signals within the set of related simulated signals have the same level of intensity and the same MUPT classes and MUP distributions but have different amounts of MUP shape and firing pattern variability.

The set of real EMG signals is of different complexities and was detected during slight to moderate levels of contraction. They have been decomposed manually by an experienced operator using a computer-based graphical display algorithm.

IX. Data Analysis Results

This section reports classification performance results of the assigned, rejected, correctly classified, and erroneously classified MUP waveforms and presents them in terms of mean and mean absolute deviation (MAD) of the assignment rate A_r, error rate E_r, and correct classification rate CC_r defined in Section VIII-C. The results obtained using the base classifiers and the hybrid classifier fusion approaches are presented in this section.

The wavelet-domain MUP feature vectors used for classification purposes were extracted through multiresolution analysis with Daubechies 4 wavelet filters and six scale levels. The used wavelet-domain MUP feature vector is a subset of the full feature vector containing only the detail coefficients at the high scale levels. For $M = 6$ levels, we noted that taking only the detail coefficients at levels 4, 5, and 6 gave better classification performance.

The base classifiers used for experimentation belong to the types described in Section III. For each kind, we used four classifiers, i.e., for ACC classifiers e_1, e_2, e_3, e_4, for AFNNC classifiers e_5, e_6, e_7, e_8, for ANCCC classifiers $e_9, e_{10}, e_{11}, e_{12}$, and for ApCC classifiers $e_{13}, e_{14}, e_{15}, e_{16}$. Classifiers e_1, e_5, e_9, e_{13} were fed with time-domain first-order discrete derivative features and using MUP patterns with sequential assignment for seeding. Classifiers e_2, e_6, e_{10}, e_{14} were fed with time-domain first-order discrete derivative features and using high-certainty MUP patterns for seeding. Classifiers e_3, e_7, e_{11}, e_{15} were fed with wavelet-domain first-order discrete derivative features and using MUP patterns with sequential assignment for seeding. Classifiers e_4, e_8, e_{12}, e_{16} were fed with wavelet-domain first-order discrete derivative features and using the highest shape certainty MUP patterns for seeding.

We performed a sensitivity analysis for deciding how many base classifiers to use in an ensemble. An exhaustive search for the best performing classifier ensemble was performed, and as we tried different sizes of base classifier ensembles, we concluded that choosing an ensemble of six base classifiers gives a better performing ensemble.

With the diversity-based hybrid classifier fusion scheme, we performed two experiments. In the first experiment, we used a base classifier pool containing eight base classifiers $e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8$ from which we selected six classifiers to work as a team in the ensemble for every signal at each stage combiners. The number of classifier ensembles that can be created is $\binom{8}{6} = 28$ ensembles.

The performance with respect to each signal, the mean performance, and the MAD of the performance of the diversity-based hybrid classifier fusion scheme, i.e., the hybrid combiner consisting of majority voting and average fixed combination rule and the hybrid combiner consisting of majority voting and Sugeno fuzzy integral with the prestage classifier selection module for each combiner, are reported in [4].

In the second experiment, we used a base classifier pool containing 16 base classifiers $e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9, e_{10}, e_{11}, e_{12}, e_{13}, e_{14}, e_{15}, e_{16}$ from which we selected six classifiers to work as a team in the ensemble for every signal at each stage combiners. The number of classifier ensembles that can be created is $\binom{16}{6} = 8008$ ensembles.

Results for the individual base classifiers for each EMG signal in the data set were reported in [4]. For comparison purposes, Table II shows the mean performance and the MAD of all the base classifiers and the diversity-based hybrid classifier fusion schemes across all the EMG signals for each data set.

Based on the criterion used for deciding which classifier has better performance, as stated in Section VIII-C, Table III shows a performance comparison of the different classification approaches across the three EMG signal data sets.

From Table III, we see that the adaptive diversity-based majority voting with average fixed rule hybrid classifier fusion scheme outperforms the average performance of the constituent base classifiers for all the data sets and also outperforms the best base classifier except across the independent simulated signals where it has the same performance.

X. Discussion

The performance of the base classifiers used and the performance of the classifier fusion systems can be explained considering the different complexities of simulated and real EMG signals. Simulated EMG signals allowed the performance to be evaluated in a controlled way. Signals of known compositions

TABLE II
MEAN AND MAD OF ASSIGNMENT RATE A_r, ERROR RATE E_r, AND CORRECT CLASSIFICATION RATE CC_r FOR THE DIFFERENT CLASSIFICATION APPROACHES ACROSS THE THREE EMG SIGNAL DATA SETS

