

Scientific Paper

Doi: <http://dx.doi.org/10.1590/1809-4430-Eng.Agric.v43n2e20220014/2023>

RESEARCH ON IDENTIFICATION AND CLASSIFICATION METHOD OF IMBALANCED DATA SET OF PIG BEHAVIOR

Min Jin¹, Bowen Yang¹, Chunguang Wang^{2*}

^{2*}Corresponding author Inner Mongolia Agricultural University, Hohhot 010018, China.

E-mail: imauwgcg@126.com | ORCID ID: <https://orcid.org/0000-0001-6709-3422>

KEYWORDS

triaxial accelerometer, behavior identification and classification, under-sampling, over-sampling, artificial neural network.

ABSTRACT

To address the problem of the low accuracy and poor robustness of modeling methods for imbalanced data sets of pig behavior identification and classification, the three commonly used re-sampling methods of under-sampling, SMOTE and Borderline-SMOTE are compared, and an adaptive boundary data augmentation algorithm AD-BL-SMOTE is proposed. The activity of the pigs was measured using triaxial accelerometers, which were fixed on the backs of the pigs. A multilayer feed-forward neural network was trained and validated with 21 input features to classify four pig activities: lying, standing, walking, and exploring. The results showed that re-sampling methods are an effective way to improve the performance of pig behavior identification and classification. Moreover, AD-BL-SMOTE could yield greater improvements in classification performance than the other three methods for balancing the training data set. The overall major mean accuracy of lying, standing, walking, and exploring by pigs A, B and C was significantly improved by using AD-BL-SMOTE, reaching 91.8%, 93.0% and 96.0%, respectively.

INTRODUCTION

With the rapid development of the livestock and poultry industry, the traditional breeding model is gradually changing to an intensive, scale and precision model (He et al., 2016). Accurate and quantitative animal behavior detection is the key to precision farming, and animal activity monitors have been shown to be useful in the detection and diagnosis of illness, as well as the potential early prediction of estrus and breeding (Chambers et al., 2021). However, on most farms, a typical weaner-grower-finishing pig may only be briefly inspected once or twice a day as part of a large group, and breeders still mainly rely upon experience to judge whether the pig is behaving abnormally (Bergamini et al., 2021). This method not only takes a lot of time and energy, but also often fails to make an effective diagnosis and early identification of abnormal behavior in pigs due to human negligence, such that some abnormal behaviors are overlooked and found to be serious or irreversible, resulting in illness and even death (Shen et al., 2014). Pig behavior is the external expression of a pig's physical health condition. However, due to pigs' living habits, there is a problem of

imbalanced data sets, where the training set contains significantly fewer samples of one or more class(es) with respect to the other class(es). Machine learning classifiers are traditionally trained to maximize the overall accuracy and are therefore prone to overpredict the majority class if trained on imbalanced data. Consequently, instances of the positive class may be erroneously classified as negative (Esposito et al., 2021). Furthermore, in practical application, minority categories often contain more useful information that is worth exploring. For instance, the time spent by the pigs walking, feeding, drinking, and excreting can reveal their state of health and welfare, which is beneficial for the early detection of abnormal behavior and reducing economic losses (Larsen et al., 2019; Barwick et al., 2018). Therefore, it is very important to solve the problem of imbalanced data sets and improve the identification and classification accuracy of pig behavior.

One of the most common strategies to solve the imbalance problem is re-sampling (Galar et al., 2012). The essence of re-sampling is to construct a 1:1 data set and delete the excess parts of the majority categories (i.e., under-sampling), or the minority categories for bootstrap sampling,

¹ College of Electrical Engineering, ChangZhou Vocational Institute of Mechatronic Technology, ChangZhou, JiangSu 213164, China.

Area Editor: Héilton Pandorfi

Received in: 2-5-2022

Accepted in: 4-27-2023

to increase the number of minority categories until it matches the number of majority categories (i.e., over-sampling) (Dal Pozzolo et al., 2010). In under-sampling, its randomness is uncontrollable, and this will inevitably lead to the loss of some important information which would be helpful for classification when cutting most of the majority categories. However, over-sampling often generates a large number of repeated samples due to the put-back sampling, which is prone to overlap between categories, leading to model overfitting (Homburger et al., 2014; Smith et al., 2016; Abell et al., 2017).

The problem of imbalanced data sets concerns not only the behavior of pigs, but also cattle, equine, sheep and canine behavior (Sakai et al., 2019; Fogarty et al., 2020; Barwick et al., 2020; Carslake et al., 2021; Mao et al., 2021; Chambers et al., 2021). Learning from imbalanced data generally is a challenge for classification algorithms.

In this paper, a wearable pig behavior information acquisition system with a triaxial accelerometer was designed to conduct real-time and continuous monitoring of pigs' four behaviors: lying, standing, walking, and exploring. The objective of the study was to examine the feasibility of utilizing the re-sampling method to balance the data set, and the four behaviors of the pigs were classified and identified based on a BP neural network. The proposed algorithm has widespread practical benefits when used in animal activity monitoring. The results could provide a basis for establishing an abnormal behavior warning system.

