Ensemble Learning from Imbalanced Data Set for Video Event Detection

Yimin Yang, Shu-Ching Chen
School of Computing and Information Sciences
Florida International University
Miami, FL 33199, USA
Email: {yyang010,chens}@cs.fiu.edu

***Abstract*—Learning from imbalanced data sets is a hot and challenging research topic with many real world applications. Many studies have been conducted on integrating sampling-based techniques and ensemble learning for imbalanced data sets. However, most existing sampling methods suffer from the problems of information loss, over-fitting, and additional bias. Moreover, there is no single model that can be applied to all scenarios. Therefore, a positive enhanced ensemble learning (PEEL) framework is presented in this paper for effective video event detection. The proposed PEEL framework involves a novel sampling technique combined with an ensemble learning mechanism built upon the base learning algorithm (BLA). Exploratory experiments have been conducted to evaluate the related parameters and performance comparisons. The experimental results demonstrate the effectiveness of the proposed PEEL framework for video event detection.**

***Keywords*-Imbalanced data set; sampling; ensemble learning; multiple correspondence analysis (MCA); video event detection**

I. Introduction

Learning from imbalanced data sets for binary classification problems has been a hot and challenging topic in the research communities and has many real-world applications such as fraud detection, medical diagnosis, intrusion detection, face recognition, information retrieval, and video event detection [1][2][3][4][5][6][7][8]. The class imbalance problem has been amplified and aggravated as the world steps into the big data era. The underlying nature of the class imbalance issue is that the number of samples (instances) in the majority (negative) class dramatically exceeds that of the minority (positive) class of interest, which undermines the classification process. For example, the positive to negative ratio is about 1:100 and 5:1000 for fraud detection and video event detection [2][8], respectively. Many attempts have been made to address the class imbalance problems in different occasions [9]. However, there is no single method that triumphs in all scenarios. In this paper, the focus is on addressing the class imbalance challenge for video event detection.

A video event is defined as an activity of a particular user interest, for example, a goal event in a soccer video. The rareness of a video event (positive instance) makes the detection task extremely difficult because of the aforementioned class imbalance issue [10][11][12][13]. By further analyzing the problem, it is found that most of the false alarms (or false positives) are pretty close to the real events in a certain sense (e.g., goal attempt and foul) which might also attracts users' interests. A good video event detection framework should retrieve as many true positive instances as possible, although it might potentially include more false positive instances. In other words, the video event detection learner should enhance the favor of the positive class. With this objective in mind, a positive-enhanced ensemble learning (PEEL) framework is presented for video event detection. The proposed framework integrates the sampling-based technique and the ensemble learning mechanism, and is able to detect most of the real events at the expense of including a small amount of related events. The proposed framework outperforms most of the well-known single models and ensemble classifiers under the Receiver Operating Characteristic (ROC) or the Area Under the Curve (AUC) criterion [14].

The paper is organized as follows. Section II provides an overview of the existing work for solving the class imbalance issue. Section III discusses the details of the proposed PEEL framework. Section IV presents a thorough experimental analysis. Finally, section V gives the conclusion.

II. Related Work

A considerable amount of efforts have been done in the research community on learning from the imbalanced data sets, especially for binary classification problems. He et al. [9] presented an overview of those methods and generally grouped them into three categories, namely (1) sampling-based methods [15][16], (2) cost-sensitive methods [17][18], and (3) kernel-based and active learning methods [19][20]. Among those approaches, the sampling-based methods and the integration with ensemble learning ones have been widely studied and have shown their success over the years [21]. Therefore, they will be the focus of this paper. Studies have demonstrated that a balanced data set usually outperforms an imbalanced one, which justifies the use of various sampling methods [22], such as random under-sampling and over-