Classifier	**Independent simulated signals**			**Related simulated signals**			**Real signals**		
	$A_r\%$	$E_r\%$	$CC_r\%$	$A_r\%$	$E_r\%$	$CC_r\%$	$A_r\%$	$E_r\%$	$CC_r\%$
Base Classifier e_1	85.2 (5.2)	1.8 (0.9)	83.7 (5.8)	84.0 (3.6)	4.9 (2.2)	79.9 (4.7)	83.9 (2.8)	2.9 (2.7)	81.4 (3.6)
Base Classifier e_2	86.9 (5.2)	1.6 (0.9)	85.5 (5.8)	86.0 (3.7)	5.1 (2.3)	81.6 (4.1)	80.5 (4.4)	4.7 (5.8)	76.7 (6.1)
Base Classifier e_3	86.9 (4.6)	2.5 (1.3)	84.8 (5.4)	86.2 (3.8)	5.7 (2.6)	81.3 (4.2)	82.6 (2.2)	5.9 (5.2)	77.8 (5.8)
Base Classifier e_4	88.6 (4.5)	2.1 (1.3)	86.8 (5.3)	88.3 (3.2)	6.3 (2.7)	82.7 (4.2)	80.6 (5.5)	7.5 (7.1)	74.6 (8.5)
Base Classifier e_5	93.1 (3.0)	4.1 (2.4)	89.3 (4.9)	89.5 (2.5)	8.4 (2.5)	82.1 (3.5)	93.2 (1.6)	6.4 (2.6)	87.3 (3.5)
Best Base Classifier e_6	96.6 (1.9)	3.2 (1.8)	93.6 (3.5)	91.4 (4.0)	5.6 (1.9)	86.3 (4.2)	91.2 (5.9)	7.1 (5.4)	84.6 (8.0)
Base Classifier e_7	92.0 (3.0)	4.7 (2.4)	87.8 (4.8)	89.8 (2.6)	8.7 (2.3)	82.0 (2.8)	91.4 (1.1)	9.4 (5.3)	82.9 (5.7)
Base Classifier e_8	95.8 (1.9)	3.6 (2.0)	92.5 (3.7)	90.8 (3.3)	6.2 (1.8)	85.2 (3.6)	90.2 (5.1)	8.8 (6.8)	82.2 (9.2)
Base Classifier e_9	93.4 (3.9)	7.4 (3.6)	86.7 (6.8)	88.4 (4.4)	15.5 (6.6)	74.7 (6.2)	90.2 (3.4)	11.1 (5.6)	80.1 (4.3)
Base Classifier e_{10}	94.2 (3.8)	7.1 (3.9)	87.7 (7.0)	86.2 (5.5)	17.1 (7.1)	71.7 (9.8)	88.1 (6.3)	13.4 (5.2)	76.4 (5.9)
Base Classifier e_{11}	92.4 (3.6)	8.6 (4.3)	84.7 (7.0)	88.7 (4.3)	15.5 (5.7)	74.9 (6.7)	92.1 (2.6)	17.4 (6.2)	76.1 (6.8)
Base Classifier e_{12}	93.7 (3.5)	8.4 (4.8)	86.1 (7.5)	86.5 (5.6)	16.4 (7.0)	72.4 (7.8)	90.4 (3.0)	21.5 (10.4)	71.0 (10.1)
Base Classifier e_{13}	86.8 (4.9)	4.8 (2.6)	82.8 (6.0)	85.1 (4.1)	12.2 (4.7)	74.8 (5.4)	86.0 (3.3)	7.9 (6.1)	79.2 (6.5)
Base Classifier e_{14}	88.2 (4.3)	5.6 (3.6)	83.4 (6.4)	83.9 (4.3)	13.3 (5.7)	72.7 (6.6)	83.2 (4.4)	9.5 (8.1)	75.6(10.8)
Base Classifier e_{15}	90.2 (3.8)	6.8 (3.5)	84.3 (6.4)	86.5 (4.0)	12.6 (5.3)	74.8 (5.4)	85.4 (3.6)	9.5 (6.2)	77.5 (7.7)
Base Classifier e_{16}	91.4 (3.2)	7.3 (4.4)	84.9 (6.6)	85.4 (3.2)	14.3 (5.4)	73.2 (6.3)	83.1 (7.1)	15.0 (12.6)	71.6 (16.2)
Average of Base Classifiers	91.0 (3.0)	5.0 (2.2)	86.5 (2.3)	87.2 (2.0)	10.5 (4.1)	78.2 (4.5)	87.2 (3.8)	9.9 (3.6)	78.4 (3.6)
AMVAFRD-6/8	95.8 (2.3)	2.2 (1.2)	93.8 (3.0)	93.5 (1.9)	4.7 (1.9)	89.1 (2.6)	93.0 (2.3)	3.9 (2.2)	89.4 (3.1)
AMVSFID-6/8	95.8 (2.3)	2.3 (1.2)	93.7 (3.1)	93.4 (1.7)	4.9 (1.8)	88.9 (2.6)	92.9 (2.3)	4.0 (2.0)	89.2 (2.9)
AMVAFRD-6/16	94.5 (3.9)	2.3 (1.3)	92.3 (4.5)	93.2 (1.8)	5.2 (2.1)	88.4 (2.8)	90.3 (1.8)	3.5 (2.1)	87.2 (2.7)
AMVSFID-6/16	94.4 (3.9)	2.5 (1.2)	92.1 (4.5)	93.1 (1.8)	5.5 (2.2)	88.0 (3.0)	90.0 (2.0)	3.8 (2.5)	86.6 (2.9)

ADMVAFR stands for adaptive diversity-based majority voting with average fixed rule hybrid classifier fusion scheme.
ADMVSFI stands for adaptive diversity-based majority voting with Sugeno fuzzy integral hybrid classifier fusion scheme.
6/8, 6/16 stands for selecting six base classifiers from the classifier pool containing 8 or 16 classifiers, respectively.

TABLE III
MEAN AND MAD OF THE DIFFERENCE BETWEEN CORRECT CLASSIFICATION RATE CC_r AND ERROR RATE E_r FOR THE DIFFERENT CLASSIFICATION APPROACHES ACROSS THE THREE EMG SIGNAL DATA SETS

Classifier	**Independent simulated signals**	**Related simulated signals**	**Real signals**
Weakest Base Classifier	77.7 (2.7)	56.0 (0.8)	49.5 (0.3)
Best Base Classifier	90.4 (1.7)	80.7 (2.3)	77.5 (2.6)
Average of Base Classifiers	81.5 (0.1)	67.7 (0.4)	68.5 (0.0)
ADMVAFR - 6/8	91.6 (1.8)	84.4 (0.7)	85.5 (0.9)
ADMVSFI - 6/8	91.2 (1.8)	84.0 (0.8)	85.2 (0.9)
ADMVAFR - 6/16	90.0 (3.2)	83.2 (0.7)	83.7 (0.6)
ADMVSFI - 6/16	89.6 (3.3)	82.5 (0.8)	82.8 (0.4)

and complexities were presented to each base classifier and an aggregator module combines their output decisions such that the performance in terms of signal characteristics was determined. Real EMG signals that were decomposed manually by a human expert were used as a reference to determine and evaluate the performance.