MATERIAL AND METHODS

Data source

The experiment was carried out on a pig farm (Figure 1) in Hohhot, Inner Mongolia, China (40°40'26"N, 111°21'46"E) from 8:00 to 18:00 every day between March 10th and April 17th, 2019. Three pigs at different fattening stages (initial weights of 35.8, 62.3 and 92.4 kg, respectively) were monitored. In addition, the pigs' activity was measured using a triaxial accelerometer with a sampling frequency of 20 Hz (SW-J4601V, China), powered with 5 V lithium-ion batteries and controlled by a CC2530F256 controller and ADXL325 chip. The triaxial accelerometer was placed in a waterproof box and tied to the backs of the pigs. This decision was made because initial tests had shown this positioning to have the least impact on the pigs' natural behavior and came with the lowest risk of the box falling off, compared to placing the box on the neck or the leg of the pigs. The installation direction of the triaxial accelerometer is shown in Figure 2.



FIGURE 1. Internal structure of the experimental pigsty.

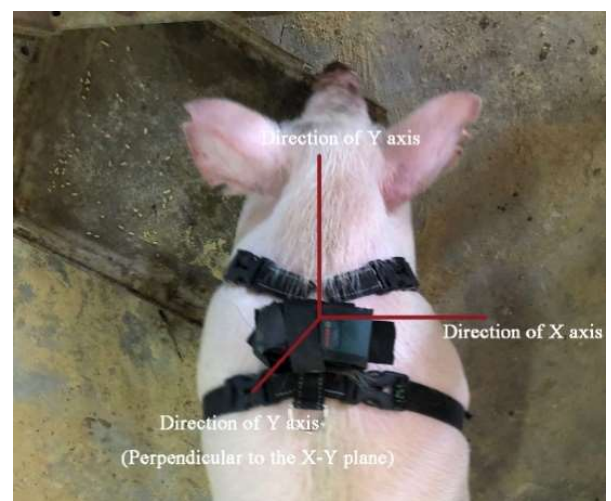


FIGURE 2. Direction of the back-mounted triaxial accelerometer. The X-axis pointed from the left to the right side of the pig's body, the Y-axis pointed from the tail to the head of the pig, and the Z-axis was perpendicular to the XY plane.

The pigs' behavior was video-recorded throughout the experiment, and the camera was time-synchronized to the computer used to initialize the accelerometers. Videos were downloaded and hand-labeled by a single observer to record the exact time and duration of each behavior bout. For this study, we focused on four behaviors of the pigs: lying, standing, walking, and exploring. As these are considered to be the main daily activities of pigs, monitoring these behaviors can provide useful information for abnormal behavior warning and environment control. The definitions and descriptions of these behavioral characteristics of the pigs are summarized in Table 1.

TABLE 1. Behavior ethogram of pigs.

Behavior	Definition and description
Lying	Lying on the side with the shoulder in direct contact with the ground, or lying with the sternum touching the ground with the breast.
Standing	Four feet touching the ground to support its body and without movement, including drinking and excreting.
Walking	A set of slow, rhythmic, symmetrical movements, supported at any moment by alternating steps of two of its four legs.
Exploring	Standing or walking through the pen, sniffing, rooting, sucking, nibbling, chewing, or scratching part of the pen above floor level with its nose.

Data recorded while a pig was transitioning from one behavior to another were removed. Additionally, data recorded during any behavior other than the four behavior categories considered in this study were removed as well. Such behaviors included, e.g., running and rubbing their bodies against the wall.

To reduce the effect on the pigs of wearing sensors, possibly affecting their behavior, the data collection started only after an acclimatization period of 3 days.

Considering that the fixed position of the triaxial accelerometer may be crucial for accurate data collection, the neck, leg and back of the pigs were selected as the fixed positions of the triaxial accelerometer according to their physical and behavioral characteristics. The results showed that when the sensor was fixed on the pig’s back, the stress generated in the pig was the least, and the sensor was not easily affected by behaviors such as lowering, raising, and shaking of the head. The stability of data collection was higher, and the differences between the various behavioral characteristics were more obvious. Consequently, the back of each pig was selected as the fixed position of the triaxial accelerometer in this study.

Data pre-processing

Data processing was done using both R (The R Core Team, 2013) and MATLAB (2017). Modeling and statistical analysis were done in R. Missing values were removed from the time series of the accelerometer data.

Data re-sampling of pig behavior

Data distribution of pig behavior

Relevant studies show that the frequency and duration of various behaviors of livestock and poultry differ. In a day, pigs spend 75%~85% of the time lying, 5%~10% of the time feeding, and the rest of the time walking, standing, and exploring (Li, 2014). As a result, the behavioral data set of pigs is often imbalanced, which has a great impact on the performance of classification learning algorithms. The most direct impact is that most or even all the minority categories are identified as majority categories, which leads to a large increase in the misidentification rate of minority categories, but the overall accuracy is still very high, resulting in unreliable conclusions. The behavioral data statistics of the three experimental pigs in this study are shown in Table 2.