978-1-4673-6656-4/15 $31.00 © 2015 IEEE
DOI 10.1109/IRI.2015.23

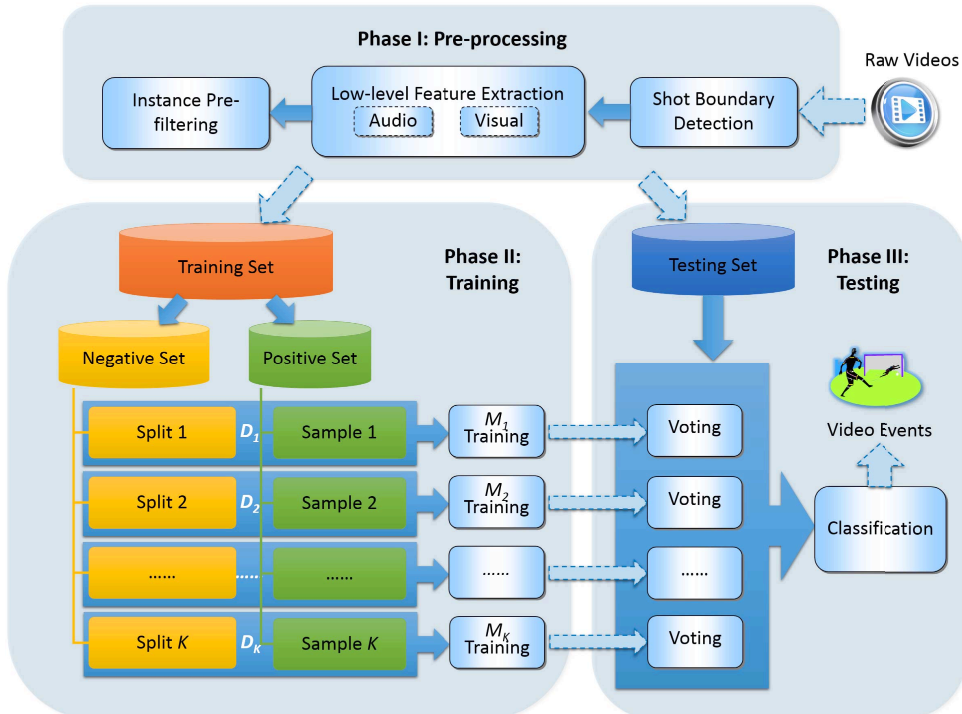

Figure 1: The proposed PEEL framework.

sampling [23][24], informed under-sampling [21][25], synthetic over-sampling [16][26][27], and clustering-based sampling [16][24][28]. The mechanics behind under-sampling and over-sampling are the random removal of the majority instances and the replication of minority instances respectively [9]. Both ways have their intrinsic problems such as the loss of majority information and over-fitting [29]. The under-sampling approaches alleviate those problems by using some statistical knowledge [21]. More recently, the clustering-based sampling methods have been proved effective by dealing with both within-class and between-class imbalance issues. For example, in [16], Barua et al. proposed the so-called Majority Weighted Minority Oversampling TEchnique (MW-MOTE), which generates the synthetic samples from the weighted minority class using a clustering approach. Although the synthetic oversampling methods provide a better balance in the distribution between the majority and minority classes, they unavoidably introduce error-prone instances [16].

To overcome the limitation of the sampling-based methods, the integration of ensemble learning mechanism (such as bagging [30] and boosting [31]) is introduced. For example, Chawla et al. [32] integrated SMOTE [26] with Adaboost [31] for boosting the performance of the minority class. In [33], Guo et al. combined the synthetic data generation technique [34] and the Adaboost algorithm [31] to improve the overall accuracy. Although the "sampling-ensemble" methods have been proved to be efficient and effective, there is no single approach that can be applied to all scenarios.

III. Ensemble Learning Framework

As illustrated in Figure 1, the proposed PEEL framework contains three phases, i.e., pre-processing, training, and testing. In phase I, the input raw videos are pre-processed to generate a pre-filtered candidate instance set with the extracted features. In phase II, the proposed PEEL framework is applied to obtain an ensemble of the base learners. Finally, in phase III, the ensemble learner is applied to classify the target video event. The details of each of the three phases are discussed in the following subsections.