For comparison between classifiers performances, we used the definition of *better performance* stated in Section VIII-C. The classifier is with better performance, relative to another one, if it is with the highest correct classification rate $CC_r\%$ and lowest error rate $E_r\%$. In situations when there is no clear distinction, we take the difference between the correct classification rate $CC_r\%$ and error rate $E_r\%$ for each classifier and consider the classifier with the higher difference as the one with the better performance.

The diversity-based hybrid classifier fusion results, as shown in Tables II and III, provide performance that exceed the performance of any of the base classifiers forming the ensemble and also reduce the classification errors for all data sets relative to the base classifiers.

The improvement in classification performance obtained using the diversity-based hybrid classifier fusion, consisting of the majority voting and the average fixed rule, is demonstrated through presenting a schematic EMG signal decomposition summary and MUP trace for a test EMG signal and comparing with the performance of the best base classifier in the pool of base classifiers.

The test signal used is signal 3-2 from the related simulated EMG signals set. It has an intensity of 135 pps, ten MUPT classes, and was simulated to have a jitter value of 75 μs and an

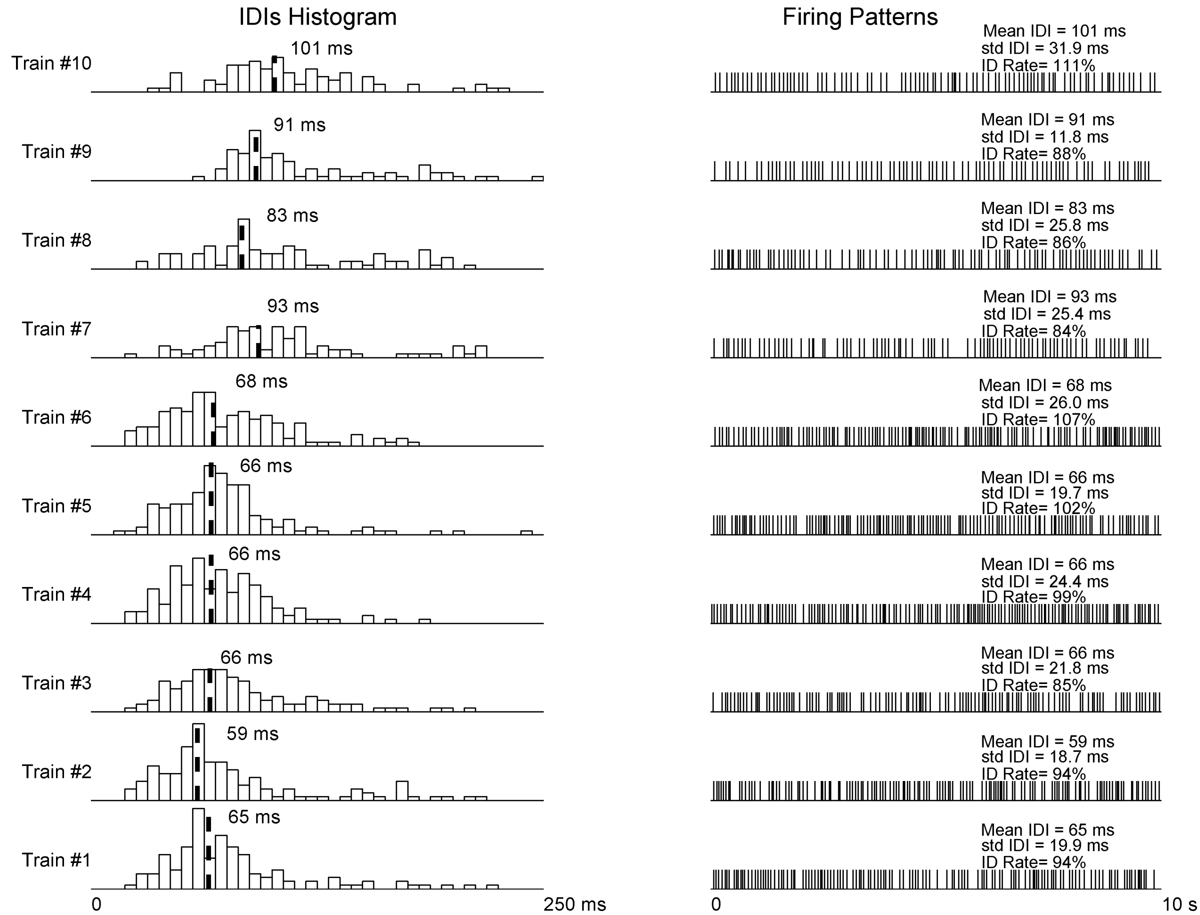


Fig. 5. Decomposition summary for the tested EMG signal 3-2 in terms of IDI histogram, MUP firing times, and statistics using diversity-based hybrid classifier fusion. Estimates of MUPT class IDI mean and standard deviation have been obtained using the EFE algorithm [31], [32].

IDI CV of 0.3. Table IV shows the performance of the base classifiers and the diversity-based hybrid classifier fusion schemes using the test signal.

Concerning the first experiment performed and reported in Section IX, from the eight base classifiers in the pool, the developed system chooses the ensemble $e_1, e_2, e_3, e_4, e_6, e_7$ to be fused at the first combiner that implement the majority voting scheme as this ensemble is the one with the best level of agreement having $\hat{\bar{\kappa}} = 0.81$. At the second combiner implementing the average fixed rule, the system chooses the ensemble $e_2, e_3, e_4, e_5, e_6, e_8$ as this ensemble is the one with the minimum level of agreement across the unassigned category having $\hat{\bar{\kappa}} = -0.04$.