TABLE 2. Behavioral data distribution of three experimental pigs.

Pigs	Lying	Standing	Walking	Exploring
Pig A	12832	1461	594	5810
Pig B	15536	388	444	1092
Pig C	16325	480	424	2757

As can be seen from Table 2, the data size of lying in this study is far more than that of standing, walking, and exploring, and walking is the most minority category. If the machine learning algorithm is directly used for identification and classification, it will inevitably lead to the neglect of the minority categories, and the decision boundary of the classifier is more inclined to the majority categories, leading to the decline of the machine learning performance (Beyan & Fisher, 2015).

SMOTE

The Synthetic Minority Over-Sampling Technique (SMOTE) is an improved scheme based on the random over-sampling method, which uses linear interpolation to create new samples of minority categories on the line between the original minority class and its selected nearest neighbor (Chawla et al., 2009). The effect of the SMOTE on imbalanced data sets is shown in Figures 3 and 4.

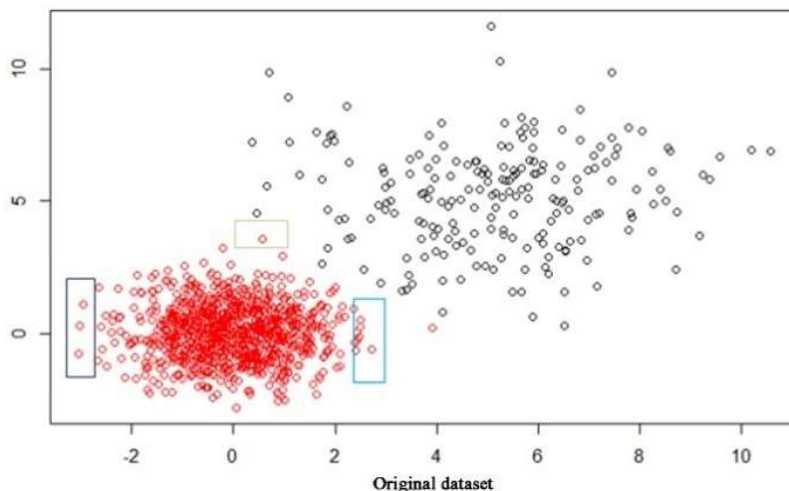


FIGURE 3. Imbalanced data set.

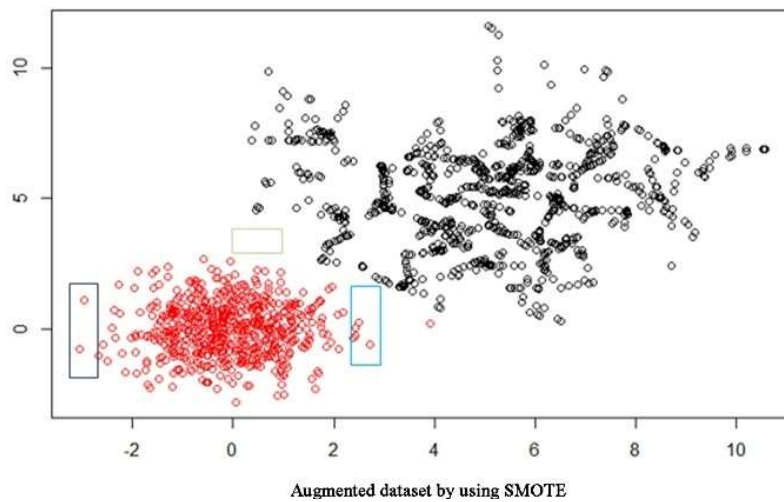


FIGURE 4. Augmented data set by using SMOTE.

As can be seen from Figures 3 and 4, in contrast to simply randomly copying and pasting directly from the majority samples, SMOTE can effectively relieve the over-fitting problem caused by random over-sampling. However, this method has two drawbacks. First, when selecting the nearest neighbor of a minority sample, the influence on the majority sample is not considered, and the process of generating new minority categories can also produce new majority categories, thereby increasing the degree of overlapping, and the contribution of samples far from the boundary to classification is weakened. The second shortcoming is that it does not consider the distribution of various types of data in the original imbalanced data set. The few samples in the boundary and the new samples synthesized by their neighbors are still in the boundary position. As the number of new samples gradually increases, the boundaries of the majority and minority categories will become more and more blurred. Although this can achieve the purpose of

balancing data sets, there is no longer any clear boundary between categories, which increases the difficulty of identification and classification.

Borderline-SMOTE

Borderline-SMOTE is the extension of SMOTE (Han et al., 2005). Compared with non-bounded samples, bounded samples are more likely to be misclassified. Hence, Borderline-SMOTE first finds the minority categories around the classification boundary, which is called the DANGER set, and gives a better indication of the overall distribution of the data set. Then each sample of the DANGER set and its nearest neighbor are linearly interpolated to reduce the overlap between the newly generated sample and the original sample. This algorithm only over-samples the samples of the DANGER set, to avoid the overlapping problem of newly generated samples, as shown in Figures 5 and 6.