A. Pre-processing

The pre-processing phase of the proposed framework consists of three main steps: shot boundary detection, low-level feature extraction, and instance pre-filtering. Usually, a video shot is treated as the basic unit for video event detection. Therefore, the first step of pre-processing is shot boundary detection, which provides the shot boundaries for video feature extraction. In this paper, the unsupervised multi-filtering method proposed in [35] is adopted for effective shot boundary detection. Due to the prevalence

Algorithm 1 Positive Enhanced Ensemble Learning Algorithm

Input: Training set Tr, BLA, positive ratio r, voting confidence $v \in [0,1]$.
Output: Ensemble learner $C(x)$.

1: **procedure** PEEL(Tr) ▷ training phase
2: $M \leftarrow \varnothing$;
3: separate Tr into positive set P and negative set Q;
4: $N_P \leftarrow |P|$; $N_Q \leftarrow |Q|$; ▷ obtain the sizes (numbers of instances) of P and Q respectively
5: $n_q \leftarrow N_P * r$; ▷ determine split size based on the given positive ratio
6: $K \leftarrow N_Q \ / \ n_q$; ▷ calculate the number of split for Q
7: evenly split Q into K subsets, denoted as $S = \{S_j \mid j = 1, \cdots, K\}$;
8: **for all** $j = 1, \cdots, K$ **do**
9: **if** $r >= 1$ **then**
10: $D_j \leftarrow S_j \cup P$; ▷ perform merge
11: **else if** $r < 1$ **then** ▷ i.i.d. sample with replacement
12: $D_j \leftarrow$ randomly sample n_q instances from P and merge with S_j;
13: **end if**
14: train model M_j based on D_j using BLA; $M \leftarrow M_j$;
15: **end for**
16: **return** the hypothesis:
17: $$C(x) = \begin{cases} 1 & \text{if } \sum_{j=1}^{K} M_j(x) > K * v, \ M_j(x) \in \{0,1\}; \\ 0 & \text{othersise} \end{cases}$$
18: **end procedure**

and effectiveness of multi-modal features for video content analysis, a set of visual and audio features are extracted for each video shot, which cover both low-level characteristics (such as pixel change) and mid-level semantics (such as grass ratio and audience volume) [36]. After feature extraction, the video data set is ready for event detection. However, the data set is highly imbalanced with a large number of irrelevant instances. As reported in [8], the interesting events (such as goal, goal attempt, and foul) only count less than 1% in the whole data set, not to mention only the goal event. As the first attempt to relieve the class imbalance issue to some extent, a pre-filtering step is performed to remove as many irrelevant instances as possible. For more details about the pre-processing process, please refer to [8].

B. Positive Enhanced Ensemble Learning

As aforementioned, most of the existing sampling algorithms (e.g., random under/over-sampling and synthetic sampling) suffer from the problems of information loss, over-fitting, and the introduction of bias. To overcome these limitations, we propose a novel sampling method which makes the full usage of all the positive and negative instances in the training set and builds an ensemble learner based on the base learning algorithm (BLA, as presented in section III-C). As shown in Algorithm 1, the proposed PEEL framework first separates the given training set Tr into the positive set P and negative set Q. Then Q is evenly split into K subsets ($S_j, \ j = 1, \cdots, K$) based on the given positive ratio r (lines 4 to 7), which represents the percentage of positive instances used in each batch ($D_j, \ j = 1, \cdots, K$) for the base model training (lines 8 to 15). When $r >= 1$ (case 1), all positive instances will be used for training in each batch with the number of negative instances increased as r goes up; otherwise, when $r < 1$ (case 2), the positive instances will be randomly sampled with replacement (assuming independent identical distribution, i.i.d.) based on the calculated n_q (line 5). Therefore, the numbers of positive and negative instances are identical for each batch in this case. In either cases, all of the negative instances in Tr will participate in the training process. When the value of r is relatively small ($<= 1$), the positive class will dominate the characteristic of each batch data set due to the superior inter-class coherency compared with the negative class, hence the name PEEL. After each base model ($M_j, \ j = 1, \cdots, K$) is properly trained, the final ensemble learner (hypothesis) is built based on the equation in line 17. As can be inferred from Algorithm 1, there are two critical parameters in this algorithm, i.e., the positive ratio r and the voting confidence v. While r decides the dominant level of the positive class in each base model, v reflects the confidence level for each model. The higher the value, the larger the number of positive outcomes is required from the base models for classifying an instance x as positive for $C(x)$. The selection and evaluation of r and v will be presented in the experimental section.