Concerning the second experiment performed and reported in Section IX, from the 16 base classifiers in the pool, the developed system chooses the ensemble $e_1, e_2, e_3, e_4, e_6, e_8$ to be fused at the first combiner that implement the majority voting scheme as this ensemble is the one with the best level of agreement having $\hat{\bar{\kappa}} = 0.81$. At the second combiner implementing the average fixed rule, the system chooses the ensemble $e_2, e_3, e_5, e_6, e_7, e_{12}$ as this ensemble is the one with the minimum level of agreement across the unassigned category having $\hat{\bar{\kappa}} = -0.07$.

The complementary action of the base classifiers of the ensemble is shown in Table IV and Figs. 5 and 7. Table IV shows

TABLE IV
PERFORMANCE OF THE BASE CLASSIFIERS AND THE DIVERSITY-BASED HYBRID CLASSIFIER FUSION SCHEMES USING THE TEST EMG SIGNAL

Classifier	$A_r\%$	$E_r\%$	$CC_r\%$	$CC_r\%$ - $E_r\%$
Base Classifier e_1	83.4	7.2	77.4	70.2
Base Classifier e_2	85.7	7.6	79.2	71.6
Base Classifier e_3	88.2	8.6	80.6	72.0
Base Classifier e_4	88.3	8.7	80.6	71.9
Base Classifier e_5	91.9	11.5	81.3	69.8
Base Classifier e_6	94.0	8.7	85.9	77.2
Base Classifier e_7	93.0	11.2	82.6	71.4
Base Classifier e_8	87.7	8.0	80.6	72.6
Base Classifier e_9	86.6	27.3	63.0	35.7
Base Classifier e_{10}	82.8	34.6	54.2	19.6
Base Classifier e_{11}	85.6	21.4	67.3	45.9
Base Classifier e_{12}	75.0	22.8	57.9	35.1
Base Classifier e_{13}	83.4	18.7	67.8	49.1
Base Classifier e_{14}	87.1	25.1	65.2	40.1
Base Classifier e_{15}	90.3	23.1	69.4	46.3
Base Classifier e_{16}	91.1	22.8	70.4	47.6
ADMVAFR - 6/8	94.6	7.1	87.9	80.8
ADMVSFI - 6/8	94.4	6.9	87.8	80.9
ADMVAFR - 6/16	95.1	7.2	88.2	81.0
ADMVSFI - 6/16	94.1	7.1	87.4	80.3

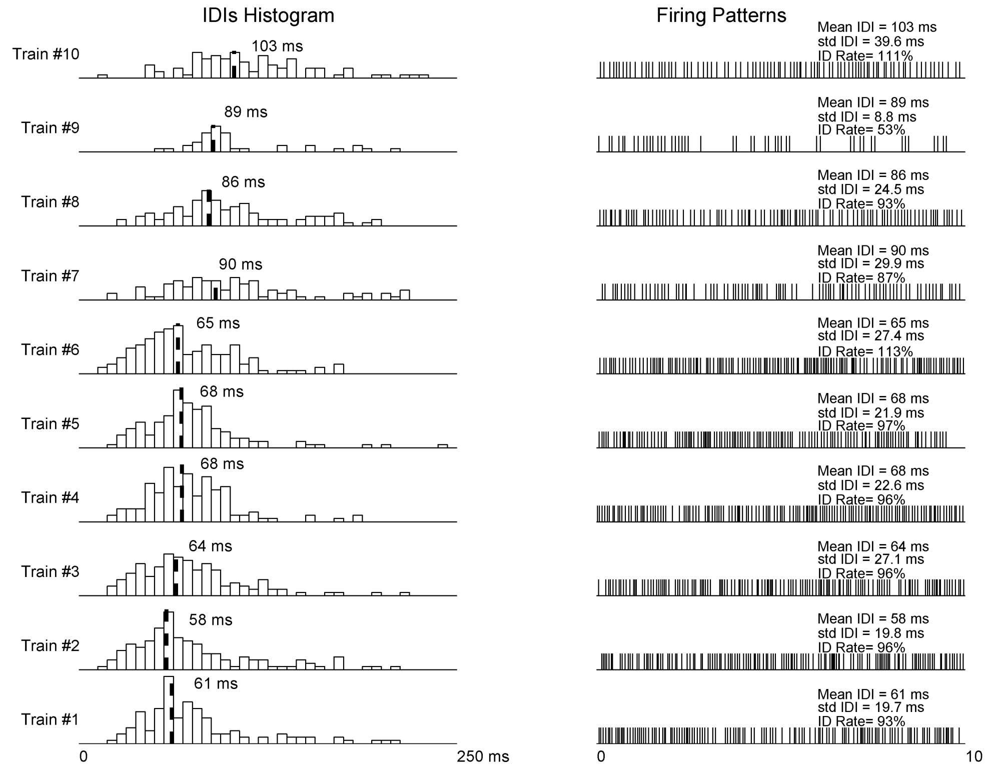


Fig. 6. Decomposition summary for the tested EMG signal 3-2 in terms of IDI histogram, MUP firing times, and statistics using the best base classifiers AFNNC e_6.