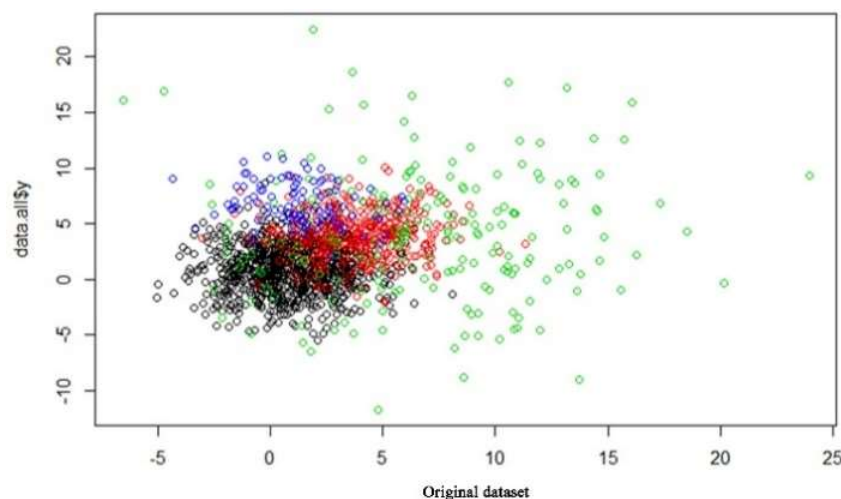


FIGURE 5. Original imbalanced data set.

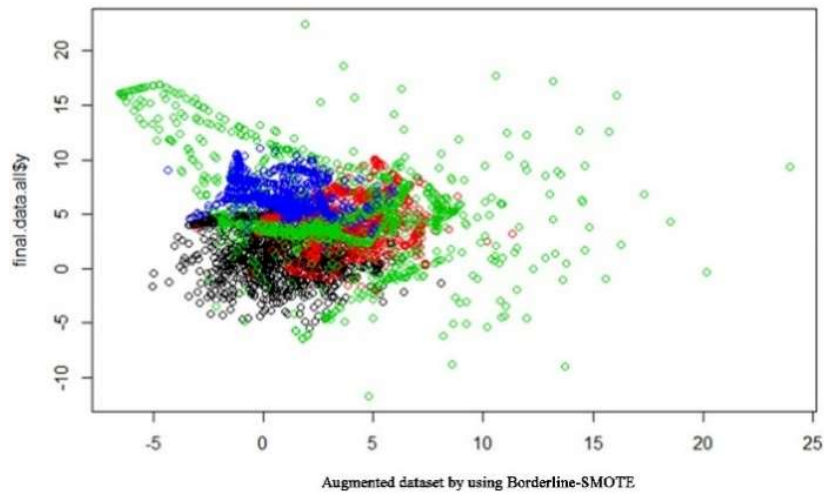


FIGURE 6. Over-sampled data set by using BL-SMOTE.

As shown in Figure 6, using Borderline-SMOTE to create a new data sample, unlike SMOTE (Figure 4), only uses a few categories near the category boundary to generate the new sample, which does not affect the number and distribution of the majority class sample. However, both the SMOTE and Borderline-SMOTE algorithms choose the number of samples that should be constructed for each minority randomly, without considering the difference between the minority samples.

AD-BL-SMOTE

In this study, considering the mean proximity distance of the boundary sample in the minority and the sample number of the majority in the proximity, we put forward an adaptive data borderline synthetic algorithm, AD-BL-SMOTE. The main idea of this algorithm is to strengthen the differentiation of boundary samples which are difficult to classify, to reduce the possibility of their misclassification as much as possible. According to the distribution of the data set and the statistical analysis of the degree of imbalance between the behavior categories, the numbers of the new synthetic samples can be determined. The “Sampling Weight” (w) is first introduced in this paper to measure the number of samples that should be synthesized for boundary samples of each minority category.

The sampling weights are set based on the level of difficulty at which the minority categories can be accurately identified. For the minority categories at the boundary, the samples that are not easily distinguished are usually close to the majority categories or far from other minority categories, so the sampling weight of such samples will be larger, and vice versa. For a sample of minority categories, the more samples of its nearest neighbor there are, the closer the sample is to that class. When the number of majority and minority categories in its k nearest neighbor is the same, the minority sample can be compared with the sum of the distances of the minority category and majority samples in its k nearest neighbor. In this paper, formula (1) is used to calculate the corresponding sampling weight w of boundary samples of minority categories:

$$w = k_1 * \frac{d_1}{k_1 - k_2} \quad (1)$$

where:

d_1 is the sum of the distances between the minority samples and all similar samples in its k nearest neighbor,

k_1 is the number of the nearest neighbors of any minority class samples in the data set, and

k_2 is the number of samples of which their nearest neighbors are majority categories.