C. Base Learning Algorithm

The BLA is constructed based on a set of weak learners ($L = \{L_h \mid h = 1, \cdots, H\}$) as shown in Algorithm 2. The output of each weak learner is linearly combined using the given weight vector $w = \{w_h \mid h = 1, \cdots, H\}$, where each element represents the confidence for the corresponding weak learner. The combined results will be used to determine the final outcome of the base learner $B(x)$ as depicted in the equation in line 6. Theoretically, a "stronger" classifier should be assigned a larger weight. If all the weak learners are with equal weights, then the base learner reduces to a majority voting algorithm. The combination of BLA and PEEL framework has an "ensemble of ensemble" flavor. Considering the small sample size of each training batch, the computation overhead of the overall PEEL framework is negligible, compared with the performance gain. The construction of BLA will be analyzed in section IV-B.

Algorithm 2 Base Learning Algorithm

Input: Training set Tr', weak learners $L = \{L_h \mid h = 1, \cdots, H\}$, weight vector $w = \{w_h \mid h = 1, \cdots, H\}$, s.t. $\sum_{h=1}^{H} w_h = 1$.
Output: Base learner $B(x)$.

1: **procedure** BLA(Tr')
2:     **for all** $h = 1, \cdots, H$ **do**
3:         `train model` L_h `on` Tr'`;`
4:     **end for**
5:     **return** `the hypothesis:`
6: $$B(x) = \begin{cases} 1 & \texttt{if } \sum_{h=1}^{H} L_h(x) * w_h > 1/2; \\ 0 & \texttt{othersise} \end{cases}$$
7: **end procedure**

IV. Experimental Analysis

The proposed framework was extensively tested upon a large data set, which contains 58 soccer videos collected from the FIFA World Cup of 2003, 2010, and 2014. The total number of frames is over 4.7 millions and the total duration of the videos is about 52 hours. Among the total 32k video shots, only 105 of them contain the goal event, which contributes less than 0.5% to the total number of shots. A summary of the data set is shown in Table I.

Table I: Data set summary.

No. Files	No. Frames	Total Time	No. Shots	No. Goal Events
58	4,731,807	51 hours 48 min.	32,463	105

A. Evaluation Criteria

The receiver operating characteristic (ROC) curve is chosen as the evaluation method (under stratified cross-validation scheme) over the precision recall (PR) curve since we care more about the true positive rate (recall) than the precision [14]. In other words, a low precision is more tolerable than a low recall. This is because some false positives are also of user interests, especially in the video event detection scenario as mentioned before. Therefore, when determining the threshold for classification, we tend to retrieve a high true positive rate (or low false negative rate) and reduce the impact of negative data on the total classification costs. Table II shows the definition of the confusion matrix (CM) and Equation 1 presents the basic metrics for the analysis.