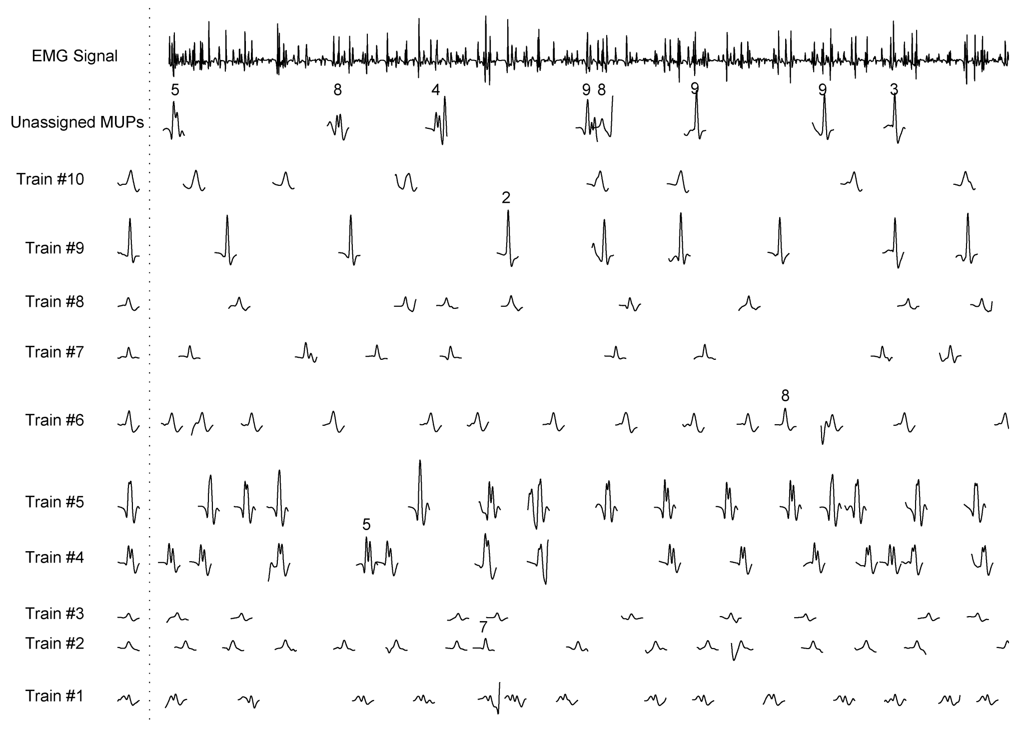


Fig. 7. MUP trace for a 1-s interval (from second 3 to second 4) of decomposition results for the tested EMG signal using the diversity-based hybrid classifier fusion, consisting of the majority voting and the average fixed rule. The MUP trace shows the MUP waveforms displayed according to their firing times within the 1-s signal segment. From top, row 1 shows the 1-s-long EMG signal segment; row 2 shows the MUP patterns that were left unassigned; the remaining rows show the MUP patterns assigned to the MUPT classes as recognized by the classifier. The left column shows the template waveforms for each MUPT class for visualization comparisons. Unassigned MUP patterns and erroneous MUP classification are indicated by displaying the number of the correct MUPT class next to the MUP. For improved visualization, MUP waveforms have been displayed using a time scale expanded by a factor of 10 relative to the time scale used to depict their firing times.