The steps of AD-BL-SMOTE are as follows:

(1) T is the original training set, the minority class is N , the sample size is n , the majority class is M , and the sample size is m . The formula of the imbalance degree α of data set T is shown in [eq. (2)]:

$$\alpha = \frac{m}{n} \quad (2)$$

S is the number of new samples to be synthesized in the training set and is shown in [eq. (3)]:

$$S = n * (r - 1) \quad (3)$$

where:

the range of the oversampling rate is $1 \leq r \leq \alpha$.

(2) For each minority class in the training set, if its K nearest neighbors are all majority categories, the sample is classified as a noise sample. If the sample number of the majority class exceeds the minority class in the K nearest neighbors, the sample is classified as a boundary sample. Otherwise, it is a safe sample.

(3) Compute the k nearest neighbor of each boundary sample in a minority set. Then, the number of the new synthetic minority class sample size N_{syn} can be calculated by [eq. (4)]:

$$N_{syn} = \omega * S \quad (4)$$

(4) A new balanced data set is obtained by combining the new synthesized sample of minority categories and the original training set; the new balanced data set is only used as the training set.

Figure 7 shows the distribution of all categories in the data set after using AD-BL-SMOTE. As we can see, most of them are synthesized from minority boundary samples that are difficult to classify, while the number of new samples synthesized from boundary samples that are easy to classify is relatively reduced. The distinct difference between SMOTE and AD-BL-SMOTE is that AD-BL-SMOTE does not affect the distribution or anything else in the majority in the process of generating new samples. Also, it is found that when the data set is large, AD-BL-SMOTE has better CPU efficiency, saving a lot of time and having better robustness, whereas Borderline-SMOTE takes longer and gives a lot of missing values.

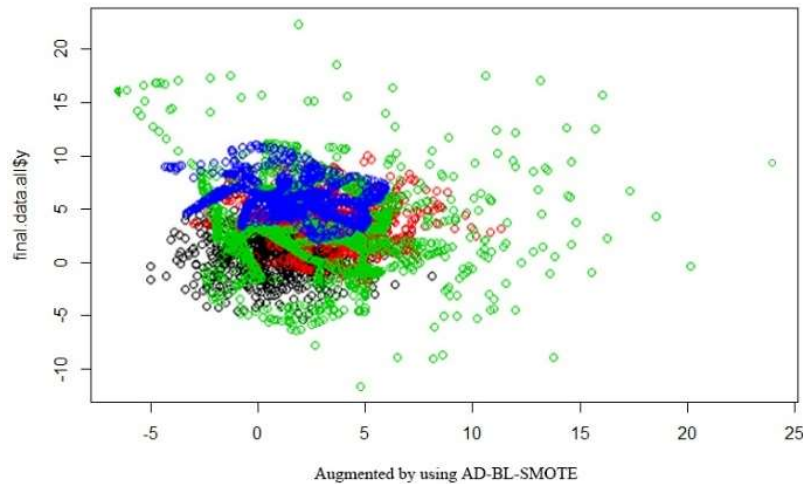


FIGURE 7. Over-sampled data set by using AD-BL-SMOTE.

Identification and classification of pig behavior based on BP neural network

The BP (back-propagation) neural network is a multi-layer feed-forward neural network trained by the error back-propagation algorithm (BP algorithm) (Zhang et al., 2021). Compared with other algorithms, the fully connected feed-forward neural network, as a general function approximation, has a strong learning ability and adaptability, low computational cost, and high computational efficiency (Hou et al., 2018).

Artificial neural network architecture

Fully connected feed-forward ANNs were trained using the back-propagation algorithm, and using the function "mx.model.FeedForward.create" from the R package "mxnet".

The ANNs trained in this study consisted of an input layer, two hidden layers and an output layer. Two hidden layers were chosen as the structure in this study since this is known to be superior to ANNs with only one hidden layer in terms of the number of parameters needed for the training (Meng & Li, 2020). Meanwhile, the number of neurons in the input layer was set as 21, including the values of the three axes (X, Y, Z) and the six moving summary statistics

calculated for each axis. Considering that the number of neurons in the hidden layers is crucial to the overall neural network architecture, too few neurons will not be sufficient to express the complex nonlinear relationship of the system, while too many neurons will lead to over-fitting and result in the decline of the generalizability of the ANN (Bennison et al., 2017). This study optimized the number of nodes in the two hidden layers as follows: for the first hidden layer, we tried using 2/3, 1 and 4/3 times the number of nodes in the input layer. Similarly, for the second hidden layer, we tried using 2/3, 1 and 4/3 times the number of nodes in the first hidden layer (Larsen et al., 2019). The best architecture of the ANN was chosen based on the highest accuracy, as shown in Table 3.

Rectified linear units (ReLU) was used as the activation function in the hidden layers, while the softmax function was used as the activation function in the output layer. The output layer had four nodes, corresponding to the four categories of pig behavior that were considered in this study. The softmax function adjusts the values of the four outputs, so that they are all between 0 and 1 and always sum to 1. Thus, each of the four output values can be interpreted as the probability of the respective behavior. The final prediction for a given observation was the behavior class with the highest probability value.