Table II: Confusion Matrix

CM	Predicted positive	Predicted negative
Actual positive	*TP*	*FN*
Actual negative	*FP*	*TN*

$$\text{True Positive Rate (TPR)} = \frac{TP}{TP+FN} \quad (1)$$

$$\text{False Positive Rate (FPR)} = \frac{FP}{FP+TN} \quad (2)$$

$$\text{False Negative Rate (FNR)} = \frac{FN}{FN+TP} \quad (3)$$

B. Selection of Weak Learners for BLA

The multiple correspondence analysis (MCA) approach [37][38] has found its success in various video analysis tasks, especially the interesting event detection problem [8][39]. In this paper, it is combined with the traditional decision tree (DT) algorithm [40] for constructing the BLA, since DT is usually used as a weak learner in the ensemble learning mechanism and it has been proved effective for goal event detection [36][41]. In our experiment, MCA and DT are assigned with equal weights. MCA is a continuous classifier which outputs probability-like ranking scores for the testing instances. Thus, the selection of a proper threshold for binary classification greatly affects the performance of MCA. To evaluate the impact of the threshold for MCA, the ROC curve is plotted in Figure 2 using a subset of the training data set. As can be seen from the figure, the MCA algorithm has a satisfactory performance for video event detection with an AUC value of 0.918. The AUC of the Conv Hull (shorted for convex hull) illustrates the theoretical maximum performance of the target algorithm for the corresponding evaluation data set. For the comparison purpose, the performance of the DT algorithm (as a discrete classifier with the binary

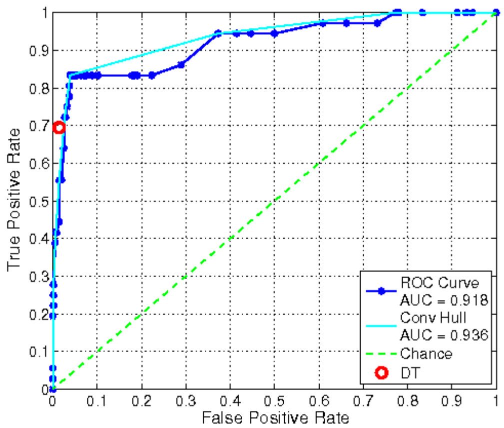

Figure 2: MCA ROC curve.

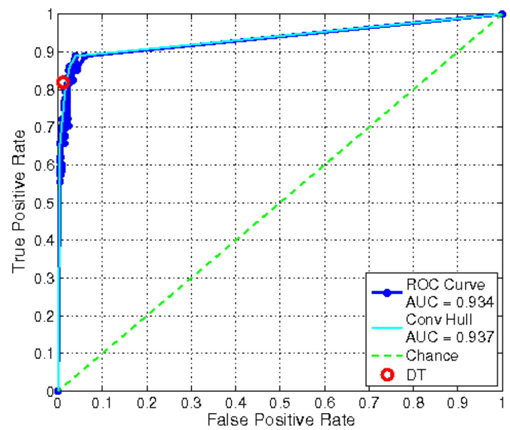

Figure 3: ROC curve on positive ratio (r).

output) on the same testing set is also depicted in the figure (as a red circle), where the green dotted line represents a random (by chance) classifier. As can be inferred from the figure, the MCA has over 10% gain of TPR over the DT in the ideal situation. The optimal threshold is obtained by minimizing the average expected cost of classification at point (y,z) in the ROC space as follows.

$$Cost(y,z) = (1-p) * \alpha * y + p * \beta * (1-z) \tag{4}$$

where α and β are the penalties of a false positive and a false negative respectively, and p is the positive portion calculated as

$$p = \frac{N_P}{N_P + N_Q} \tag{5}$$

where N_P and N_Q are the numbers of positives and negatives in the training as illustrated in Algorithm 1. In our scenario, α and β are assigned with the values of 0.2 and 0.8 respectively in order to emphasize the importance of TPR.

C. Analysis of Positive Ratio

To evaluate the performance and impact of the positive ratio r, the ROC curve over r is plotted with a fixed value of v (=0.5) as shown in Figure 3. There are two main observations and conclusions from the figure. First, the PEEL framework outperforms the individual weak learner (i.e., DT) in the sense of TPR by about 10% while maintaining comparable FPR. Second, the performance of PEEL boosted rapidly with relatively low FPR. Based on our experimental analysis, the PEEL framework achieves the best performance when the value of r is around 1.0, which means the positives and negatives are comparable. In other words, the training set is relatively balanced for each batch.