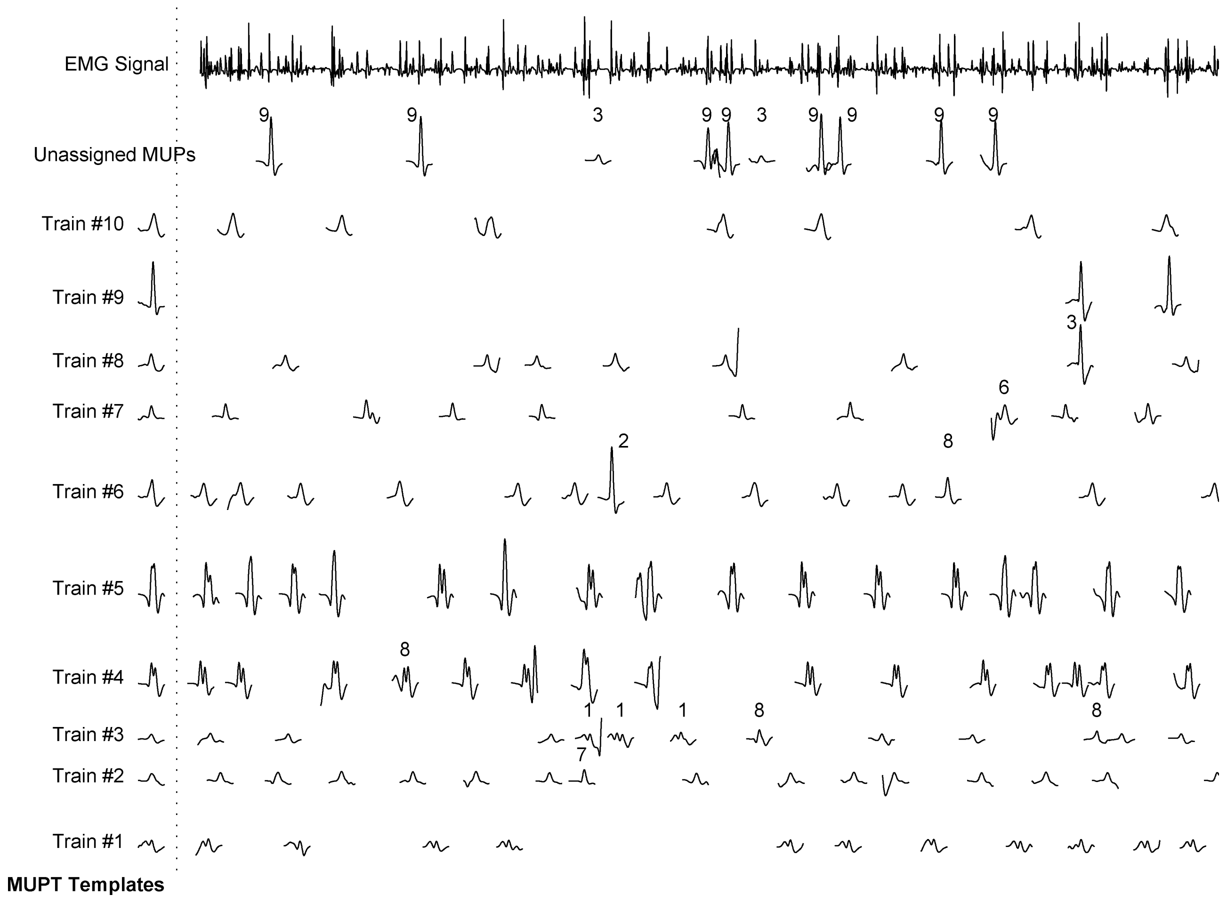


Fig. 8. MUP trace for a 1-s interval (from second 3 to second 4) decomposition results of the tested EMG signal using the best base classifier e_6.

from the first experiment that the diversity-based hybrid classifier fusion, consisting of the majority voting and the average fixed rule, gives a correct classification rate (CC_r) of 87.9% and an error rate (E_r) of 7.1% compared to the best base classifier e_6, which gives correct classification rate of 85.9% and an error rate of 8.7%. It also shows from the second experiment a correct classification rate of 88.2% and an error rate of 7.2%.

Fig. 5 presents decomposition results for signal 3-2 using the diversity-based hybrid classifier fusion, consisting of the majority voting and the average fixed rule. The effect of the MUP shape and MU firing pattern variability can be seen in the MUP trace, as shown in Fig. 7, and IDI histogram and MUP firing time plots, respectively. When comparing with decomposition results of the best base classifiers AFNNC e_6 shown in Fig. 6, we see that the diversity-based hybrid classifier fusion decomposition is better than that for the best base classifiers e_6 in terms of the ID rate for each MUPT class. Specifically, MUPT class 9 is now a nearly complete train, while the best base classifier recognizes a somewhat sparse train, where the ID rate for e_6 = 53% but for the hybrid classifier fusion is 88%. This demonstrates that the complementary action of the base classifiers of the ensemble is able to correct the errors.

Fig. 7 displays a MUP trace for a 1-s interval (from second 3 to second 4) of the decomposition results for the ten MUPT classes of the test signal and the unassigned MUP patterns as decomposed by the diversity-based hybrid classifier fusion system, consisting of the majority voting and the average fixed rule. When compared with the MUP trace of the best base classifier e_6, as shown in Fig. 8, we see that the diversity-based hybrid classifier fusion system during this epoch of time has a MUP trace with fewer MUP patterns errors as it erroneously assigned only four MUP patterns compared to ten MUP patterns in case of the best base classifier. Also, it rejected by passing to the unassigned category less MUP patterns, where it unassigned only eight MUP patterns compared to ten MUP patterns in case of the best base classifier.

XI. Conclusion

A diversity-based hybrid classifier fusion approach was proposed to overcome the limitation of selecting the base classifiers comprising an ensemble, where there is a need to perform an exhaustive search for the best accuracy classifier ensemble. The improvement in accuracy of a classifier fusion system does not depend only on the fusion method used for combining the base classifiers but also on the selection of classifiers used for the combination. Based on this, the hybrid classifier fusion approach exploits a diversity measure for designing classifier teams. We chose the kappa statistics measure for this purpose to estimate the level of agreement between the base classifier outputs, i.e., to measure the degree of decision similarity between the base classifiers.

The diversity-based hybrid classifier fusion approaches have been described and evaluated using simulated and real EMG signals and compared with the performance of the ensemble of constituent base classifiers. In these situations and across the used EMG signal datasets, the diversity-based hybrid approaches outperform the best base classifier.

When combining base classifiers having the best performance, it does not mean that the classifier fusion system will give the optimal performance. Also, if some of the base classifiers in the ensemble have weak classification performance, it does not mean that the weak ones when selected to work in a team with others will not improve the performance. This behavior was noted when the ensemble contains some MTFs base classifiers having weak performances relative to the certainty-based and the fuzzy k-NN classifiers.

Acknowledgment

The authors would like to thank the anonymous reviewers for providing the positive comments.

References

[1] J. M. Kittler, M. Hatef, R. P. V. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 955–966, Mar. 1998.

[2] G. Giacinto and F. Roli, "An approach to the automatic design of multiple classifier systems," *Pattern Recognit. Lett.*, vol. 22, pp. 25–33, 2001.

[3] S. Rasheed, D. Stashuk, and M. Kamel, "A hybrid classifier fusion approach for motor unit potential classification during EMG signal decomposition," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 9, pp. 1715–1721, Sep. 2007.

[4] S. Rasheed, "A multiclassifier approach to motor unit potential classification for EMG signal decomposition," Ph.D. dissertation, Dept. Syst. Des. Eng., Univ. Waterloo, Waterloo, ON, Canada 2006 [Online]. Available: http://etd.uwaterloo.ca/etd/srasheed2006.pdf

[5] D. W. Stashuk, "Decomposition and quantitative analysis of clinical electromyographic signals," *Med. Eng. Phys.*, vol. 21, pp. 389–404, 1999.

[6] G. M. Paoli, "Estimating certainty in classification of motor unit action potentials" Master's thesis, Dept. Syst. Des. Eng., Univ. Waterloo, Waterloo, ON, Canada, 1993.

[7] D. W. Stashuk, "EMG signal decomposition: How can it be accomplished and used?," *J. Electromyogr. Kinesiol.*, vol. 11, pp. 151–173, 2001.

[8] D. W. Stashuk and G. M. Paoli, "Robust supervised classification of motor unit action potentials," *Med. Biol. Eng. Comput.*, vol. 36, no. 1, pp. 75–82, 1998.