TABLE 3. Architecture of the ANN.

Structural parameter	Application value
Number of input variables	21
Number of hidden layers	2
Number of output variables	4
Number of hidden layer nodes	28, 28
Learning rate	0.01
Initial weight	-1 to 1
Activation function	ReLU
Output layer transfer function	softmax
Momentum factor	0.9
Maximum number of training steps	120

Model training and evaluation

The ANNs were trained with labeled samples for 120 iterations. In this study, the data of 3 different pigs were first combined, then the whole data set was randomly divided as three parts and three-fold cross validation was used to train and validate the models. Two of the three data sets were combined in turn and used to train a model iteratively, then the model was tested on the remaining data set, respectively.

Accuracy is one of the most used evaluation metrics in classification. The calculation of the accuracy uses the four quantities (*TP*, *TN*, *FP* and *FN*), which give a better summary of the performance of classification algorithms, as defined in [eq. (5)]:

$$ACC = \frac{TP + TN}{TN + TP + FN + FP} \tag{5}$$

where:

TP (True Positives) represents actual positives that are correctly predicted positives;

TN (True Negatives) is actual negatives that are correctly predicted negatives;

FP (False Positives) is actual negatives that are wrongly predicted as positives;

FN (False Negatives) is actual positives that are wrongly predicted as negatives.

In this paper, the main performance metric was the major mean accuracy. For each behavior class, the per-class accuracy was calculated as the observed instances of that class which were correctly predicted to be of that class. The major mean accuracy was then calculated as the simple mean of the four per-class accuracies.

RESULTS AND DISCUSSION

The “mxnet” and “dplyr” packages of R were used to realize the pig behavior identification and classification based on the BP neural network. To assess the usage of the four different re-sampling methods, the results of random under-sampling, SMOTE, Borderline-SMOTE and AD-BL-SMOTE on pig behavior identification and classification were compared.

The behavior data of each experimental pig were repeated and returned for 20 random samples using under-sampling. Each time, the lowest number of categories in the data set is used as a baseline, and the same number of categories as the minority are randomly selected from the other three categories. The 20 newly generated data sets were only used for training the model, and the original imbalanced data set was used for validation, and three-fold cross-validation was carried out. The major mean accuracies of the 20 groups were calculated, and the results are shown in Table 4.

TABLE 4. Classification results based on three over-sampling methods and BP neural networks.

Re-sampling methods	Pigs	Lying	Standing	Walking	Exploring	Major mean accuracy
Without re-sampling	Pig A	75.5%	16.4%	12.6%	19.9%	31.1%
	Pig B	79.5%	18.1%	21.4%	28.5%	36.9%
	Pig C	84.1%	27.1%	28.1%	22.6%	40.5%
	Mean accuracy	79.7%	20.5%	20.7%	23.7%	36.2%
Under-sampling	Pig A	48.6%	15.3%	58.1%	26.6%	37.1%
	Pig B	69.0%	61.6%	15.8%	24.5%	42.9%
	Pig C	70.1%	44.2%	35.6%	20.5%	42.6%
	Mean accuracy	62.6%	40.4%	46.4%	23.9%	40.9%
SMOTE	Pig A	95.0%	72.3%	68.8%	76.5%	78.2%
	Pig B	93.4%	77.5%	76.2%	80.3%	81.9%
	Pig C	92.1%	82.3%	79.8%	82.5%	84.2%
	Mean accuracy	93.5%	77.4%	74.9%	79.8%	81.4%
BL-SMOTE	Pig A	90.2%	79.9%	87.8%	82.6%	85.1%
	Pig B	91.4%	81.7%	79.7%	88.3%	85.3%
	Pig C	93.2%	85.4%	81.2%	88.7%	87.1%
	Mean accuracy	91.6%	82.3%	82.9%	86.5%	85.8%
AD-BL-SMOTE	Pig A	96.8%	90.1%	87.3%	92.9%	91.8%
	Pig B	98.4%	88.6%	90.3%	94.7%	93.0%
	Pig C	99.7%	94.7%	91.8%	97.6%	96.0%
	Mean accuracy	98.3%	91.1%	89.8%	95.1%	93.6%

Firstly, it can be seen from Table 4 that, compared with the results obtained without using any re-sampling methods, the over-sampling method has a significant effect on balancing the training set and thus improves the identification and classification accuracy of pig behavior, especially the minority categories.

Secondly, when the original imbalanced data set was balanced by using under-sampling, the overall major mean accuracy of pig A, pig B and pig C was changed from 31.1% to 37.1%, 36.9% to 42.9%, and 40.5% to 42.6%, respectively, which proves that balancing data sets by the re-sampling method can relieve the problem of the classification performance of the algorithm being biased to the majority categories. Although the accuracies of identification and classification of various behaviors have been slightly

improved, the overall results are still far from ideal, however, which may be related to the reduction of a large amount of data.