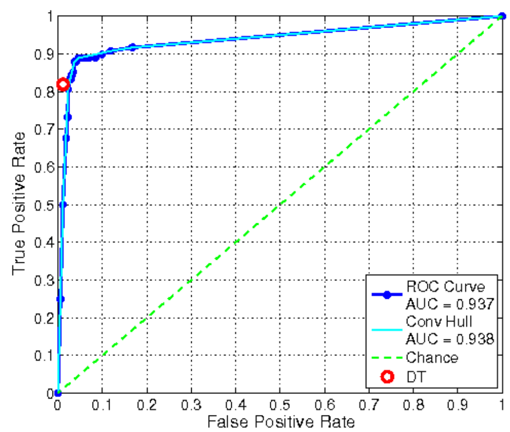

Figure 4: ROC curve on voting confidence (v).

D. Analysis of Voting Confidence

The ROC curve over the voting confidence v for the proposed PEEL framework is shown in Figure 4 with $r = 0.8$. As can be seen from the figure, Figure 4 is similar to Figure 3. The AUC (=0.937) is slightly better than in Figure 3 (with AUC=0.934), which means v has a relatively higher impact than $r$on the performance of PEEL. It is also observed that FPR degrades relatively faster with varying v values than r values. Based on the experimental results, the best performance is achieved when v is about 0.5, which is equivalent to majority voting among the base learners (M_j).

E. Comparison with Other Methods

Finally, the proposed PEEL framework is compared with various traditional single models (e.g., KNN, SVM, Naive Bayes, and DT) and other ensemble learners (e.g., Adboost, Bagging, and RandomForest).

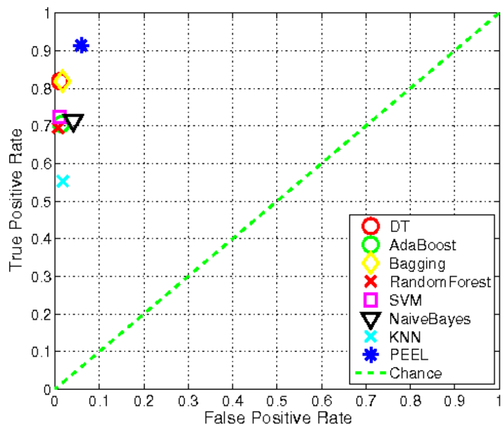

Figure 5: Comparison on various methods.

All the comparison methods (treated as discrete classifiers) are based on the implementation of WEKA [42] with the default parameter settings. As can be seen from Figure 5, our PEEL framework outperforms all the other methods with over 90% of TPR and comparable FPR. To be specific, it achieves about 10% TPR gain over the DT and Bagging algorithms; 20% TPR gain over the SVM, NaiveBayes, RandomForest, and Adaboost algorithms; and finally almost 40% TPR gain over the KNN algorithm.

V. Conclusion

In this paper, an effective ensemble learning algorithm called PEEL is proposed for video event detection. The PEEL framework contains a novel sampling method which makes the full use of all negative instances while enhancing the impact of the positive class for base learner training in the ensemble mechanism. The experimental analysis demonstrates the effectiveness of the proposed PEEL framework. In the future, more data sets and additional measurements should be applied to further evaluate the framework. Moreover, the within-class distribution should also be explored to develop better sampling mechanisms. In addition, it has great significance to study the optimization strategies for critical parameter estimation. Finally, it becomes gradually important to introduce big data analytics and technologies to accommodate ever-growing data sets.

Acknowledgment

This research is partially supported by DHS under grant Award Number 2010-ST-062-000039, DHS's VACCINE Center under Award Number 2009-ST-061-CI0001, NSF HRD-0833093, CNS-1126619, and NSF CNS-1461926.

References

[1] C. Chen, Q. Zhu, L. Lin, and M.-L. Shyu, "Web media semantic concept retrieval via tag removal and model fusion," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 4, no. 4, p. 61, 2013.