[9] S. Rasheed, D. Stashuk, and M. Kamel, "Adaptive certainty-based classification for decomposition of EMG signals," *Med. Biol. Eng. Comput.*, vol. 44, no. 4, pp. 298–310, 2006.

[10] J. M. Keller, M. R. Gray, and J. A. Givens, Jr., "A fuzzy k-nearest neighbor algorithm," *IEEE Trans. Syst., Man, Cybern.*, vol. 15, no. 4, pp. 580–585, Jul./Aug. 1985.

[11] S. Rasheed, D. Stashuk, and M. Kamel, "A software package for interactive motor unit potential classification using fuzzy k-NN classifier," *Comput. Methods Programs Biomed.*, vol. 89, no. 1, pp. 56–71, 2008.

[12] S. Rasheed, D. Stashuk, and M. Kamel, "Adaptive fuzzy k-NN classifier for EMG signal decomposition," *Med. Eng. Phys.*, vol. 28, no. 7, pp. 694–709, 2006.

[13] L. D. Stefano and S. Mattoccia, "Fast template matching using bounded partial correlation," *Mach. Vis. Appl.*, vol. 13, pp. 213–221, 2003.

[14] J. R. Florestal, P. A. Mathieu, and A. Malanda, "Automatic decomposition of simulated EMG signals," in *Proc. 28th Conf. Can. Medican Biol. Eng. Soc.*, 2004, pp. 29–30.

[15] J. R. Florestal, P. A. Mathieu, and A. Malanda, "Automated decomposition of intramuscular electromyographic signals," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 5, pp. 832–839, May 2006.

[16] D. Partridge and W. B. Yates, "Engineering multiversion neural-net systems," *Neural Comput.*, vol. 8, pp. 869–893, 1996.

[17] A. J. C. Sharkey, N. E. Sharkey, U. Gerecke, and G. O. Chandroth, "The test and select approach to ensemble combination," in *Multiple Classifier Systems* (Lecture Notes in Computer Science), vol. 1857, J. Kittler and F. Roli, Eds. Gagliari, Italy: Springer-Verlag, 2000, pp. 30–44.

[18] F. Roli and G. Giacinto, "Design of multiple classifier systems," in *Hybrid Methods in Pattern Recognition*, H. Bunke and A. Kandel, Eds. Cambridge, MA: World Scientific, 2002, pp. 199–226.

[19] F. Roli, G. Giacinto, and G. Vernazza, "Methods for designing multiple classifier systems," in *Proceedings of the Second International Workshop on Multiple Classifier Systems* (Lecture Notes in Computer Science), vol. 2096, J. Kittler and F. Roli, Eds. New York: Springer-Verlag, 2001, pp. 78–87.

[20] J. L. Fleiss, B. Levin, and M. C. Paik, *Statistical Methods for Rates and Proportions*, 3rd ed. New York: Wiley, 2003.

[21] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 2, pp. 37–46, 1960.

[22] L. Lam and C. Y. Suen, "Application of majority voting to pattern recognition: An analysis of its behavior and performance," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 27, no. 5, pp. 553–568, Sep. 1997.

[23] A. Verikas, A. Lipnickas, K. Malmqvist, M. Bacauskiene, and A. Gelzinis, "Soft combination of neural classifiers: A comparative study," *Pattern Recognit. Lett.*, vol. 20, pp. 429–444, 1999.

[24] S. Hashem, "Optimal linear combinations of neural networks," *Neural Netw.*, vol. 10, no. 4, pp. 599–614, 1997.

[25] J. M. Keller, P. Gader, H. Tahani, J.-H. Chiang, and M. Mohamed, "Advances in fuzzy integration for pattern recognition," *Fuzzy Sets Syst.*, vol. 65, pp. 273–283, 1994.

[26] M. Sugeno, "Fuzzy measures and fuzzy integrals—A survey," in *Fuzzy Automata Decision Processes*. Amsterdam, The Netherlands: North Holland, 1977, pp. 89–102.

[27] R. P. W. Duin, "The combining classifier: To train or not to train?," in *Proc. 16th Int. Conf. Pattern Recognit.*, 2002, vol. 2, pp. 765–770.

[28] S. Rasheed, D. Stashuk, and M. Kamel, "Multi-classification techniques applied to EMG signal decomposition," in *Proc. 2004 IEEE Int. Conf. Syst., Man Cybern. (SMC 2004)*, The Hugue, The Netherland, vol. 2, pp. 1226–1231.

[29] H. Etawil and D. Stashuk, "Resolving superimposed motor unit action potentials," *Med. Biol. Eng. Comput.*, vol. 34, no. 1, pp. 33–40, 1996.

[30] G.-H. Chiang, "Choquet fuzzy integral-based hierarchical networks for decision analysis," *IEEE Trans. Fuzzy Syst.*, vol. 7, no. 1, pp. 63–71, Feb. 1999.

[31] Y. Qu, "Automatic clustering of motor unit action potentials using shape and time information," Master's thesis, Dept. Syst. Des. Eng., Univ. Waterloo, Waterloo, ON, Canada, 1993.

[32] D. W. Stashuk and Y. Qu, "Robust method for estimating motor unit firing-pattern statistics," *Med. Biol. Eng. Comput.*, vol. 34, no. 1, pp. 50–57, 1996.

[33] D. Farina, R. A. Colombo, R. Merletti, and H. B. Olsen, "Evaluation of intra-muscular EMG signal decomposition algorithms," *J. Electromyogr. Kinesiol.*, vol. 11, pp. 175–187, 2001.

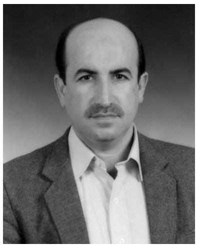
[34] D. Farina, A. Crosetti, and R. Merletti, "A model for the generation of synthetic intramuscular EMG signals to test decomposition algorithms," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 1, pp. 66–77, Jan. 2001.