Additionally, there are three other over-sampling methods besides random under-sampling; SMOTE, Borderline-SMOTE and AD-BL-SMOTE are used to classify and identify the pig behavior. The major mean accuracy of pig A by using these three over-sampling methods reaches 78.2%, 85.1% and 91.8%, respectively. The major mean accuracy of pig B is 81.9%, 85.3% and 93.0%, respectively. The major mean accuracy of pig C is 84.2%, 87.1% and 96.0%, respectively. Therefore, when using the AD-BL-SMOTE algorithm for the identification and classification of pig behavior, the overall performance is significantly improved, which proves that this method is an effective way to improve the identification and classification of pig behavior.

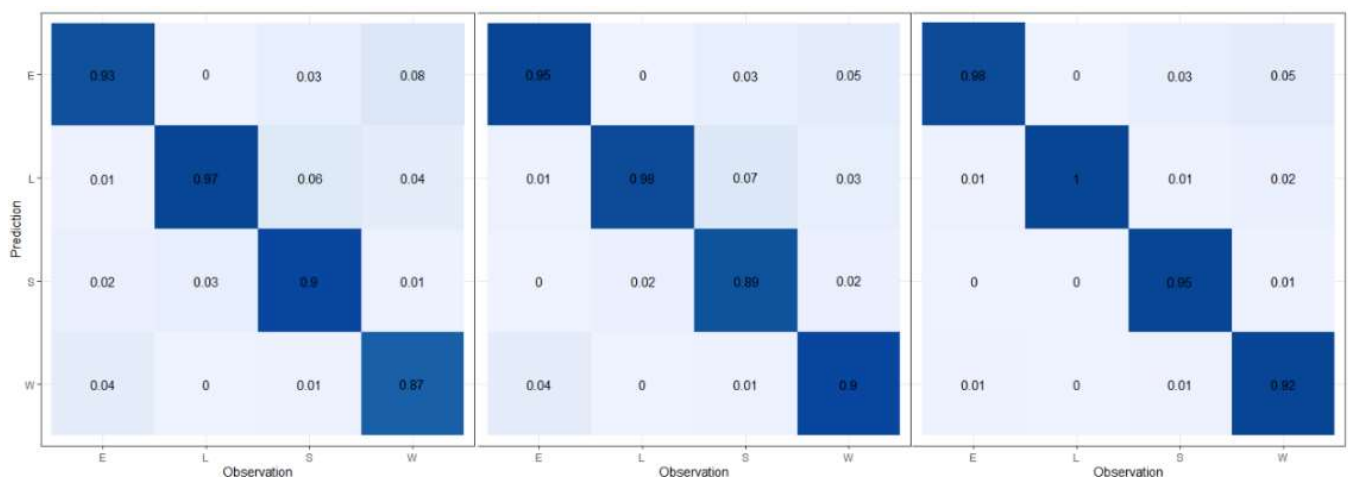


FIGURE 8. Behavior classification results of pigs A, B and C after using AD-BL-SMOTE.

As shown in Figure 8, for all three experimental pigs, lying, standing, and walking are easy to confuse with exploring, and exploring is often misclassified as standing and walking, which may be related to the motion amplitude of pigs. When the pig is standing but with its head slightly sniffing or rubbing against the wall, exploring, and standing are easily confused because the sensor is fixed on the pig's back. When the pig remained motionless but its head moved violently, the exploring was easily misidentified as walking, and vice versa. In addition, lying was often misidentified as standing, since both behaviors are static in nature and have similar behavior patterns. Meanwhile, walking behavior consists of semi-regular, repetitive steps at regular intervals. When pig walking and standing, walking, and exploring occurs repeatedly, considering that the three-axis accelerometer itself has a certain size and weight, and the pig's back is not completely flat, when the pig is in a state of lying or standing, breathing and body shaking will produce acceleration data, which also raises the possibility of misclassifying the pig behavior. To solve this problem, further research will consider adding the transition state between the two kinds of behaviors into the analysis. The enrichment of the data sets and data types may help to improve the learning performance of the classifier.

CONCLUSIONS

Based on the degree of imbalance of the pig behavior data set and the deficiency of the two over-sampling methods (SMOTE, Border-line SMOTE), this paper presents the AD-BL-SMOTE algorithm to classify and identify pig behavior. Re-sampling methods, and especially over-sampling methods, have been proven to yield accurate classification accuracy over a range of pig behaviors using triaxial-accelerometer data from a back-mounted device. The effect of using AD-BL-SMOTE is more pronounced than balancing the training data by SMOTE and Borderline-SMOTE. The overall performance is consistently and significantly improved, which proves that this method is an effective way to improve the identification and classification of pig behavior. The results could provide technical support for further improving the welfare of pigs and aiding pig farms in making management decisions.

ACKNOWLEDGMENTS

This research is supported by the 12th Five-year National Science and Technology Support Program of China [2014BAD08B05] and Inner Mongolia Autonomous Region Graduate Student Scientific Research Innovation Projects [B2018111948].