[2] T. Fawcett and F. Provost, "Adaptive fraud detection," *Data Mining and Knowledge Discovery*, vol. 1, no. 3, pp. 291–316, 1997.

[3] T. Meng, A. T. Soliman, M.-L. Shyu, Y. Yang, S.-C. Chen, S. Iyengar, J. S. Yordy, and P. Iyengar, "Wavelet analysis in current cancer genome research: A survey," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 10, no. 6, pp. 1442–14 359, 2013.

[4] L. Peng, Y. Yang, X. Qi, and H. Wang, "Highly accurate video object identification utilizing hint information," in *IEEE International Conference on Computing, Networking and Communications (ICNC)*, 2014, pp. 317–321.

[5] M.-L. Shyu, T. Quirino, Z. Xie, S.-C. Chen, and L. Chang, "Network intrusion detection through adaptive sub-eigenspace modeling in multiagent systems," *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, vol. 2, no. 3, p. 9, 2007.

[6] M.-L. Shyu, S.-C. Chen, M. Chen, C. Zhang, and K. Sarinnapakorn, "Image database retrieval utilizing affinity relationships," in *Proceedings of the 1st ACM international workshop on Multimedia databases*, 2003, pp. 78–85.

[7] Y. Yang, H.-Y. Ha, F. C. Fleites, and S.-C. Chen, "A multimedia semantic retrieval mobile system based on HCFGs," *IEEE MultiMedia*, vol. 21, no. 1, pp. 36–46, 2014.

[8] Y. Yang, S.-C. Chen, and M.-L. Shyu, "Temporal multiple correspondence analysis for big data mining in soccer videos," in *The First IEEE International Conference on Multimedia Big Data (BigMM)*, 2015, pp. 64–71.

[9] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[10] S.-C. Chen, M.-L. Shyu, C. Zhang, L. Luo, and M. Chen, "Detection of soccer goal shots using joint multimedia features and classification rules," *International Workshop on Multimedia Data Mining (MDM/KDD2003), in conjunction with the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 36–44, 2003.

[11] M. Chen, S.-C. Chen, M.-L. Shyu, and K. Wickramaratna, "Semantic event detection via multimodal data mining," *IEEE Signal Processing Magazine*, vol. 23, no. 2, pp. 38–46, 2006.

[12] L. Lin, C. Chen, M.-L. Shyu, and S.-C. Chen, "Weighted subspace filtering and ranking algorithms for video concept retrieval," *IEEE MultiMedia*, vol. 18, no. 3, pp. 32–43, 2011.

[13] M.-L. Shyu, Z. Xie, M. Chen, and S.-C. Chen, "Video semantic event/concept detection using a subspace-based multimedia data mining framework," *IEEE Transactions on Multimedia*, vol. 10, no. 2, pp. 252–259, 2008.

[14] J. Davis and M. Goadrich, "The relationship between precision-recall and roc curves," in *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 233–240.

[15] G. E. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM Sigkdd Explorations Newsletter*, vol. 6, no. 1, pp. 20–29, 2004.

[16] S. Barua, M. M. Islam, X. Yao, and K. Murase, "Mwmote–majority weighted minority oversampling technique for imbalanced data set learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 2, pp. 405–425, 2014.

[17] C. Elkan, "The foundations of cost-sensitive learning," in *International Joint Conference on Artificial Intelligence*, vol. 17, no. 1. Citeseer, 2001, pp. 973–978.

[18] K. M. Ting, "An instance-weighting method to induce cost-sensitive trees," *IEEE Transactions on Knowledge and Data Engineering*, vol. 14, no. 3, pp. 659–665, 2002.

[19] G. Wu and E. Y. Chang, "KBA: Kernel boundary alignment considering imbalanced data distribution," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 786–795, 2005.

[20] S. Ertekin, J. Huang, and C. L. Giles, "Active learning for class imbalance problem," in *Proceedings of the 30th annual international ACM SIGIR conference on research and development in information retrieval*, 2007, pp. 823–824.