[35] C. J. De Luca, "Reflections on EMG signal decomposition," in *Computer-Aided Electromyography and Expert Systems*, J. E. Desmedt, Ed. Amesterdam, The Netherlands: Elsevier, 1989, pp. 33–37.

[36] K. C. McGill, "A method for quantitating the clinical electromyogram," Ph.D. dissertation, Dept. Electr. Eng., Stanford Univ., Stanford, CA, 1984.

[37] S. Usaui and I. Amidror, "Digital low-pass differentiation for biological signal processing," *IEEE Trans. Biomed. Eng.*, vol. BME-29, no. 10, pp. 686–693, Oct. 1982.

[38] A. Hamilton-Wright and D. W. Stashuk, "Physiologically based simulation of clinical EMG signals," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 2, pp. 171–183, Feb. 2005.



Sarbast Rasheed (M'96) received the B.Sc. degree in electrical engineering and the M.Sc. degree in computer and control engineering from the University of Baghdad, Baghdad, Iraq, in 1981 and 1985, respectively, and the Ph.D. degree in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 2006.

From 1987 to 1989, he was an R&D Engineer in the Iraqi Electronics Industries Company. In September 1989, he joined as a Lecturer Member with the Department of Computer and Control Engineering, University of Baghdad, and then with the Department of Computer Science, Applied Science University, Amman, Jordan until 1999. From 2006 to 2007, he was a Postdoctoral Fellow in the area of microelectromechanical systems (MEMS) modeling. From 2007 to 2009, he was a Senior Software Specialist in the visual inspection area at the German company Erhardt and Leimer Ltd. and then a Vision System Engineering Specialist at the Canadian company Rotoflex International, Inc., Mississauga, ON, Canada.

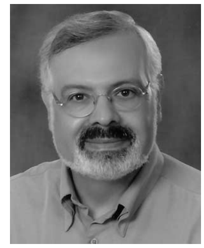
Since July 2009, he has been a Researcher in the Department of Systems Design Engineering, University of Waterloo, where he is working in the area of Medical Imaging. His current research interests include magnetic resonance imaging (MRI) and 3-D optical coherence tomography (OCT).

Dr. Rasheed is a member of the Engineering in Medicine and Biology Society of the IEEE and also is a member of the International Society for Magnetic Resonance in Medicine (ISMRM).

Daniel W. Stashuk (M'89) received the B.A.Sc. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 1978, and the M.Eng. and Ph.D. degrees in electrical engineering from McMaster University, Hamilton, ON, in 1982 and 1985, respectively.

From 1985 to 1988, he was a Research Assistant Professor at Boston University Neuromuscular Research Center, where he was a supervisor of the Motor Unit Laboratory. Since 1988, he has been a faculty member in the Department of Systems Design Engineering, University of Waterloo, where he is currently a Professor. His current research interests include detection and analysis of biological signals, especially the detection and analysis of electromyographic signals for the extraction of clinically relevant information.

Prof. Stashuk is a member of the Canadian Medical and Biological Society and the Engineering in Medicine and Biology Society of the IEEE.



Mohamed S. Kamel (S'74–M'80–SM'96–F'05) received the B.Sc. (Hons.) degree in electrical engineering from Alexandria University, Alexandria, Egypt, the M.A.Sc degree from McMaster University, Hamilton, ON, Canada, and the Ph.D. degree from the University of Toronto, Toronto, ON.

In 1985, he joined the University of Waterloo, Waterloo, ON, Canada, where he is currently a Professor and the Director of the Pattern Analysis and Machine Intelligence Laboratory, Department of Electrical and Computer Engineering, and holds the University Research Chair. He also holds the Canada Research Chair in Cooperative Intelligent Systems. He is a cofounder of Virtek Vision, Inc., Waterloo, where he is the Chair of the Technology Advisory Group. He was a Consultant for General Motors, NCR, International Business Machines Corporation (IBM), Northern Telecom, and Spar Aerospace. His current research interests include computational intelligence, pattern recognition, machine learning, and cooperative intelligent systems. He has authored or coauthored over 390 papers in journals and conference proceedings, 11 edited volumes, two patents, and numerous technical and industrial project reports. He is the Editor-in-Chief of the *International Journal of Robotics and Automation*. He is an Associate Editor of the *Pattern Recognition Letters*, *Cognitive Neurodynamics*, and *Pattern Recognition*. He is also a member of the Editorial Advisory Board of the *International Journal of Image and Graphics* and *Intelligent Automation and Soft Computing*. He was an Associate Editor of the *Simulation* (the journal of the Society for Computer Simulation).

Prof. Kamel received the NCR Inventor Award. He was the recipient of the Systems Research Foundation Award for outstanding presentation in 1985 and the International Symposium on Robotics and Manufacturing (ISRAM) Best Paper Award in 1992. In 1994, he was awarded the IEEE Computer Society Press Outstanding Referee Award. He was also a coauthor of the Best Paper in the 2000 IEEE Canadian Conference on electrical and computer engineering. He was also the recipient of the University of Waterloo Outstanding Performance Award twice and the Faculty of Engineering Distinguished Performance Award. He is a member of the Association for Computing Machinery (ACM) and the Professional Engineers Ontario (PEO). He is a Fellow of the Engineering Institute of Canada (EIC), the Canadian Academy of Engineering (CAE), and selected to be a Fellow of the International Association of Pattern Recognition (IAPR) in 2008. He was a member of the board from 1992 to 2008 and the Vice President (VP) of research and development from 1987 to 1992. He is an Associate Editor of the IEEE SYSTEMS, MAN, AND CYBERNETICS (SMC), PART A.