REFERENCES

- Abell KM, Theurer ME, Larson RL, White BJ, Hardin DK, Randle RF (2017) Predicting bull behavior events in a multiple-sire pasture with video analysis, accelerometers, and classification algorithms. *Computers and Electronics in Agriculture* 136:221-227.
- Barwick J, Lamb DW, Dobos R, Schneider D, Welch M, Trotter M (2018) Predicting lameness in sheep activity using tri-axial acceleration signals. *Animals* 8:1-16.
- Barwick J, Lamb DW, Dobos R, Welch M, Schneider D, Trotter M (2020) Identifying sheep activity from tri-axial acceleration signals using a moving window classification model. *Remote Sensing* 12:1-13.
- Bennison A, Bearhop S, Bodey TW, Votier SC, Grecian WJ, Wakefield ED, Hamer KC, Jessopp M (2017) Search and foraging behaviors from movement data: a comparison of methods. *Ecology and Evolution* 11:1-13.
- Bergamini L, Pini S, Simoni A, Vezzani R, Calderara S, D'Eath RB, Fisher RB (2021) Extracting accurate long-term behavior changes from a large pig data set. *VISIGRAPP 2021*. In: *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. Setubal, Science and Technology Publications, Proceedings... v5, p524-533. 5:524-533.
- Beyan C, Fisher R (2015) Classifying imbalanced data sets using similarity based hierarchical decomposition. *Pattern Recognition* 48:1653-1672.
- Carslake C, Vázquez-Diosdado AJ, Vázquez-Diosdado J (2021) Machine learning algorithms to classify and quantify multiple behaviours in dairy calves using a sensor: moving beyond classification in precision livestock. *Sensors* 21:1-14.
- Chambers RD, Yoder NC, Carson AB, Junge C, Allen DE, Prescott LM, Bradley S, Wymore G, Lloyd K, Lyle S (2021) Deep learning classification of canine behavior using a single collar-mounted accelerometer: Real-world validation. *Animals* 11:1-19.
- Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2009) SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* 16:321-357.
- Dal Pozzolo A, Caelen O, Bontempi G (2010) Comparison of balancing techniques for unbalanced data sets. *Machine Learning Group, Université Libre de Bruxelles, Belgium* 16:732-735.
- Esposito C, Landrum GA, Schneider N, Stief S, Riniker S (2021) GHOST: Adjusting the decision threshold to handle imbalanced data in machine learning. *Journal of Chemical Information and Modeling* 61:2623-2640.
- Fogarty ES, Swain DL, Cronin GM, Moraes LE, Trotter M (2020) Behaviour classification of extensively grazed sheep using machine learning. *Computers and Electronics in Agriculture* 169:1-10.
- Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F (2012) A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* 42: 463-484.
- Han H, Wang W, Mao B (2005) Borderline-SMOTE: a new over-sampling method. *International Conference on Intelligent Computing* 878-887.
- He DJ, Liu D, Zhao KX (2016) Review of perceiving animal information and behavior in precision livestock farming. *Transactions of the Chinese Society for Agricultural Machinery* 47: 231-244.

Homburger H, Schneider MK, Hilfiker S, Lüscher A (2014) Inferring behavioral states of grazing livestock from high-frequency position data alone. *PLoS One* 9:1-22.

Hou YT, Cai XH, Wu ZQ, Dong ZG (2018) Research and implementation of cattle behavior character recognition method-based on support vector machine. *Journal of Agricultural Mechanization Research* 8:36-41.

Larsen MLV, Bertelsen M, Pedersen LJ (2019) Pen fouling in finisher pigs: Changes in the lying pattern and pen temperature prior to fouling. *Frontiers in Veterinary Science* 6:1-6.

Li Y (2014) Normal and abnormal behaviors of swine under production conditions. Available:

<https://thepigsite.com/articles/normal-and-abnormal-behaviours-of-swine-under-production-conditions>

Mao A, Huang ED, Gan HM, Parkes RSV, Xu WT, Liu K (2021) Cross-modality interaction network for equine activity recognition using imbalanced multi-modal data. *Sensors* 21:1-17.

MATLAB (2017) Matrix&Laboratory. MathWorks, 1984. Available: <https://www.mathworks.com/products/matlab.html>

Meng ZL, Li TL (2020) Review and prospect of machine learning technology. *Application of IC* 37:56-57.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available: <https://www.r-project.org/>

Sakai K, Oishi K, Miwa M, Kumagai H, Hirooka H (2019) Behavior classification of goats using 9-axis multi sensors: the effect of imbalanced data sets on classification performance. *Computers and Electronics in Agriculture* 166:105027.

Shen MX, Liu LS, Yan L, Lu MZ, Yao W, Yang XJ (2014) Review of monitoring technology for animal individual in animal husbandry. *Transactions of the Chinese Society for Agricultural Machinery* 45: 245-251.

Smith D, Rahman A, Bishop-Hurley GJ, Hills J, Shahriar S, Henry D, Rawnsley R (2016) Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems. *Computers and Electronics in Agriculture* 131:40-50.

Zhang C, Guo Y, Li M (2021) Review of development and application of artificial neural network models. *Computer Engineering and Application* 57:57-69.