[21] X.-Y. Liu, J. Wu, and Z.-H. Zhou, "Exploratory undersampling for class-imbalance learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 39, no. 2, pp. 539–550, 2009.

[22] A. Estabrooks, T. Jo, and N. Japkowicz, "A multiple resampling method for learning from imbalanced data sets," *Computational Intelligence*, vol. 20, no. 1, pp. 18–36, 2004.

[23] I. Mani and I. Zhang, "KNN approach to unbalanced data distributions: a case study involving information extraction," in *Proceedings of Workshop on Learning from Imbalanced Datasets*, 2003.

[24] T. Jo and N. Japkowicz, "Class imbalances versus small disjuncts," *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 40–49, 2004.

[25] R. C. Prati, G. E. Batista, and M. C. Monard, "Class imbalances versus class overlapping: an analysis of a learning system behavior," in *MICAI 2004: Advances in Artificial Intelligence*. Springer, 2004, pp. 312–321.

[26] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, no. 1, pp. 321–357, 2002.

[27] H. He, Y. Bai, E. A. Garcia, and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning," in *IEEE International Joint Conference on Neural Networks*, 2008, pp. 1322–1328.

[28] D. A. Cieslak and N. V. Chawla, "Start globally, optimize locally, predict globally: Improving performance on imbalanced data," in *IEEE International Conference on Data Mining (ICDM)*, 2008, pp. 143–152.

[29] D. Mease, A. J. Wyner, and A. Buja, "Boosted classification trees and class probability/quantile estimation," *The Journal of Machine Learning Research*, vol. 8, pp. 409–439, 2007.

[30] L. Breiman, "Bagging predictors," *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.

[31] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[32] N. V. Chawla, A. Lazarevic, L. O. Hall, and K. W. Bowyer, "Smoteboost: Improving prediction of the minority class in boosting," in *Knowledge Discovery in Databases (PKDD)*. Springer, 2003, pp. 107–119.

[33] H. Guo and H. L. Viktor, "Learning from imbalanced data sets with boosting and data generation: the databoost-im approach," *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 30–39, 2004.

[34] ——, "Boosting with data generation: Improving the classification of hard to learn examples," in *Innovations in Applied Artificial Intelligence*. Springer, 2004, pp. 1082–1091.

[35] S.-C. Chen, M.-L. Shyu, and C. Zhang, "Innovative shot boundary detection for video indexing," *Video data management and information retrieval*, pp. 217–236, 2005.

[36] S.-C. Chen, M.-L. Shyu, M. Chen, and C. Zhang, "A decision tree-based multimodal data mining framework for soccer goal detection," in *IEEE International Conference on Multimedia and Expo (ICME)*, vol. 1, 2004, pp. 265–268.

[37] L. Lin, M.-L. Shyu, and S.-C. Chen, "Enhancing concept detection by pruning data with mca-based transaction weights," in *IEEE International Symposium on Multimedia (ISM)*, 2009, pp. 304–311.

[38] Y. Yang, H.-Y. Ha, F. Fleites, S.-C. Chen, and S. Luis, "Hierarchical disaster image classification for situation report enhancement," in *IEEE International Conference on Information Reuse and Integration (IRI)*, 2011, pp. 181–186.

[39] S.-C. Chen, M.-L. Shyu, and N. Zhao, "An enhanced query model for soccer video retrieval using temporal relationships," in *IEEE International Conference on Data Engineering (ICDE)*, 2005, pp. 1133–1134.

[40] J. R. Quinlan, *C4.5: programs for machine learning*. Elsevier, 2014.

[41] S.-C. Chen, M.-L. Shyu, C. Zhang, and M. Chen, "A multimodal data mining framework for soccer goal detection based on decision tree logic," *International Journal of Computer Applications in Technology*, vol. 27, no. 4, pp. 312–323, 2006.

[42] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *ACM SIGKDD explorations newsletter*, vol. 11, no. 1, pp. 10–18, 2009.